

<<基于知识元的知识发现>>

图书基本信息

书名：<<基于知识元的知识发现>>

13位ISBN编号：9787560625683

10位ISBN编号：7560625681

出版时间：2011-5

出版时间：西安电子科技大学出版社

作者：温有奎，焦玉英 著

页数：191

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>

<<基于知识元的知识发现>>

内容概要

本书从代表未来知识发现模式——知识元语义链接的角度论述文献知识元间隐含关联的潜在知识发现的基本概念、方法和技术。

全书共9章。

第1章论述了潜在新知识发现问题的提出；第2章给出了语义文本知识挖掘的进展；第3章介绍了文本知识元模型的演变；第4章讨论了知识元模型的数学理论基础，建立了知识科学方程组，引入范畴论表示知识元的语义三角形模型；第5章讨论了文献知识元的分布与提取模型；第6章研究了数字化文本知识元的映射模型；第7章建立了文本主题成因的知识元本体转换模型；第8章以电子病历为例介绍了文本知识元的挖掘应用案例；第9章为结论与展望。

本书既是一本专著，也可以作为高等学校信息管理、知识管理、教育学、机械设计、网页挖掘、计算机应用、图书情报与档案及其相关专业的研究生和本科高年级教材，还可供从事知识系统教学、研究、开发和应用的科技工作者参考。

<<基于知识元的知识发现>>

作者简介

温有奎，男，1951年12月出生，博士，西安电子科技大学经济管理学院教授。

研究方向：语义文本挖掘和知识管理。

校学术带头人，中国科技情报学会理事。

主持国家自然科学基金项目、国家社会科学基金项目、教育部人文社会科学基金项目、陕西省科学技术研究发展计划项目等9项。

出版著作7本，发表期刊论文70余篇(EI检索2篇)。

获“陕西省哲学社会科学优秀成果奖”三等奖2项(第一完成人)。

获中华人民共和国实用新型专利1项(专利号：ZL93212829.7)。

焦玉英，女，1942年出生。

武汉大学信息管理学院教授，博士生导师。

主要从事信息检索理论与实践及信息服务系统建设、企业管理咨询理论与实践、企业竞争情报与服务等方向的教学与科学研究。

主持过国家教委、国家社会科学基金、国家自然科学基金等级别科研项目9项；发表国内外学术论文100余篇；出版教材、专著11部；参编专著、教材20余部。

<<基于知识元的知识发现>>

书籍目录

第1章 绪论

1.1 知识元问题的提出

1.1.1 网络资源构建需要关注知识元

1.1.2 文献单元向知识单元深化需要探讨知识元

1.1.3 知识计量的开拓离不开知识元的探讨

1.1.4 语义Web知识发现需要研究知识元

1.2 知识元研究的意义

1.3 国内外文本知识元研究现状分析

1.3.1 国内外文本知识元研究成果统计

1.3.2 国内外文本知识元研究领域划分

1.4 国内外语义文本挖掘研究工作综述

1.4.1 国内外语义文本挖掘工作成果统计

1.4.2 国内外语义文本挖掘研究领域划分

1.5 国内外知识元研究中存在的问题与不足

1.6 本书的主要内容与创新之处

1.6.1 本书的主要内容

1.6.2 本书的创新之处

第2章 语义文本知识挖掘概述

2.1 文本知识挖掘的进展

2.1.1 文本挖掘的目标

2.1.2 文本挖掘的关键技术与过程

2.1.3 文本挖掘与相关概念的比较

2.2 传统文本挖掘模型

2.2.1 文本向量空间挖掘模型

2.2.2 潜在语义文本挖掘模型

2.3 语义本体文本挖掘模型

2.3.1 语义文本挖掘

2.3.2 文本数据挖掘与文本知识挖掘的区别

2.3.3 基于领域本体的文本挖掘

第3章 文本知识元模型的演变

3.1 知识元的认知模型

3.1.1 知识元的“数据元”模型

.....

第4章 知识元模型的数学理论基础

第5章 文献知识元的特征分布与提取模型

第6章 文本知识元的语义映射模型

第7章 主题成因的知识元本体转换模型设计

第8章 期末知识元的文本模型与挖掘实例

第9章 总结与展望

参考文献

后记

<<基于知识元的知识发现>>

章节摘录

版权页：插图：4.文本挖掘的层次从非结构化的数据中提取可以表达成结构化数据的知识，其难度是相当高的。

但从另一方面看，需要提取的知识本身又是分层次的。

因此，大体上可以把文本挖掘分为元数据、特征和模板三个层次。

元数据：顾名思义，是“关于数据的数据”。

在面向知识网格中描述知识的一种方法是采用元数据。

从简化的角度看，元数据包括了用于注释内容的描述性声明。

元数据的目标是机器对数据的处理和声明。

在数据呈现非结构化形态的情况下，元数据是帮助人们从最基本的角度把握非结构化数据内容的最直接的线索。

常见的元数据有标题、作者、发表时间、体裁、分类、关键词、摘要等公共“字段”，通过这些字段，把有关的线索组织成结构化的数据。

元数据提取可看成是文本组织的结构化过程，元数据层是文本挖掘的最低层次。

特征：文本和多媒体数据的特征，是反映内容属性的量化指标，一般用向量的形式表示，当然向量也是结构化数据的一种特殊形式。

因此，特征提取可看成是文本内容的结构化过程，特征层是文本挖掘的中间层次。

知识模板：它是在语义层面表示数据内容的核心部分的一种结构化表示。

一般的知识模板采用类似于“框架”的表述方式。

框架本身是结构化的，但要通过深层次的理解来获得并填充模板中相应的“槽”。

因此，知识模板生成可看成是文本内容知识的结构化过程，知识模板是文本挖掘的更高层次。

5.文本挖掘的特点 谌志群等对文本挖掘模型做了归纳，其特点表现如下： 文本挖掘处理的是大规模的文本集合，而不是一个或少量的文本文档。

文本挖掘发现的知识是隐藏在大量文本文档中的，是新的、以前未知的模式或关系。

文本挖掘抽取的知识是以真实世界为基础的，具有潜在价值，是直接可用的，它或者是某个特定用户感兴趣的，或者是对于解答某个特定问题有用的。

<<基于知识元的知识发现>>

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>