

ОТЗЫВ

на автореферат диссертации И.В. Соченкова «Реляционно-ситуационные структуры данных, методы и алгоритмы решения поисково-аналитических задач», представленной на соискание ученой степени кандидата физико-математических наук по специальности 05.13.17 – «Теоретические основы информатики».

Диссертационная работа И.В. Соченкова посвящена разработке и исследованию методов оценки сходства текстов на естественных языках на основе лингвистической информации, включающей в себя результаты морфологического, синтаксического и семантического анализа текстовых документов. Помимо этого, в рамках созданной информационно-аналитической системы, были реализованы структуры данных и алгоритмы, предназначенные для анализа текстовых документов с учетом лингвистической информации. Данное направление исследований, несомненно, является актуальным в связи с ростом количества обрабатываемой информации и повышением требований к качеству информационно поиска.

Важно отметить многокритериальность предлагаемого метода оценки сходства текстовых документов, основой для которой выступает предложенная автором структура поискового индекса, позволяющая хранить и использовать не только морфологическую, но и синтаксическую и семантическую информацию. Данное новшество позволяет строить эффективные процедуры как полнотекстового, так и фразового поиска без потери скорости работы информационно-аналитической системы.

Также важно подчеркнуть, что предлагаемые автором модели, методы анализа и структуры данных не привязаны к конкретному естественному языку, что в перспективе позволяет использовать данный подход при анализе документов не только на русском и английском языках.

Приведенные в автореферате результаты оценки качества созданной информационно-аналитической системы, а также корректное использование аппарата теории множеств, алгебры логики и методов объектно-ориентированного проектирования позволяют заключить, что разработанная система может быть использована для эффективного решения широкого круга задач анализа текстовой информации на различных естественных языках.

Результаты исследований опубликованы в 14 работах, из них 6 статей в рецензируемых научных журналах, входящих в список ВАК РФ, а также доложены на ряде научных конференций. Особенно хотелось бы отметить, что созданное программное обеспечение внедрено и успешно используется в целом ряде промышленных систем.

По автореферату следует сделать ряд **замечаний**:

1. На странице 11 вводится бинарное отношение между словоупотреблениями и тегами. Исходя из определения получается, что конкретное словоупотребление может встретиться в документе только в одной области (заголовок, список авторов), иначе условие, что каждому словоупотреблению соответствует единственный тег, не может быть выполнено. Не ясно как на практике обрабатываются случаи, когда одна и та же словоформа встречается в разных областях текста.
2. В работе не сказано, какой алгоритм ранжирования используется при выдаче поисковых результатов созданной ИАС.

Приведенные замечания не снижают общего качества представленной работы.

Диссертационная работа удовлетворяет требованиям ВАК РФ в соответствии с «Положением о порядке присуждения учёных степеней», предъявляемым к кандидатским диссертациям, а ее автор, Соченков И.В., заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 05.13.17 — «Теоретические основы информатики».

Доцент кафедры Защиты информации
МГТУ им. Н.Э. Баумана,
к.т.н.
abolkhovityanov@bmstu.ru

Подпись Болховитянова А.В. заверяю -



Болховитянов А.В.
9 октября 2014 г.