

УТВЕРЖДАЮ

Директор Федерального государственного  
бюджетного учреждения науки

Институт высшей нервной деятельности и

эмоциональной физиологии РАН

А.Ю. Малышев

2025г.

## ОТЗЫВ ВЕДУЩЕЙ ОРГАНИЗАЦИИ

на диссертацию Большаковой Светланы Анатольевны на тему:

«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)

### Актуальность темы исследования

Современные технологии стремительно продвигаются вперед в сфере понимания естественного языка. Этому способствуют три основных фактора: взрывной рост количества текстовых данных, потребность компаний в автоматизации коммуникации и необходимость извлекать знания из неструктурированной информации.

Проблемы омонимии и синонимии существенно осложняют обработку текстов. Точное распознавание значений способствует улучшению качества машинного перевода, повышению эффективности поисковых систем и созданию удобных интерфейсов взаимодействия с ИИ-сервисами.

Таким образом, совершенствование методов обработки текстовой информации является приоритетным направлением, напрямую влияющим на



развитие современных цифровых технологий и качество пользовательского опыта, а тема диссертационного исследования является актуальной.

## **Связь работы с планом соответствующих отраслей науки и народного хозяйства**

Диссертационная работа выполнена в рамках утверждённой тематики и плана научных исследований федерального государственного бюджетного научного учреждения «Институт проблем искусственного интеллекта» (ФГБНУ «ИПИИ»). Полученные результаты нашли применение при выполнении следующих фундаментальных научно-исследовательских работ:

- «Исследование и разработка методов снятия омонимии в естественно-языковых текстах внутри парадигмы русского слова» (№ Г/Р 0121D000017);
- «Исследование и разработка методов семантического анализа и интерпретации потоков данных интеллектуальными системами» (№ Г/Р 0118D000003);
- «Исследование и разработка методов обработки данных и естественно-языковых текстов с применением онтологий» (государственный регистрационный номер в ЕГИСУ НИОКТР – 123092600030-4).

## **Новизна полученных результатов и выводов**

Научная новизна диссертационного исследования заключается в следующем:

1. Дальнейшее развитие методов снятия частеречной омонимии на основе гибридного подхода, сочетающего декларативные лингвистические знания (в виде словарей) и продукционные правила, позволило достичь высокой точности (99,3%) при анализе сложных конструкций русского языка – предикативов, деепричастий и групп «наречие + существительное». Предложенный подход демонстрирует преимущество перед чисто статистическими методами в условиях ограниченной обучающей выборки и

высокой морфологической неоднозначности.

2. Впервые предложен и реализован метод лексического упрощения текста, основанный на специально размеченной базе синонимов и наборе правил синтаксического контроля. Метод обеспечивает замену лексики с сохранением семантики и грамматической корректности текста, что подтверждается точностью выше 96%. Такой подход не имеет аналогов в отечественной практике и представляет собой существенный вклад в развитие технологий доступного языка и адаптивных образовательных систем.

3. Развиты методы автоматического разбиения текста на семантически однородные фрагменты за счёт введения величины, учитывающей частоту употребления слова и протяжённость его вхождений в текст. Данный метод повышает точность сегментации по сравнению с традиционными подходами и может быть использован в задачах тематического анализа, суммаризации и индексации документов.

### **Апробация работы и публикации**

Основные положения диссертационного исследования отражены в 17 публикациях, включая 2 статьи, опубликованные в рецензируемых научных изданиях, входящих в перечень, утверждённый ВАК РФ, для публикации результатов диссертаций на соискание учёных степеней кандидата и доктора наук. Ещё 7 работ опубликованы в других рецензируемых научных журналах и сборниках. Кроме того, получено свидетельство о государственной регистрации одной программы для ЭВМ.

Результаты диссертационной работы докладывались и обсуждались на авторитетных конференциях и научных семинарах:

– II Всероссийская школа Национального центра физики и математики для студентов, аспирантов, молодых ученых и специалистов по искусенному интеллекту и большим данным в технических, промышленных, природных и социальных системах – г. Саров, 25.11.2024-29.11.2024;

- Донецкий международный научный круглый стол «Искусственный интеллект: теоретические аспекты и практическое применение» – г. Донецк, 30.05.2024;
- Донецкий международный научный круглый стол «Искусственный интеллект: теоретические аспекты и практическое применение» – г. Донецк, 24.05.2023;
- Донецкий международный научный круглый стол «Искусственный интеллект: теоретические аспекты и практическое применение» – г. Донецк, 25.05.2022;
- VII Международная научно-техническая конференция «Современные информационные технологии в образовании и научных исследованиях» – г. Донецк, 23.11.2021;
- Донецкий международный научный круглый стол «Искусственный интеллект: теоретические аспекты и практическое применение» – г. Донецк, 27.05.2021;
- Донецкий международный научный круглый стол «Искусственный интеллект: теоретические аспекты и практическое применение» – г. Донецк, 27.05.2020.

**Обоснованность научных положений и выводов,  
сформулированных в диссертации**

Достоверность и обоснованность результатов и выводов диссертационной работы подтверждаются корректным применением методов исследования, а также реализацией всех предложенных решений в виде работоспособных программных систем. Научные положения, выводы и рекомендации опираются на методологическую базу, построенную на системном подходе. Их обоснованность обеспечивается чёткой постановкой задач, достаточным объёмом и внутренней согласованностью теоретических и экспериментальных данных, а также согласованием полученных результатов с данными, представленными в авторитетных научных источниках по рассматриваемой тематике.

