

<<多语言领域本体学习研究>>

图书基本信息

书名：<<多语言领域本体学习研究>>

13位ISBN编号：9787305106705

10位ISBN编号：7305106704

出版时间：2012-10

出版时间：南京大学出版社

作者：章成志

页数：206

字数：322000

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>

<<多语言领域本体学习研究>>

内容概要

《多语言领域本体学习研究/信息管理专业前沿论丛》编著者章程成志。

《多语言领域本体学习研究/信息管理专业前沿论丛》内容提要：多语言领域本体是一种解决瓦联网信息资源语义化和多语言化需求问题的重要资源，在跨语言信息检索、机器翻译等多语言科技信息服务中具有重要作用。

本书介绍当前国内外关于多语言本体学习方法、工具以及应用项目等的相关动态。

本书围绕多语言领域本体学习中的两个关键问题，即：双语术语抽取与概念层次体系构建问题，进行了深入研究。

该书研究内容主要包括：基于领域平行语料抽取的双语核心术语抽取研究、基于多层特征的一体化策略术语抽取研究、基于术语度约束的双语术语对齐研究、基于术语聚类的概念层次体系生成研究、基于多语文本聚类的主题层次体系生成研究。

本书可为图书馆学、情报学、计算机科学与技术，信息管理和信息系统等专业的研究生和高年级本科生，以及从事语义网、文本挖掘、知识组织、数字图书馆等方向的科研人员，提供教学参考和技术指导。

<<多语言领域本体学习研究>>

书籍目录

第1章 引言

1.1 研究背景

1.2 研究内容概况

1.2.1 课题概述

1.2.2 多语言本体学习框架与开发平台的设计

1.3 本书的内容与章节安排

参考文献

第2章 多语言领域本体学习研究综述

2.1 多语言词汇资源概述

2.2 多语言本体学习研究现状

2.2.1 本体学习综述

2.2.2 典型的本体学习工具概述

2.2.3 多语本体构建相关研究概述

2.3 多语言本体应用项目概述

2.3.1 基本原理

2.3.2 相关应用项目概述

2.4 本章小结

参考文献

第3章 基于领域平行语料的双语核心术语抽取研究

3.1 相关研究概述

3.2 基于专业领域平行语料的双语核心术语抽取

3.2.1 总体流程与关键技术

3.2.2 抽取结果与分析

3.2.3 模块运行界面

3.3 基于领域分类知识库的术语分类

3.3.1 术语分类概述

3.3.2 基于领域分类知识库的术语分类

3.4 本章小结

参考文献

第4章 基于多层术语度的一体化术语抽取研究

4.1 相关研究概述

4.2 基于多层术语度的一体化术语抽取

4.2.1 一体化术语抽取策略

4.2.2 条件随机场模型概述

4.2.3 多层术语度特征度量

4.2.4 术语抽取结果与分析

4.2.5 模块运行界面

4.3 本章小结

参考文献

第5章 基于术语度约束的双语术语对齐研究

5.1 相关研究概述

5.2 基于术语度约束的双语术语对齐

5.2.1 基本原理

5.2.2 基于术语度约束的词对齐结果优化

5.2.3 基于术语度约束的双语术语对齐

<<多语言领域本体学习研究>>

5.2.4 模块运行界面

5.3 本章小结

参考文献

第6章 基于多语术语聚类的概念层次体系生成研究

6.1 相关研究概述

6.2 基于多语术语聚类的概念层次体系生成

6.2.1 总体流程与关键技术

6.2.2 实验结果与分析

6.2.3 模块运行界面

6.3 本章小结

参考文献

第7章 基于多语文本聚类的主题层次体系生成研究

7.1 相关研究概述

7.2 基于多语言文本聚类的主题式概念层次生成

7.2.1 主题式概念层次生成基本原理

7.2.2 基于多语言文本聚类的主题式层次体系生成策略

7.2.3 实验结果与分析

7.2.4 模块运行界面 ”

7.3 本章小结

参考文献

第8章 结束语

8.1 总结

8.2 进一步工作

8.2.1 双语术语抽取方面

8.2.2 多语言层次关系生成方面

8.2.3 概念间其他关系的发现

附录

附录1 Segtag汉语文本词性标注标记集

附录2 Penn Treebank词性标注标记集

附录3 核心术语训练阶段类间分布熵计算结果样例

附录4 候选核心术语抽取样例(法律类Top—40)

附录5 双语候选核心术语对齐样例(法律类Top—40)

附录6 《人民日报》(1998年上半年)语料词频排序(Top—40)

附录7 NTCIR新闻语料词频排序(Top—40)

附录8 新闻语句样例中的术语度

附录9 专业领域语句样例中的术语度

附录10 术语抽取模型训练标注语料片段

附录11 用于术语抽取的CRF++特征模板

附录12 术语抽取10折交叉验证结果

附录13 Giza++结果中术语度比值排序(Top—40)

附录14 IT领域1万句对的双语术语抽取结果(Top—40)

附录15 基于术语聚类的多语言概念层次体系示例

附录16 基于文本聚类的主题层次体系示例

附录17 语料库与开源工具列表

后记

索引

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>