

ОТЗЫВ НА АВТОРЕФЕРАТ ДИССЕРТАЦИИ

Лукашевич Натальи Валентиновны

“Модели и методы автоматического анализа текстов на основе лингвистических онтологий в информационно-аналитических системах”
по специальности 05.13.17 – Теоретические основы информатики

В настоящее время для обработки огромных потоков текстовой информации в информационных системах применяются в основном статистические методы, в минимальной степени учитывающие знания о мире и предметной области.

В диссертации Лукашевич Н.В. предлагается специализированная модель описания понятийной модели предметной области, которая направлена на использование для автоматической обработки текстов в рамках информационно-поисковых и информационно-аналитических систем, включая такие приложения как автоматическая классификация документов, автоматическое аннотирование одного или многих документов, автоматическое расширение запросов, визуализация поисковой выдачи.

Предложенная модель была построена в результате многих экспериментов на реальных текстовых данных и стала основой для нескольких больших компьютерных ресурсов для обработки текстов, включая Общественно-политический тезаурус, тезаурус русского языка РуТез, Онтология по естественным наукам и технологиям (ОЕНТ), Авиа-Онтология и др. Недавно первая версия тезауруса РуТез, содержащая 100 тысяч слов и выражений, была опубликован и доступна всем желающим для некоммерческого использования.

В диссертационной работе Лукашевич Н.В. уделяется большое внимание тестированию предложенных ресурсов и технологий. В частности, предложенная технология автоматической рубрикации документов была протестирована на семинаре по информационному поиску РОМИП-2007, когда за восемь человеко-часов была создана система рубрикации по новому рубрикатору, заданному организаторами. Созданная система рубрикации

продemonстрировала высокое качество автоматической рубрикации по сравнению с другими участниками тестирования. Полученные результаты базируются на заранее накопленных знаниях, описанных в Общественно-политическом тезаурусе. Таким образом, модель лингвистической онтологии, описанная в диссертации Лукашевич Н.В., продемонстрировала высокую надежность при переходе к новой поставленной задаче.

Считаю, что диссертация Лукашевич Натальи Валентиновны является законченной самостоятельной научно-исследовательской работой, совокупность результатов которой можно квалифицировать как существенное продвижение в решении актуальной научной проблемы сверхбольших лингвистических ресурсов для автоматической обработки текстов, автоматического анализа структуры связного текста.

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

Д.ф.-м.н., профессор, ведущий научный сотрудник
Казанского федерального университета
420008, Казань, ул. Кремлевская, д. 18, ИТИС,
89196910489, maki.solovyev@mail.ru
Соловьев Валерий Дмитриевич

В.С.Сол
21.09.2016



Валерий Соловьев В.Д. заверю

21.09.2016

Михаил Леоненко Д.В. Леон