## **Соответствие содержания диссертации автореферату и указанной специальности**

Диссертационная работа соответствует паспорту научной специальности 2.3.1. «Системный анализ, управление и обработка информации, статистика» по ключевым направлениям исследований, что подтверждается как содержанием работы, так и её научными результатами, методологией и практической реализацией:

- п. 3 паспорта специальности, поскольку в ней разработаны и применены модели и критерии оценки эффективности алгоритмов в задачах анализа и интерпретации информации.
- п. 4, поскольку все ключевые результаты – это конкретные методы и алгоритмы, направленные на решение задач обработки информации и искусственного интеллекта.
- п 5, поскольку в ней представлено специальное алгоритмическое и программное обеспечение, предназначенное для систем анализа и интерпретации естественно-языковой информации.

## **Значимость результатов для науки и производства**

Теоретическая значимость полученных научных результатов состоит в развитии методов компьютерной обработки русскоязычных текстов путём создания лингвистических баз знаний, предназначенных для снятия омонимии и обеспечения лексической адаптации.

Практическая ценность работы заключается в создании эффективных методов обработки русскоязычных текстов, которые могут быть использованы в различных прикладных системах. Особенno важным представляется достижение высокой точности разрешения омонимии (99,3%) и эффективности лексической адаптации (более 96%).

Методы компьютерной обработки текстовой информации нашли применение в работе федерального государственного бюджетного научного учреждения "Республиканский академический научно-исследовательский и

проектно-конструкторский институт горной геологии, геомеханики, геофизики и маркшейдерского дела" (ФГБНУ "РАНИМИ") при обработке массивов текстовой информации.

### **Соответствие автореферата основным положениям диссертации**

Автореферат оформлен в соответствии с установленными требованиями и достоверно отражает цели, методы, научную новизну и результаты диссертационного исследования. Все основные положения работы представлены адекватно, без искажений и упрощений.

### **Замечания и предложения по диссертационной работе**

При высокой научной и практической значимости диссертационной работы следует отметить ряд аспектов, требующих уточнения или дополнения:

1. Для повышения достоверности и обобщённости результатов рекомендуется расширить тестовые выборки и провести дополнительные эксперименты на разнородных текстовых корпусах (научных, публицистических, разговорных).

2. В отношении методов автоматической лексической адаптации отсутствует сопоставление с существующими системами. Целесообразно провести косвенное сравнение – например, с универсальными системами упрощения текста (включая адаптированные англоязычные модели) или с методами замены лексики в рамках NLP-задач.

3. Аналитическое сравнение методов снятия омонимии проведено с двумя наиболее популярными и эффективными системами. Однако для более полной оценки конкурентоспособности предложенного подхода желательно расширить круг сравниваемых решений, включив, по возможности, и менее распространённые, но научно обоснованные методы (в том числе гибридные и нейросетевые).

4. Оценка точности некоторых методов снятия омонимии основана на ручной разметке текстов, поскольку Национальный корпус русского языка не поддерживает аннотацию словосочетаний как единой лексической единицы. При этом ручная разметка требует дополнительной валидации. Рекомендуется указать процедуру контроля качества разметки (например, участие нескольких лингвистов, подсчёт меры согласия кодеров и т.п.).

5. В исследовании не рассматриваются методы предварительной обработки текста, в частности, исправление опечаток. Учёт орфографических ошибок является важным фактором для повышения устойчивости систем к реальным текстовым данным. Целесообразно обсудить возможность интеграции модуля коррекции орфографии в общую архитектуру системы.

6. Разработанные методы ограничены в работе с лексикой, отсутствующей в используемых словарях. Это снижает их применимость к динамически изменяющимся или специализированным текстам (например, жаргон, неологизмы, имена собственные). Рекомендуется рассмотреть пути повышения робастности системы – например, за счёт использования методов расширения словарей на основе контекста или встраивания.

Все указанные замечания носят рекомендательный характер и не снижают общей научной ценности работы. Они направлены на дальнейшее развитие представленных методов и повышение их применимости в реальных условиях.

## **Заключение**

Диссертация Большаковой С.А. является законченной научной работой, в которой решена актуальная задача повышения эффективности обработки и анализа текстовой информации в контексте решения задач снятия омонимии и применения способов лексической адаптации. Диссертация отвечает требованиям Положения о присуждении ученых степеней, утвержденного Постановлением Правительства Российской Федерации от 24.09.2013 № 842 с изменениями (ред. от 16.10.2024), а именно пунктам 9, 10, 11, 13, 14,

степени кандидата технических наук по специальности 2.3.1. Системный анализ, управление и обработка информации, статистика (технические науки).

Отзыв обсужден и единогласно одобрен на научном семинаре лаборатории нейроонтогенеза Федерального государственного бюджетного учреждения науки Институт высшей нервной деятельности и нейрофизиологии РАН «11» декабря 2025г., протокол № 1.

старший научный сотрудник  
лаборатории нейроонтогенеза ИВНД и НФ РАН,  
доктор технических наук  
(05.13.01 «Системный анализ, управление  
и обработка информации (по отраслям)»)

Харlamov Александр Александрович

Федеральное государственное бюджетное учреждение науки  
Институт Высшей нервной деятельности и нейрофизиологии РАН  
Адрес: ул. Бутлерова, дом 5А, г. Москва, 117485  
Тел.: +7 (495) 334-70-00  
e-mail: [kharlamov@ihna.ru](mailto:kharlamov@ihna.ru), [kharlamov@analyst.ru](mailto:kharlamov@analyst.ru)  
официальный сайт: <https://ihna.ru>

Я, Харlamов Александр Александрович, даю согласие на автоматизированную обработку моих персональных данных, указанных в отзыве, и размещение их на сайте ФГБОУ ВО «ДонНТУ».

«13» декабря 2025г.

Александр Александрович