

<<语义网数据管理技术及应用>>

图书基本信息

书名：<<语义网数据管理技术及应用>>

13位ISBN编号：9787030264688

10位ISBN编号：7030264681

出版时间：2010-2

出版时间：科学出版社

作者：金海，袁平鹏 著

页数：304

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

## <<语义网数据管理技术及应用>>

### 前言

国家重点基础研究发展计划项目“语义网格的基础理论、模型与方法研究”从2003年开始围绕“规范重构，语义互联，智能聚融”这三个科学问题系统地探索了未来互联环境的语义基础问题，取得了一系列成果，本书是其中的重要成果之一。

自从计算机诞生以来，数据管理一直是计算机科学的核心问题之一。特别是随着信息技术的快速发展与普及，个人与社会每时每刻都在大量产生数据。人类将生活在一个充满数据的世界里。

如何对海量数据进行有效管理就成为一个重要的研究问题。

要实现海量数据的有效管理首先要解决语义问题。

有效利用海量数据更加依赖语义问题的解决。

计算机科学乃至信息科学的许多问题在本质上都可以归结为语义问题。

虽然人们自觉不自觉地都在使用语义，但不同的领域对语义有不同的认识。

例如，语义网研究领域希望对万维网中的数据增加机器可理解的语义，以使机器能快速、准确地从万维网上获取全面有用的信息并进行自动处理。

但大量加上标注的数据如何进行管理就成为了一个新的问题。

传统的数据库模型难以适应如今的网络数据管理，从而研究规范组织的语义网数据模型就成为问题的核心。

该书结合作者的研究体会，吸收了新的技术和方法，大跨度地阐述了语义数据管理的技术和研究问题，涉及诸如数据模型、抽取、分类、存储、检索、呈现等。

适应了语义网数据管理开发、应用和研究的需要。

该书内容丰富，希望读者能从中获得对语义数据管理的知识，对相关研究有所启示。故向读者郑重推荐。

## <<语义网数据管理技术及应用>>

### 内容概要

万维网已经深刻影响到社会生活的各个方面。

《语义网数据管理技术及应用》围绕Web科学的核心问题之一——数据管理技术进行阐述。

首先，介绍了语义网的意义及语义网的标准规范；然后，研究了语义数据模型及存储技术，以实现语义数据的高效存储、快速查询和动态添加。

由于现实中的数据通常为半结构化或非结构化数据，为了从中获取语义数据，《语义网数据管理技术及应用》介绍了信息抽取技术、数据分类技术。

在此基础上，介绍了语义数据检索、语义数据可视化技术，以便为用户提供高质量的检索结果，并向用户可视化地呈现丰富的数据间语义关联。

最后，介绍了一些语义网应用系统。

《语义网数据管理技术及应用》可供计算机科学与技术、模式识别等相关专业的研究人员、工程技术人员、教师、研究生和本科生学习参考。

## <<语义网数据管理技术及应用>>

### 作者简介

袁平鹏，博士，副教授，中国计算机学会高级会员，1972年生。2002年于浙江大学获得工学博士学位。担任多个国内外期刊与学术会议审稿人。湖北省创新群体主要成员之一。主要从事语义网、分布式系统等方面的研究。在国内外发表学术论文40余篇。其中被SCI、EI等收录20余篇次。获得国

