

«УТВЕРЖДАЮ»

Директор Федерального

государственного учреждения

«Федеральный исследовательский
центр «Информатика и управление»

Российской академии наук»,



М.А. Посыпкин

«08» сентября 2025 г.

ЗАКЛЮЧЕНИЕ

Федерального государственного учреждения «Федеральный исследовательский центр
«Информатика и управление» Российской академии наук»

Диссертация Чистовой Елены Викторовны на тему «Методы анализа риторической структуры текстов на русском языке» выполнена в отделе №72 «Интеллектуальный анализ информации» отделения №7 Федерального государственного учреждения «Федеральный исследовательский центр «Информатика и управление» Российской академии наук» (ФИЦ ИУ РАН).

Чистова Елена Викторовна, 1996 года рождения, гражданка России, в 2020 году окончила ФГАОУ ВО «Российский университет дружбы народов» по специальности «Фундаментальная информатика и информационные технологии».

Тема диссертационного исследования была утверждена на заседании ученого совета Федерального исследовательского центра «Информатика и управление» Российской академии наук, протокол №1 от 08.02.2021 г.

В период подготовки диссертации соискатель Чистова Елена Викторовна работала в отделе №72 ФИЦ ИУ РАН в должности программиста с 2018 по 2023 гг. и младшего научного сотрудника с 2023 г. по настоящее время.

Справка о сдаче кандидатских экзаменов по специальности 1.2.1 «Искусственный интеллект и машинное обучение» выдана в 2025 году в ФИЦ ИУ РАН.

Научный руководитель — доктор технических наук, доцент Смирнов Иван Валентинович, заведующий отделом №72 Федерального исследовательского центра «Информатика и управление» Российской академии наук.

По итогам обсуждения принято следующее заключение.

Актуальность темы

Многие задачи обработки и понимания естественного языка требуют рассмотрения текста не только на уровне слов, но и на уровне связности предложений, представляемой в виде риторической структуры. Существующие методы анализа риторической структуры ориентированы на англоязычные тексты, поэтому разработка методов анализа риторической структуры текстов на русском языке является актуальной и практически востребованной научной задачей, требующей для своего решения новых научных подходов.

Обоснованность научных положений

Положения, выносимые Чистовой Е. В. на защиту в полной мере обоснованы в проведенном научном исследовании и подтверждены полученными практическими результатами.

Личное участие соискателя ученой степени в получении результатов,

изложенных в диссертации

Все результаты диссертационной работы получены автором самостоятельно.

Степень достоверности результатов проведенных соискателем ученой степени исследований

Достоверность результатов диссертационной работы подтверждается использованием надежных и апробированных методов исследования, согласованностью результатов проведенных экспериментальных исследований с известными теоретическими положениями в области риторического анализа текстов. Результаты прошли апробацию на международных и российских конференциях, представлены высокорейтинговыми публикациями.

Научная новизна работы

- Впервые разработаны методы риторического анализа текстов на русском языке, включая анализ локальных и полнотекстовых дискурсивных структур.
- Впервые исследованы возможности кросс-языковой адаптации риторического анализа на материале большого параллельного корпуса дискурсивной разметки.
- Предложен метод риторического дискурсивного анализа, позволяющий достичь качества полнотекстового анализа риторической структуры на русском и английском языках, превышающего качество предыдущих систем.
- Впервые предложен метод реализации риторического анализа при помощи глубокого обучения на материалах разнородной риторической разметки.
- Разработан метод классификации текстов с учетом риторических структур, показана его эффективность в задачах классификации тональности и аргументации.
- Разработан метод разрешения кореференции с учётом риторической структуры.
- Разработан метод построения структур аргументации на основе риторических структур. Экспериментально показано, что использование нескольких вариантов дискурсивной структуры при обучении анализатора аргументации улучшает качество построения структур аргументации в рассуждениях на русском и английском языках.

Теоретическая значимость

Созданы научно-обоснованные подходы к поверхностному и полнотекстовому риторическому анализу текстов на русском языке, предложены подходы к риторическому анализу на кросс-языковой и разнородной риторической разметке. Предложены подходы к применению риторического анализа к базовым задачам обработки естественного языка.

Практическая значимость результатов проведенных соискателем ученой степени исследований

Разработанные в рамках диссертации методы риторического анализа текстов на русском языке реализованы в открытом дискурсивном анализаторе. Разработанные методы риторического анализа текстов на русском языке являются основой для классификации, разрешения кореференции, построения структур аргументации и других прикладных задач обработки текстов на естественном языке. Разработанное программное обеспечение, реализующее методы анализа риторической структуры текстов на русском языке, внедрены в программные средства лингво-статистического и психолингвистического анализа текстов ООО «РИ ТЕХНОЛОГИИ», что подтверждается справкой об использовании.

Апробация работы

Основные материалы диссертации и результаты исследования докладывались на следующих конференциях:

- Международная конференция «Диалог 2019», (Россия, Москва, июнь 2019 г.), “Classification models for RST discourse parsing of texts in Russian”;
- Workshop on Discourse Relation Parsing and Treebanking 2019, (США, Миннеаполис, июнь 2019 г.), “Towards the Data-driven System for Rhetorical Parsing of Russian Texts”;
- Мультиконференция по проблемам управления МКПУ, (Россия, Геленджик,

сентябрь 2019 г.), «Модели и методы дискурсивного анализа текстов для решения задач интеллектуальной обработки и понимания текстов»;

– Analysis of Images, Social Networks and Texts: 9th International Conference, AIST 2020, (Россия, Москва, октябрь 2020 г.), “RST Discourse Parser for Russian: an Experimental Study of Deep Learning Models”;

– Международная конференция «Диалог 2022», (Россия, Москва, июнь 2022 г.), “Discourse-aware text classification for argument mining”;

– Международная конференция «Диалог 2023», (Россия, Москва, июнь 2023 г.), “Light Coreference Resolution for Russian with Hierarchical Discourse Features”;

– The 61st Annual Meeting of the Association for Computational Linguistics, (Канада, Торонто, июль 2023 г.) (A* по рейтингу CORE), “End-to-End Argument Mining over Varying Rhetorical Structures”;

– The 62nd Annual Meeting of the Association for Computational Linguistics, (Таиланд, Бангкок, август 2024 г.) (A* по рейтингу CORE), “Bilingual Rhetorical Structure Parsing with Large Parallel Annotations”.

