

ОТЗЫВ

о выпускной квалификационной работе Гаврилики Дарьи Александровны «Экспериментальное исследование методов автоматического выделения ключевых выражений в корпусах русскоязычных текстов»

Научная проблема, которую Дарья Александровна Гаврилики выбрала для исследования в своей выпускной квалификационной работе, относится к такой области прикладной лингвистики как автоматическое понимание текстов. Дарья Александровна сфокусировала свое внимание на автоматическом выделении ключевых выражений как разновидности семантической компрессии. Цель исследования – сравнить разноплановые методы автоматического выделения выражений (статистические, лингвистические и гибридные), используя при этом стилистически разнородные данные. Автор основывает свое исследование на солидном теоретическом фундаменте, объединяющем в себе основные положения корпусной лингвистики и методов анализа данных.

В ходе экспериментов автор выдвигает и подтверждает гипотезу о тяготении ключевых выражений к началу текста. Особенностью ВКР является нестандартный экспериментальный дизайн: Дарья Александровна предложила сравнивать экспертную разметку и автоматическую, но в особых условиях. Проверка гипотезы предполагает экспертную разметку ключевых выражений в начале текста, а в оставшейся части текста ключевые выражения выделяются автоматическими методами. Если гипотеза верна, тогда число расхождений между экспертной разметкой и выдачей автоматических методов будет превышать число совпадений. Для проверки этой гипотезы автору потребовалось решить ряд трудоемких задач, среди которых сборка, предобработка, морфологическая разметка авторского корпуса, его экспертный анализ и обработка автоматическими методами, анализ полученных данных. Исследовательский корпус включает в себя тексты новостей, научные тексты и тексты художественных произведений.

Значительная заслуга автора ВКР заключается в том, что помимо сравнения уже существующих методов Дарья Александровна предложила, реализовала на языке Python и протестировала свой собственный экстрактор, который по качеству превосходит стандартные инструменты. В этой связи можно считать работу Дарьи Александровны новаторской, предложенная ею разработка может рассматриваться как задел для исследования более высокого уровня. Результаты автоматической обработки данных отражены в ВКР с подробными комментариями, а также представлены в представительных приложениях. Все задачи решены успешно и в полном объеме.

Научный руководитель считает должным отметить, что все исследовательские процедуры были выполнены Дарьей Александровной самостоятельно и на высоком профессиональном уровне, результаты изложены четко, ясно и убедительно. В процессе работы над ВКР автор проявляет трудолюбие, настойчивость, пытливость, открытость новой информации. Хочется отметить, что за время обучения в магистратуре и в ходе подготовки ВКР Дарья Александровна постоянно совершенствовалась, она уверенно владеет основными методами и инструментами автоматической обработки данных, навыками программирования, статистической обработки результатов наблюдений и т.д. Дарья Александровна – молодой исследователь с широким научным кругозором. Сотрудничество с Дарьей Александровной было большой радостью!

Итак, выпускная квалификационная работа, представленная к защите, соответствует всем требованиям, предъявляемым к научным произведениям такого уровня, а ее автор заслуживает присвоения квалификации магистра лингвистики. Содержание ВКР полностью соответствует заявленной в названии теме, тема раскрыта полностью, структура ВКР соответствует цели и задачам исследования, в ВКР в полной мере освещены актуальные проблемы теоретического и практического характера. При выполнении исследования и подготовке текста ВКР автор использовал современную литературу, достижения науки и практики. Текст ВКР прекрасно структурирован, написан строгим научным стилем, автор четко формулирует теоретические предпосылки исследования, детально комментирует практический материал, дает развернутое обоснование выводов. Результаты, достигнутые автором ВКР, имеют высокую теоретическую значимость, практическую ценность и могут найти применение в самых разных областях прикладной лингвистики.

11 июня 2021 г.



Научный руководитель, кандидат филологических наук, доцент кафедры
математической лингвистики СПбГУ О.А.Митрофанова

Согласие на обработку персональных данных

Я, **Митрофанова Ольга Александровна**, даю согласие на обработку своих персональных данных оператору - Федеральное государственное бюджетное образовательное учреждение высшего образования «Санкт-Петербургский государственный университет» (далее - СПбГУ), 199034, Санкт-Петербург, Университетская наб., д. 7-9, на следующих условиях:

1. Оператор осуществляет обработку персональных данных исключительно в связи с осуществлением рецензирования и проведения защиты выпускных квалификационных работ обучающихся СПбГУ в целях реализации принципа открытости образовательной деятельности.
2. Перечень персональных данных, передаваемых Оператору на обработку:
 - фамилия, имя, отчество;
 - место работы, должность;
 - ученая степень и звание (при наличии);
 - контактный телефон и адрес электронной почты.
3. Оператор имеет право на обработку персональных данных, то есть совершение, в том числе, следующих действий: обработку (включая сбор, систематизацию, накопление, хранение, уточнение (обновление, изменение), использование, обезличивание, блокирование, уничтожение персональных данных).
4. Данным заявлением разрешаю считать общедоступными, в том числе выставлять в сети Интернет, следующие персональные данные: фамилия, имя, отчество, место работы, должность, ученая степень и звание (при наличии).
5. Обработка персональных данных осуществляется оператором в соответствии с нормами Федерального закона от 27.07.2006 № 152-ФЗ «О персональных данных» и смешанным способом.
6. Срок действия данного Согласия не ограничен.

11 июня 2021 г.



Митрофанова О.А.