

Многодокументный анализ референции в процессе извлечения информации на основе онтологии предметной области

СЕРЫЙ АЛЕКСЕЙ СЕРГЕЕВИЧ

Институт систем информатики имени А.П. Ершова СО РАН (Новосибирск), Россия
e-mail: alexey.seryj@iis.nsk.su

СИДОРОВА ЕЛЕНА АНАТОЛЬЕВНА

Институт систем информатики имени А.П. Ершова СО РАН (Новосибирск), Россия

ГАРАНИНА НАТАЛЬЯ ОЛЕГОВНА

Институт систем информатики имени А.П. Ершова СО РАН (Новосибирск), Россия

В данной работе исследуются пути развития разработанного ранее подхода к разрешению референциальной неоднозначности, возникающей в процессе извлечения информации из текстов. Рассмотрены возможности расширения модели разрешения кореференции, построенной для работы с единичным текстом, на текстовый корпус. Это позволило бы при анализе каждого текста учитывать результаты, полученные для всех текстов корпуса. Ожидается, что переход от анализа единичных документов к анализу корпуса повысит эффективность извлечения информации на основе онтологии предметной области.