

新谷 虎松
名古屋工業大学 教授

応答文章に着目したデータクリーニング手法

本研究では、コールセンターにおける質問応答文書を応答文書に含まれる自立語数によって文書集合を分割することで、データクリーニングを行い、より意味のあるパターンを発見する手法を提案した。本提案手法は、応答文書に含まれる自立語数に応じて2分割したそれぞれの文書集合に対して、応答文書と問い合わせ文書との関連を考慮することで、問い合わせの意図を表していると思われる重要なパターンを抽出することができる。また、本提案手法の応用としてFAQ作成支援システムの実装を行った。

研究成果の発表

A Paper Recommendation Mechanism for the Research Support System Papis
In the Proc. of International Workshop on Data Engineering Issues
in Electronic Commerce (DEEC2005) pp71-80 2005

A Feature Selection for Text Categorization on Research Support System Papis
Lecture Notes in Computer Science pp524-533 No.3157 2004

コールセンターにおける大規模質問応答データに基づくFAQ作成支援システムの実装
第66回情報処理学会全国大会論文集 2004