

РЕФЕРАТ

магистерской аттестационной работы

на тему:

“ Технологии семантического поиска в массивах неструктурированной информации ”

Шумейко Юрия Дмитриевича

Актуальность работы

На сегодняшний день технические средства позволяют хранить и обрабатывать большие объёмы информации. Электронная информация используется повсеместно, во всех сферах общественной деятельности: в банках, библиотеках, научно-исследовательских учреждениях и т.д. Возможность нахождения необходимой информации в кратчайшее время становится важной и нетривиальной задачей. Для решения подобных задач применяются информационно-поисковые системы. Как правило, данные системы в процессе своей работы исходят из предположения о взаимной независимости слов текста документа, что затрудняет поиск релевантных документов. Многозначность слов, большое количество синонимов, омонимов и т.д. приводит к тому, что в результате работы системы могут быть получены нерелевантные поисковому запросу документы. Решением данной проблемы может являться информационно-поисковая система, способная распознавать смысловую нагрузку анализируемого документа. Поэтому поиск путей повышения эффективности работы информационно-поисковых систем является актуальной задачей в наше время.

Цель работы

Цель работы заключается в выявлении недостатков моделей информационного поиска и разработке способов их устранения. Разработка информационно-поисковой системы на основе векторной модели информационного

поиска, расширенной семантической моделью с применением онтологий для работы в едином поисковом механизме.

Задачи, решаемые в работе

1. Исследование особенностей существующих моделей информационного поиска.
2. Определение недостатков информационно-поисковых моделей и поиск путей их устранения.
3. Разработка модели системы информационного поиска с использованием онтологий.
4. Разработка информационно-поисковой системы на основе предлагаемой модели поиска с применением онтологий.

Достигнутые результаты

Выполнив задания, поставленные в работе, автор защищает:

- результаты анализа особенностей существующих моделей систем информационного поиска;
- разработанная модель информационного поиска с использованием онтологий;
- разработка информационно-поисковой системы на основе предложенной модели информационного поиска.

Научная новизна работы

Научная новизна работы заключается в том, что:

- предложено модель информационно-поисковой системы с применением онтологий предметных областей:

- для определения релевантности документа поисковому запросу учитывается смысл текста документа, что достигается путём применения онтологий предметных областей;
- применение онтологий предметных областей позволяет включить в предлагаемую информационно-поисковую систему модуль автоматического ответа на поставленный вопрос;
- применение онтологий предметных областей позволяет тематическое структурирование найденных документов.

Практическая ценность работы

Практическая ценность работы заключается в получении систематизированной теоретической информации об организации информационного поиска в массивах неструктурированной информации, используя разработанную модель информационного поиска.

Выводы

1. Проанализированы модели систем информационного поиска, рассмотрены их основные преимущества и недостатки.
2. Предложена модель информационно-поисковой системы, использующая для определения релевантности документов поисковому запросу онтологии предметных областей.
3. Разработана информационно-поисковая система на основе предложенной модели информационного поиска.

Работа содержит 86 с., 22 рис., 35 источников.

Ключевые слова: ПОИСКОВЫЕ СИСТЕМЫ, СЕМАНТИЧЕСКИЙ ПОИСК, ОНТОЛОГИИ.