## &lt;&lt;语义网数据管理技术及应用&gt;&gt;

## 书籍目录

序前言	第一章 绪论	1.1 语义网及其面临的挑战	1.2 与web1.0和Web2.0的区别	1.3 语义网体系结构
	1.4 语义网规范	1.4.1 资源描述框架	1.4.2 RDFS / OWL	1.4.3 SPARQL
		1.4.5 RDFa	1.5 本章小结	参考文献
	第二章 语义数据模型	2.1 数据模型简介	2.1.1 语义及数据模型	2.1.2 语义数据模型特征
		2.1.3 语义数据模型分类	2.2 逻辑及物理数据模型	2.2.1 层次和网状数据模型
		2.2.2 关系数据模型	2.2.3 面向对象的数据模型	2.2.4 模糊语义模型
		2.2.5 图模型	2.3 概念数据模型	2.3.1 实体-关系模型
		2.3.2 RM / T	2.3.3 SDM	2.3.4 资源空间模型及语义链网络
	2.4 主动语义数据模型	2.4.1 主动语义数据模型概述	2.4.2 事件模型研究现状	2.4.3 FEM
		2.4.4 事件探测	2.5 本章小结	参考文献
	第三章 语义数据抽取	3.1 信息抽取技术简介	3.1.1 国内外研究现状	3.1.2 信息抽取技术分类
		3.2 科技文献元数据抽取	3.2.1 科技文献特征	3.2.2 信息抽取预处理
		3.2.3 基于模板匹配的头部信息抽取	3.2.4 基于统计的尾部信息抽取	3.2.5 信息封装
	3.3 科技文献信息抽取系统及其性能评价	3.3.1 科技文献信息抽取系统	3.3.2 评价指标	3.3.3 实验结果及分析
	3.4 语义数据清理	3.4.1 语义数据清理概述	3.4.2 歧义实体的识别	3.4.3 重复数据的发现
	3.5 本章小结	参考文献	第四章 数据分类	4.1 数据分类意义
			4.2 研究现状	4.2.1 传统文本分类
			4.2.2 层次分类	4.2.3 基于知识的分类
			4.3 多类别文本自动分类方法	4.3.1 支持向量机分类方法
			4.3.2 k近邻分类方法	4.3.3 多类别单标签分类方法
			4.3.4 多类别多标签分类方法	4.4 多类别文本分类实现技术
			4.4.1 文献预处理	4.4.2 特征降维
			4.4.3 权重计算	4.4.4 分类器设计
			4.5 多类别分类方法性能评价	4.6 本章小结
	参考文献	第五章 语义网数据存储技术	5.1 语义网数据存储技术现状	5.1.1 RDF数据模型
		5.1.2 RDF存储结构	5.1.3 RDF查询语言	5.1.4 语义数据索引技术
		5.1.5 语义网数据存储系统	5.2 语义网数据管理系统DBLink	5.2.1 DBLink的体系结构
		5.2.2 基于主存的面向对象存储模型	5.2.3 URI映射与内置数据类型	5.2.4 语义网数据分割与映射
		5.2.5 语义网数据的压缩存储	5.2.6 模式空间与实例空间分离	5.3 DBLink性能评价
		5.3.1 测试数据集	5.3.2 实验环境	5.3.3 测试查询集
		5.3.4 实验结果	5.4 本章小结	参考文献
	第六章 语义数据检索技术	6.1 语义数据检索的分类	6.1.1 传统数据 / 信息检索	6.1.2 语义网数据检索
		6.1.3 语义网格 / 知识网格中的数据检索	6.1.4 对等网络中的信息检索	6.2 语义关联数据模型及其检索的关键问题
		6.2.1 支持知识评价的语义关联数据模型	6.2.2 实用且有效的语义数据检索机制	6.2.3 模糊语义数据的表达及检索
		6.2.4 基于语义的对等网数据组织及其检索	6.3 支持知识评价的语义关联数据模型	6.3.1 语义关联数据模型
		6.3.2 数据的检索及回答	6.4 基于关键字的语义数据检索	6.4.1 问题描述
		6.4.2 近似搜索算法及分析	6.4.3 相关问题讨论	6.5 基于语义模糊性和用户偏好的检索机制
		6.5.1 后台知识库的语义模糊性	6.5.2 用户检索请求的规范化	6.5.3 检索结果的排序
		6.6 基于语义小世界的的数据检索	6.6.1 语义小世界的构建	6.6.2 数据检索算法
		6.7 本章小结	参考文献	第七章 语义数据检索可视化
		第八章 语义网数据应用系统	参考文献	

## <<语义网数据管理技术及应用>>

### 章节摘录

万维网已经成为人们获取信息的主要渠道之一，深刻影响到人类社会生活的各个方面：人们在Web上浏览新闻、搜索信息、买卖商品及服务。然而，目前的万维网是面向人而不是面向机器的。换句话说，人可以理解万维网上的内容，而机器则不能理解。随着万维网上的内容越来越多（据Google的官方Blog公布的数据，目前Google索引了IT个URL），人们准确、快速、全面获取到信息也越来越难。

对此，Www、URIs、HTTP和TMI，的发明人Berners-Lee提出了语义网概念。他认为“ The Semantic Web will bring structure to the meaningful content of web pages. creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users ”自从提出语义网之后，语义网得到了学术界和工业界的重视。Gartner认为企业语义网是3个具有影响力的万维网技术之一。许多语义网公司，如Siderean、SandPiper、Intellidimension等在兴起。一些大公司，如Adobe、HP、IBM、Nokia、Oracle、Vodafone等纷纷对语义网技术进行研发。一些国家和国际组织，如英国、欧盟、美国也投入资金资助相关研究。

语义网是关于万维网的元数据网。数据一直是语义网的核心问题，Berners-Lee曾经公开表示，他认为对语义网更准确的名称应该是数据网（data web）。语义网存在及互联的目的是使得计算机能够自动处理、集成来自不同数据源的数据，而不仅仅是为了让计算机能够正确显示。为了实现让机器或设备能够自动识别和处理网上数据，需要在内容中加入标记，即需要采用标记语言。考虑到适应不同领域的标记需要，标记语言必须是灵活、可扩展的。为了实现此目标，研究和制定在万维网上发布能够帮助计算机理解内容的设施、技术和规范成为迫切需要解决的问题。

对此，研究人员、标准化组织包括互联网工程任务组和W3C等都致力于增强、扩展语义网能力，在研究、制定以及部署能够共享语义的语言以及工具方面进行了大量工作。这些语言为语义互操作提供了基础。1997年，W3C定义了资源描述框架规范RDF。此后，相继提出了RDFS / OWL、GRDDL、RDFa等规范或建议规范。研究人员也开发了许多工具，如Kowari、RDFLib、Jena、Sesame、Protege、SWOOP等。W3C将其中的部分技术标准化，形成了一些标准，包括资源描述框架RDF、RDFS / OWL等。

<<语义网数据管理技术及应用>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>