

知识组织与知识管理

信息集成中的字符串匹配技术研究*

孙海霞¹; 成颖¹

南京大学信息管理系¹

收稿日期 2007-5-31 修回日期 2007-6-6 网络版发布日期 2007-7-26 接受日期

摘要 匹配是信息集成的核心技术之一。论述基于编辑距离、基于标记以及N元文法等为代表的字符串匹配技术的研究现状, 指出其存在的不足并提出改进思路。

Abstract Matching is one of the most important techniques of information integration. In this paper, string-based matching algorithms, mainly distance-based, token-based and the N-gram are elucidated. The deficiencies and research directions are also outlined.

关键词 [匹配](#) [信息集成](#) [字符串匹配](#)

Key words Matching; Information integration; String-based matching

分类号 [TP393](#)

DOI:

通讯作者:

孙海霞 sunqiqin1984@yahoo.com.cn

作者个人主页: 孙海霞 成颖

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF \(379KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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“匹配”的 相关文章](#)
- ▶ 本文作者相关文章
 - [孙海霞](#)
 - [成颖](#)