

ОТЗЫВ НАУЧНОГО РУКОВОДИТЕЛЯ

на диссертационную работу Чисовой Елены Викторовны «Методы анализа риторической структуры текстов на русском языке», представленную на соискание ученой степени кандидата технических наук по специальности 1.2.1 – «Искусственный интеллект и машинное обучение»

В течение последних нескольких лет, начиная с бакалавриата РУДН, Чисова Елена Викторовна активно занимается научной деятельностью в области обработки естественного языка и анализа текстов. В 2020 году Елена Викторовна защитила магистерскую работу по дискурсивному анализу текста и продолжала заниматься этой тематикой в аспирантуре ФИЦ ИУ РАН с 2020 по 2024 годы. За это время Елена Викторовна показала себя инициативным и квалифицированным исследователем, активно участвовала в проектах РФФИ, публиковала статьи и выступала на международных конференциях.

Работа Чисовой Елены Викторовны посвящена исследованию методов и разработке новых алгоритмов риторического анализа текстов на русском языке, а также изучению возможностей кросс-языкового риторического анализа и объединения разнородных размеченных корпусов для создания жанрово-универсального риторического анализатора. Понимание естественного языка машинами невозможно без привлечения знаний о строении языка и структуре текста. Изучение законов построения связного текста, а также создание методов автоматического извлечения из текстов риторических структур, отражающих связность текста, являются актуальными задачами, особенно в приложении к русскому языку.

За время работ по теме диссертации Еленой Викторовной самостоятельно получен ряд важных результатов, вносящих значимый вклад в развитие методов риторического анализа текстов на русском языке. Риторический анализ является сложной задачей, так как из текста необходимо извлечь многоуровневую древовидную структуру в условиях неоднозначности ручной разметки и нехватки размеченных корпусов. Эта задача не решалась ранее для текстов на русском языке. Елена Викторовна разработала методы извлечения из текстов как локальных, на уровне абзаца, так и глобальных, на уровне всего текста, риторических структур, исследовав при этом различные подходы к построению риторических деревьев.

При обработке естественного языка последние годы применяются кросс-языковые или межязыковые подходы, которые повышают качество решения различных задач для каждого языка в отдельности. Однако, кросс-языковой анализ риторической структуры остаётся сложной задачей, т.к. существует много несовпадающих между собой интерпретаций теории риторических структур, а параллельных корпусов с риторической разметкой почти нет. Елена Чисова решила эту задачу, разработав русскоязычную версию корпуса Джорджтаунского университета и обучив на нём двуязычные (английский-

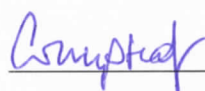
русский) модели, экспериментально показав преимущества двуязычных моделей по сравнению с моноязычными при отсутствии значительного объёма разметки текстов на русском языке. Аналогично ею решалась задача кросс-жанрового риторического анализа.

Значительная часть диссертационной работы посвящена применению риторических структур для решения некоторых ключевых задач обработки текстов с оценкой влияния учета этих структур на качество решения указанных задач. Степень влияния верхнеуровневых (дискурсивных, сюжетных) представлений текста на качество обработки текстов, особенно длинных текстов со сложными рассуждениями, недостаточно исследована для всех языков. Елена Чистова разработала методы классификации текстов и разрешения кореференции с учетом риторических структур, а также метод извлечения структур аргументации на основе риторических структур. Она экспериментально показала, что учёт риторических признаков позволяет повысить качество классификации текстов с рассуждениями, качество разрешения кореференции и анализа структуры аргументации.

Результаты работы опубликованы в высокорейтинговых изданиях и апробированы на авторитетных научных мероприятиях, включая конференции уровня А* по рейтингу CORE, что говорит о значимости и актуальности полученных результатов. Всего Елена Викторовна имеет 17 научных публикаций, основные результаты диссертационной работы представлены в 9 научных публикациях, из которых 2 в журналах из Перечня ВАК РФ, а 6 — в сборниках трудов международных конференций, индексируемых в Scopus. Добавлю, что Елена Викторовна является активным членом научного сообщества, выступает со-организатором международного Семинара по анализу аргументации (Workshop on Argument Mining), рецензентом статей на ведущих конференциях по компьютерной лингвистике Диалог и ACL, что подтверждает её высокую квалификацию.

Диссертация Чистовой Елены Викторовны является законченной научной работой, соответствует требованиям Положения о присуждении ученых степеней, предъявляемым к кандидатским диссертациям. Считаю, что Чистова Елена Викторовна заслуживает присуждения степени кандидата технических наук по специальности 1.2.1 — «Искусственный интеллект и машинное обучение».

Научный руководитель,
зав. отделом №72 ФИЦ ИУ РАН
Д.Т.Н., доцент

 И.В. Смирнов

04.09.2025



Р.В. РАЗУМЧИК
по доверенности
№ 1968-105 от 26.09.2024 г.