

ОТЗЫВ

на автореферат диссертационной работы Большаковой Светланы Анатольевны «Совершенствование методов компьютерной обработки текстовой информации в аспекте задач, связанных с омонимией и синонимией», представленной на соискание ученой степени кандидата технических наук по специальности 2.3.1 – Системный анализ, управление и обработка информации, статистика (технические науки)

Одной из актуальных и социально-значимых задач обработки текстов на естественном языке является преобразование сложных текстов в тексты, использующие более простой и понятный язык. В рассматриваемой диссертации основная научная задача заключается в повышении точности и эффективности обработки русскоязычных текстов за счёт разработки и внедрения методов автоматического разрешения частеречной омонимии и лексического упрощения текстов путём замены сложных слов и словосочетаний на более простые и частотные синонимы с сохранением синтаксиса и семантики.

Автор справедливо отмечает, что существующие системы (такие как MyStem, Pymorphy2) недостаточно эффективно справляются с омонимией в русском языке, особенно в случае предикативов и устойчивых словосочетаний.

Научная новизна и практическая значимость результатов:

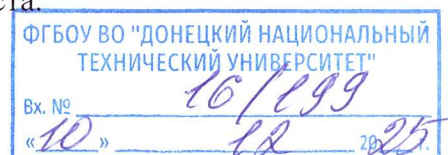
1. Предложен гибридный подход к снятию омонимии, сочетающий: декларативные знания (словари), продукционные правила, контекстный анализ и позволяющий достичь точности 99,3% для ряда типов омонимии.
2. Предложен метод лексического упрощения текста на основе специально размеченной базы синонимов и правил синтаксической адаптации. Точность замены с сохранением смысла и грамматики выше 96%.
3. Предложен статистический метод автоматического разбиения текста на семантически однородные фрагменты (абзацы) на основе частоты ключевых слов и их распределения в тексте.

Результаты работы внедрены в процессы обработки текстовых данных ФГБНУ "РАНИМИ" и ФГБНУ «Институт проблем искусственного интеллекта», о чём свидетельствуют справки о внедрении. Это подтверждает востребованность разработок и их применимость в промышленных условиях. Разработанные методы и алгоритмы, а также размеченные текстовые корпуса и базы синонимов могут быть использованы как компоненты в NLP-системах различного назначения: машинного перевода, информационного поиска, автоматического реферирования, классификации текстов.

Автор демонстрирует высокий уровень профессиональной компетенции, представляя материал в логичной и хорошо структурированной форме. Масштаб выполненных исследований убедительно свидетельствует о научной обоснованности полученных результатов.

Однако, в диссертационной работе имеются следующие недостатки:

1. Предложенный подход основан на продукционных правилах и словарях, что делает его уязвимым к изменениям в языке и требует постоянного ручного обновления.
2. Работа сосредоточена на частеречной омонимии, но почти не затрагивает другие виды неоднозначности (например, семантическую или синтаксическую), которые также критичны для полноценного понимания текста.



3. В автореферате приведено сравнение только с MyStem и Rumorphy2, но не с трансформерными моделями (например, ruBERT), которые сегодня являются доминирующими в задачах обработки естественного языка.
4. Не указано, насколько предложенные методы требовательны к вычислительным ресурсам и как они масштабируются на большие объёмы текстов.

Итоговая оценка

Диссертационная работа «Совершенствование методов компьютерной обработки текстовой информации в аспекте задач, связанных с омонимией и синонимией» представляет собой законченное научное исследование, в котором успешно решена актуальная научно-техническая задача. Диссертация является научно-квалификационной работой, полностью соответствует установленным требованиям к кандидатским диссертациям, а её автор, Большакова Светлана Анатольевна, заслуживает присуждения ученой степени кандидата технических наук по специальности 2.3.1 – Системный анализ, управление и обработка информации, статистика (технические науки).

Старший научный сотрудник СПб ФИЦ РАН,
кандидат технических наук
доцент

Зайцева Александра Алексеевна

специальность 05.13.01 – «Системный анализ, управление и обработка информации»

Я, Зайцева Александра Алексеевна, даю согласие на автоматизированную обработку моих персональных данных, указанных в отзыве, и размещение их на сайте ФГБОУ ВО ДонНТУ.

25 ноября 2025 г

Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский
Федеральный исследовательский центр Российской академии наук», www.spcras.ru
199178, Санкт-Петербург, 14 линия В.О., дом 39
тел.: (812) 508-33-11 (доб. 1231)
e-mail: cher@iias.spb.su

Подпись руки Зайцевой А.А. Зайцевой А.А. заверяю
Начальник отдела кадров СПб ФИЦ РАН
Сидорова Е.В.
«25» ноября 2025 г.