

УТВЕРЖДАЮ

Проректор Национального
исследовательского университета
«Высшая школа экономики»



К.Э.Н., доцент Рощин Сергей Юрьевич

Игорь 2015 г.

ЗАКЛЮЧЕНИЕ

федерального государственного автономного образовательного
учреждения высшего профессионального образования
«Национальный исследовательский университет
«Высшая школа экономики»

Диссертация на тему: «Автоматическая разметка семантических ролей в русском языке» выполнена соискателем Кузнецовым Ильей Олеговичем в Школе лингвистики факультета гуманитарных наук федерального государственного автономного образовательного учреждения высшего профессионального образования «Национальный исследовательский университет «Высшая школа экономики» (далее – Национальный исследовательский университет «Высшая школа экономики»).

В период подготовки диссертации соискатель Кузнецов Илья Олегович работал в Научно-образовательном центре Семантических технологий Национального исследовательского университета «Высшая школа экономики» в должности младшего научного сотрудника с мая 2011 г. по февраль 2014 г.

В 2010 г. окончил Московский государственный университет им. М.В. Ломоносова. Присуждена квалификация «Лингвист» по специальности «Теоретическая и прикладная лингвистика»

Удостоверение о сдаче кандидатских экзаменов выдано в 2015 г. управлением аспирантуры и докторантуры Национального исследовательского университета «Высшая школа экономики»

Научный руководитель – кандидат филологических наук, доцент Бонч-Осмоловская Анастасия Александровна, доцент Школы лингвистики Национального исследовательского университета «Высшая школа экономики», старший научный сотрудник НОЦ Семантических технологий Национального исследовательского университета «Высшая школа экономики».

По результатам рассмотрения диссертации : «Автоматическая разметка семантических ролей в русском языке»

принято следующее заключение.

Диссертационное исследование посвящено автоматической разметке семантических ролей в русском языке.

Актуальность выбранной темы обусловлена большим прикладным значением исследований в выбранной области и популярностью темы в современной компьютерной лингвистике. Предложенные в работе методы впервые апробированы на русском материале. Кроме того, это первая работа, которая использует открытый ресурс FrameBank для выполнения автоматической разметки семантических ролей.

Научная новизна работы состоит в том, что ранее подобных исследований на русском материале не проводилось. Предложенное исследование – первый опыт применения систем на основе машинного обучения к корпусу примеров FrameBank. Ряд частных решений также применяется к русскому языку впервые, кроме того, это первое известное нам полноценное описание подобной системы, достаточно подробное для успешной реимплементации и усовершенствования предложенного метода.

Теоретическая значимость исследования состоит в оценке вклада различных лингвистических свойств в качество работы классификатора. Система опирается на ряд свойств, в т.ч. на путь в дереве зависимостей, падеж актанта, предлог, которым оформлен актант, лемму и кластер актанта и др. Представленный анализ демонстрирует важность синтаксических свойств для автоматической разметки актантов в русском языке. Роль лексических свойств оказывается второстепенной, подробно рассматриваются возможные причины такого поведения системы. Для русского языка подобное исследование проводится впервые.

Практическая значимость исследования состоит в подробном качественном и количественном анализе результатов работы системы. Кроме того, работа содержит детальное описание компонентов системы, а также ряд рекомендаций для дальнейших экспериментальных исследований в этой области.

Достоверность выводов диссертации обеспечивается опорой на фундаментальные работы в области лингвистики и машинного обучения. Научные положения работы обоснованы и подтверждены в ходе эксперимента по оценке качества системы, который включал в себя автоматическую оценку на основе тестового набора данных, так и детальный экспертный анализ результатов.

Личный вклад соискателя состоит в проектировании и описании системы автоматической разметки семантических ролей, оценке качества работы системы и интерпретации полученных результатов.

Работа была **апробирована** в ходе обсуждений на заседаниях и семинарах кафедры, а также на конференциях

- АИСТ-2013 (4-6 апреля 2013 г., Екатеринбург)
- АИСТ-2015 (9-11 апреля 2015 г., Екатеринбург)
- Диалог-2013 (29 мая – 2 июня 2013 г., Бекасово)
- ECIR Doctoral Consortium (24 марта 2013 г., Москва)

Основные результаты диссертационного исследования **опубликованы** в 6 работах общим объёмом 6,7 п.л. Личный вклад автора составляет 5,08 п.л.

