

基金项目

基于Ontology的面向主题的网络信息采集算法

刘 军<sup>1</sup>;凌云<sup>1</sup>;王 勋<sup>1,2</sup>

浙江工商大学计算机与信息工程学院<sup>1</sup>

收稿日期 2005-8-25 修回日期 网络版发布日期 2006-5-15 接受日期

摘要 介绍基于内容评价的、基于链接结构评价的和基于巩固学习的三种采集算法的优缺点;介绍一种依据词典构建主题Ontology的方法,该方法有助于提高Ontology的构建速度;最后,在分析传统采集算法的基础上,提出一种新的基于Ontology的面向主题的网页采集算法,并通过试验证明其优越性。

关键词 [网页采集](#) [Ontology](#) [rdfs](#)

分类号 [G352](#)

DOI:

对应的英文版文章: [2005-2398](#)

通讯作者:

刘 军<sup>1</sup> [qixingzhanzi@163.com](mailto:qixingzhanzi@163.com)

作者个人主页: 刘 军<sup>1</sup> 凌云<sup>1</sup> 王 勋

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF](#) (OKB)

▶ [\[HTML全文\]](#) (OKB)

▶ [参考文献\[PDF\]](#)

▶ [参考文献](#)

服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [引用本文](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

相关信息

▶ [本刊中 包含“网页采集”的 相关文章](#)

▶ [本文作者相关文章](#)

- [刘军](#)
- [凌云](#)
- [王勋](#)
-