Полнота изложения материалов диссертации в публикациях

Основные результаты по теме диссертации изложены в 9 научных публикациях общим объемом 12,8 п.л; личный вклад автора составляет 12,5 п.л. Зарегистрирована 1 программа для ЭВМ.

1. Classification models for RST discourse parsing of texts in Russian / E. Chistova [et al.] // Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue”. — 2019. — P. 163—176. **(Scopus)**
2. Towards the Data-driven System for Rhetorical Parsing of Russian Texts / E. Chistova [et al.] // Proceedings of the Workshop on Discourse Relation Parsing and Treebanking 2019. — Minneapolis, MN : Association for Computational Linguistics, 2019. — P. 82—87.
3. RST discourse parser for Russian: an experimental study of deep learning models / E. Chistova [et al.] // Analysis of Images, Social Networks and Texts: 9th International Conference, AIST 2020. — Springer. Moscow, Russia, 2021. — P. 105—119. **(Scopus)**
4. Chistova E., Smirnov I. Discourse-aware text classification for argument mining // Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue”. — 2022. — P. 93—105. **(Scopus)**
5. Chistova E., Smirnov I. Light Coreference Resolution for Russian with Hierarchical Discourse Features // Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue”. — 2023. — P. 34—41. **(Scopus)**
6. Chistova E. End-to-End Argument Mining over Varying Rhetorical Structures // Findings of the Association for Computational Linguistics: ACL 2023. — Toronto, Canada : Association for Computational Linguistics, 2023. — P. 3376—3391. **(Scopus, CORE A*)**
7. Chistova E. Bilingual Rhetorical Structure Parsing with Large Parallel Annotations // Findings of the Association for Computational Linguistics: ACL 2024. — Bangkok, Thailand, virtual meeting : Association for Computational Linguistics, 2024. — P. 9689—9706. **(Scopus, CORE A*)**
8. Чистова Е. В. Методы анализа риторических структур в текстах на русском языке // Искусственный интеллект и принятие решений. — 2024. — № 4. — С. 79—92. **(ВАК K1)**
9. Чистова Е. В. Влияние признаков иерархического дискурса на разрешение кореференции в русском языке // Искусственный интеллект и принятие решений. — 2025. — № 1. — С. 95—102. **(ВАК K1)**
10. Свидетельство о гос. регистрации программы для ЭВМ №2024618391

Российская Федерация. Программа для анализа дискурсивной риторической структуры текстов // Е. В. Чистова; правообладатель: Общество с ограниченной ответственностью «РИ Технологии» – Заявка № 2024616947 04.04.2024; опубл. 11.04.2024.

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

Ценность научных работ соискателя ученой степени

Ценность научных работ соискателя определяется возможностью их применения в продвинутых системах обработки естественного языка и интеллектуального анализа текстов, требующих интерпретируемости результатов. Полученные в диссертации результаты могут быть использованы также в научных исследованиях в области компьютерной лингвистики и психолингвистики.

Результаты диссертационного исследования соответствуют паспорту специальности

1.2.1 «Искусственный интеллект и машинное обучение», а именно:

- пункту 4 «Разработка методов, алгоритмов и создание систем искусственного интеллекта и машинного обучения для обработки и анализа текстов на естественном языке, для изображений, речи, биомедицины и других специальных видов данных»;
- пункту 5 в части «Методы и технологии поиска, приобретения и использования знаний и закономерностей, в том числе – эмпирических, в системах искусственного интеллекта»;
- пункту 7 в части «Разработка специализированного математического, алгоритмического и программного обеспечения систем искусственного интеллекта и машинного обучения».

Диссертация соответствует требованиям п.14 Положения о присуждении ученых степеней (Положение правительства РФ №842 от 24.09.2013 в действующей редакции) и не содержит заимствованного материала без ссылки на авторов.

Диссертация Чистовой Елены Викторовны «Методы анализа риторической структуры текстов на русском языке» является законченной научно-квалификационной работой, результаты которой можно квалифицировать как научное достижение в области интеллектуального анализа текстов на естественном языке, имеющее важное значение для создания эффективных систем текстовой аналитики. Диссертация соответствует критериям Положения о присуждении ученых степеней, предъявляемым к кандидатским диссертациям, и рекомендуется к защите на соискание ученой степени кандидата технических наук по специальности 1.2.1 «Искусственный интеллект и машинное обучение» в диссертационном совете 24.1.224.03 при Федеральном государственном учреждении «Федеральный исследовательский центр «Информатика и управление» Российской академии наук».

Заключение принято на заседании семинара Отделения №7 ФИЦ ИУ РАН 28 августа 2025г., протокол №7. Присутствовало на заседании 11 человек. Результаты голосования: «за» – 11 человек, «против» – нет, «воздержалось» – нет.

Председательствующий на заседании:
Руководитель отделения 7
доктор технических наук



О.Г. Григорьев

Секретарь семинара
с.н.с., руководитель Отдела 73
к.ф.-м.н.



Д.А. Девяткин