Работы, опубликованные автором в ведущих рецензируемых научных журналах и изданиях, рекомендованных ВАК:

1. Кузнецов И. Автоматическое выделение глагольных актантов: теоретическая основа и актуальные подходы. // НТИ. Информ. процессы и системы. ВИНТИ РАН. – 2012 № 12. – С. 34-44 – 1,39 п.л.
2. Кузнецов И. Автоматическое извлечение двусловных терминов по тематике «Нанотехнологии в медицине» на основе корпусных данных. // НТИ. Информ. процессы и системы. ВИНТИ РАН. – 2013 № 5. – С. 25-33 – 1,27 п.л.
3. Акинина Ю., Кузнецов И., Толдова С. Сравнение двух методов автоматического извлечения участников из неструктурированных источников. // НТИ. Информ. процессы и системы. ВИНТИ РАН. – 2013 № 6. – С. 24-34 – 1,16 п.л. (личный вклад – 0,39 п.л.)

Работы, опубликованные автором в других изданиях:

1. Кузнецов И. Распознавание и классификация актантов в русском языке // Доклады Всероссийской научно-практической конференции «Анализ Изображений, Сетей и Текстов». Екатеринбург, 2013. – С. 205–210 – 0,92 п.л.
2. Akinina Y., Kuznetsov I., Toldova S. The impact of syntactic structure on verb-noun collocation extraction // Компьютерная лингвистика и интеллектуальные технологии: По материалам ежегодной Международной конференции «Диалог» (Бекасово, 29 мая – 2 июня 2013 г.). Вып. 12 (19). – М.: Изд-во РГГУ, 2013. С. 2-16 – 1,27 п.л. (личный вклад - 0,42 п.л.)
3. Kuznetsov I. Semantic Role Labeling for Russian Language based on Russian FrameBank // Analysis of Images, Social Networks and Texts. 4th International Conference, AIST 2015, Yekaterinburg, Russia, April 9–11, 2015, Revised Selected Papers / Ed. by M. Y. Khachay, N. Konstantinova, A. Panchenko, D. I. Ignatov, V. Labunets. Vol. 542: Series: Communications in Computer and Information Science. Springer International Publishing, 2015. P. 337-348. – 0,69 п.л.

Публикации полностью соответствуют теме диссертационного исследования и раскрывают её основные положения.

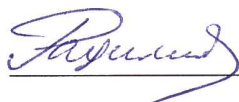
Диссертационная работа Кузнецова Илья Олеговича на тему «Автоматическая разметка семантических ролей в русском языке» – это законченная научно-квалификационная работа, соответствующая требованиям п. 9,10 Положения о присуждении ученых степеней, Паспорту научной специальности 10.02.21 «Прикладная и математическая лингвистика» группы специальностей 10.02.00 «Языкознание» отрасли 10.00.00 «Филологические науки» Номенклатуры специальностей научных работников.

Диссертация на тему: «Автоматическая разметка семантических ролей в русском языке» Кузнецова Ильи Олеговича соответствует всем требованиям ВАК РФ, предъявляемым к диссертациям на соискание ученой степени кандидата филологических наук и рекомендуется к защите на соискание ученой степени кандидата филологических наук по специальности 10.02.21 «Прикладная и математическая лингвистика».

Заключение принято на заседании Школы лингвистики факультета гуманитарных наук Национального исследовательского университета «Высшая школа экономики».

Присутствовало на заседании 10 чел. Результаты голосования: «за» – 10 чел., «против» – 0 чел., «воздержалось» – 0 чел., протокол № 2/1 от «9» июня 2015 г.

Доктор филологических наук,
профессор,
руководитель школы лингвистики
факультета гуманитарных наук НИУ ВШЭ,



Екатерина Владимировна Рахилина



Юркина заверяю
Заместитель начальника
Управления персонала
Шатула Т.А.