

ОТЗЫВ

на магистерскую диссертацию магистра
факультета ПМ-ПУ СПбГУ

Чжао Шугуан

на тему: "Классификация текстов электронных писем для
упрощения поиска"

Перед автором работы поставлена задача разработки прототипа системы, которая бы использовала для упрощения навигации и поиска электронных писем такую форму представления данных, как облако тегов. Для решения данной задачи исследуются алгоритмы обработки текстов, которые позволили бы выделить из них наиболее популярные ключевые слова.

В качестве компонент разрабатываемой системы используется СУБД Oracle, Web сервер, поддерживающий технологию JSP. Процедуры обработки данных реализованы как хранимые процедуры и триггера базы данных, браузер же и JSP страницы используются для навигации по обработанным данным.

В части реализации самого комплекса отмечается его законченность, результаты были продемонстрированы и апробированы. При этом автором не исследована и не решена задача адаптации используемых алгоритмов и техник кластеризации таких специфических текстов, как сообщения электронной почты. Возникают нарекания к производительности решения, скорость обработки даже небольшого количества текстов очень низкая. Наконец, не проводится аналитика сценариев добавления в коллекцию новых и удаление существующих текстов, что соответствует специфике использования электронной почты. Всё это снижает практическую значимость полученного результата и не даёт возможности использовать полученный результат в промышленных условиях. Отсутствие перечисленной функциональности и аналитики являются недостатками работы.

Как замечание, стоит отметить сложность при прочтении текста работы, ввиду наличия большого количества ошибок и опечаток – сказывается недостаточный опыт написания работ на русском языке.

Учитывая указанные недостатки, считаю, что работа Чжао Шугуан заслуживает оценки «хорошо».

Научный
руководитель,
ст. преподаватель



Севрюков С.Ю.

25.06.2011