

## ОТЗЫВ НА АВТОРЕФЕРАТ ДИССЕРТАЦИИ

Лукашевич Натальи Валентиновны

“ Модели и методы автоматического анализа текстов на основе лингвистических онтологий в информационно-аналитических системах”,  
представленной на соискание ученой степени доктора технических наук  
по специальности 05.13.17 – Теоретические основы информатики

В настоящем отзыве-на-автореферат весьма трудно игнорировать широко известную специалистам положительную информацию о научной работе Н.В.Лукашевич. Последние 30 лет Наталья Валентиновна является ведущим специалистом в области разработки информационных ресурсов понимания русского естественного языка. Особенное значение ее работа получила

- в связи с развитием сети Интернет,
- в связи со взрывообразным ростом объемов слабоструктурированной информации, а также
- в связи активным развитием информационно-поисковых и информационно-аналитических систем способных накапливать текстовые данные.

Известно, что для автоматической обработки потоков текстовой информации применяются алгоритмы автоматической классификации, кластеризации и аннотирования текстов. Для качественной работы таких систем в конкретных предметных областях необходимо использовать знания о предметной области, в частности, знания о ее терминологии.

В диссертационной работе Н.В.Лукашевич рассматривается модель построения специализированной онтологии для автоматической обработки больших потоков текстов, поступающих в информационную систему. Базовая модель была проверена и усовершенствована при построении многочисленных терминологических ресурсов в различных предметных областях. Кроме того, на базе предложенной модели разработан тезаурус русского языка РуТез.

Созданные Н.В.Лукашевич ресурсы позволяют автоматически анализировать структуру связного текста. В работе предложен алгоритм построения тематического представления документа, состоящего из автоматически собранных групп близких по смыслу терминов (тематических узлов), которые

подразделяются на локальные и основные тематические узлы (соответствующие основной теме документа). Построенное тематическое представление используется для автоматического расширения поисковых запросов в информационно-поисковых системах, автоматической рубрикации и аннотирования документов.

В диссертации Н.В.Лукашевич также исследуются методы автоматического извлечения терминов из коллекции текстов предметной области. Предложена модель, основанная на нескольких типах признаков терминов, которые комплексированы с методами машинного обучения.

Диссертация Лукашевич Натальи Валентиновны является законченной самостоятельной научно-квалификационной работой, в которой разработаны теоретические положения, совокупность которых можно квалифицировать как научное достижение в решении актуальной научной проблемы разработки сверхбольших лингвистических ресурсов для автоматической обработки текстов, а также для автоматического анализа структуры связного текста.

Принимая во внимание актуальность темы диссертации, ее научную новизну и практическую значимость, считаю, что диссертационная работа «Модели и методы автоматического анализа текстов на основе лингвистических онтологий в информационно-аналитических системах» удовлетворяет требованиям ВАК при Минобрнауки России, предъявляемым к докторским диссертациям, а ее автор, Лукашевич Наталья Валентиновна, заслуживает присуждения ученой степени доктора технических наук по специальности 05.13.17 – "Теоретические основы информатики".

Профессор кафедры алгоритмических языков  
факультета ВМК МГУ имени М.В.Ломоносова,  
доктор физико-математических наук, профессор  
140002, г.Люберцы, Октябрьский пр-т 12, кв.53  
Тел.: +7(909)9254642. E-mail: glosscom@mail.ru

20 сентября 2016 г.

Соловьев  
Сергей  
Юрьевич



УДОСТОВЕРЯЮ  
СПЕЦИАЛИСТ  
Г.КОВАЛЕНКО