

## **Компьютерная лингвистика и интеллектуальные технологии**

По материалам ежегодной международной конференции  
«Диалог» (2025)

Выпуск 23  
Дополнительный том

## **Computational Linguistics and Intellectual Technologies**

Papers from the Annual International Conference “Dialogue” (2025)

Issue 23  
Supplementary volume

Редакционная  
коллегия: *В. П. Селегей (главный редактор), В. И. Беликов, И. М. Богуславский,  
Б. В. Добров, Д. О. Добровольский, И. М. Кобозева, Е. Б. Козеренко,  
М. А. Кронгауз, Н. В. Лукашевич, Д. Маккарти, П. Наков, Й. Нивре,  
В. Раскин, Э. Хови, Т. О. Шаврина, С. А. Шаров, Т. Е. Янко*

**Компьютерная лингвистика** и интеллектуальные технологии: По материалам ежегодной международной конференции «Диалог». Вып. 23, дополнительный том. 2025. С. I–1136.

Сборник включает 9 докладов международной конференции по компьютерной лингвистике и интеллектуальным технологиям «Диалог 2025», представляющих широкий спектр теоретических и прикладных исследований в области описания естественного языка, моделирования языковых процессов, создания практически применимых компьютерных лингвистических технологий.

Для специалистов в области теоретической и прикладной лингвистики и интеллектуальных технологий.

## Предисловие

23-й выпуск ежегодника «Компьютерная лингвистика и интеллектуальные технологии» содержит избранные материалы 30-й международной онлайн-конференции «Диалог». В 2025 году для публикации в дополнительном томе сборника редколлегией были отобраны 9 докладов. Работы, представленные в сборнике, отражают те направления исследований в области компьютерного моделирования и анализа естественного языка, которые по традиции представляются на Диалоге:

- **Интеллектуальный анализ документов;**
- **Лингвистический анализ текста;**
- **Глубокое обучение в компьютерной лингвистике:** методики и примеры применения в лингвистических исследованиях, содержательная интерпретация работы LLM;
- **Компьютерные лингвистические ресурсы:** включая новые датасеты и новые сценарии и типы разметки, Evaluation Benchmarks;
- **Компьютерный анализ Social Media;**
- **Корпусная лингвистика и корпусометрия:** методики создания, использования и оценки корпусов;
- **Компьютерная семантика:** аналитические и дистрибуционные модели, связь между ними;
- **Лингвистические онтологии и автоматическое извлечение знаний;**
- **Мультимодальная коммуникация:** аналитические и нейронные модели речевого акта;
- **Модели общения и диалоговые агенты;**
- **Компьютерная лексикография;**
- **Многоязычие:** language transfer, новые технологии работы с малоресурсными языками.

В соответствии с традициями «Диалога» отбор работ основывается на представлении о важности соединения новых методов и технологий анализа языковых данных с полноценным лингвистическим анализом. Диалог является де-факто крупнейшим форумом по проблемам создания современных компьютерных ресурсов, моделей и технологий для русского языка, поэтому ключевым событием «Диалога» является подведение итогов технологических соревнований между разработчиками систем лингвистического анализа русскоязычных текстов — *Dialogue Evaluation*. В сборнике представлены статьи участников и организаторов 3-х соревнований:

- **CoBaLD Parsing:** Соревнование по автоматической лингвистической разметке;
- **RuOpinionNE-2024:** Соревнование по извлечению мнений из новостных текстов;
- **RuTermEval-2024:** Соревнование по выявлению терминов в научных статьях на русском языке.

Статьи в сборнике публикуются на русском и английском языках. При выборе языка публикации действует следующее правило:

- доклады по компьютерной лингвистике подаются на английском языке. Это расширяет их аудиторию и позволяет привлекать к рецензированию международных экспертов;
- доклады, посвященные лингвистическому анализу русского языка, предполагающие знание этого языка у читателя, подаются на русском языке (с обязательной аннотацией на английском).

Несмотря на традиционную широту тематики представленных на конференции и отобранных в сборник докладов, они не могут дать полной картины направлений «Диалога». Ее можно получить с помощью сайта конференции <https://dialogue-conf.org/>, на котором представлены обширные электронные архивы «Диалогов» последних лет и все результаты проведенных тестирований Dialogue Evaluation.

Мы обращаем внимание авторов и читателей сборника, что с 2018 года Редсовет отказался от печати сборника на бумаге. Все сборники размещаются на сайте конференции. С 2014 года основной том индексируется Scopus.

Программный комитет конференции «Диалог»  
Редколлегия сборника «Компьютерная лингвистика и интеллектуальные технологии»

## Рецензенты

Баюк Александра Михайловна  
Баюк Илья Сергеевич  
Беликов Владимир Иванович  
Богданова-Бегларян Наталья Викторовна  
Богуславский Игорь Михайлович  
Бухаров Ян Михайлович  
Васильев Виталий Геннадьевич  
Галицкий Борис Васильевич  
Добров Борис Викторович  
Добровольский Дмитрий Олегович  
Жарков Андрей  
Ильвовский Дмитрий Алексеевич  
Инькова Ольга Юрьевна  
Киосе Мария Ивановна  
Киселева Ксения Львовна  
Клышинский Эдуард Станиславович  
Клячко Елена Леонидовна  
Князев Сергей Владимирович  
Кобозева Ирина Михайловна  
Козеренко Елена Борисовна  
Коротаев Николай Алексеевич  
Котельников Евгений Вячеславович  
Котов Артемий Александрович  
Кутузов Андрей Борисович  
Лапошина Антонина Николаевна  
Лепехин Михаил

Левонтина Ирина Борисовна  
Лобанов Борис Мефодьевич  
Лукашевич Наталья Валентиновна  
Ляшевская Ольга Николаевна  
Мамонтова Ангелина  
Митрофанова Ольга Александровна  
Мичурина Мария Александровна  
Никишина Ирина Юрьевна  
Пазельская Анна Германовна  
Переверзева Светлана Игоревна  
Петрова Мария Владимировна  
Пиперски Александр Чедович  
Подлеская Вера Исааковна  
Рыгаев Иван Петрович  
Селегей Владимир Павлович  
Смирнов Иван Валентинович  
Смулов Иван Михайлович  
Татевосов Сергей Георгиевич  
Урысон Елена Владимировна  
Федорова Ольга Викторовна  
Феногенова Алена Сергеевна  
Хохлова Мария Владимировна  
Циммерлинг Антон Владимирович  
Шаврина Татьяна Олеговна  
Шаров Сергей Александрович  
Янко Татьяна Евгеньевна

## Contents<sup>1</sup>

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |      |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|
| Бонч-Осмоловская А. А., Гладили С. А., Козеренко А. Д., Ляшевская О. Н., Морозов Д. А., Кузнецова Ю. Н.,<br>Махова А. А., Пискунова С. В., Буйлова Н. Н., Бородина Д. Г., Виноградова И. И., Сизов В. Г., Дьяченко П. В.,<br>Казенников А. О., Власова Н. А., Глазкова А. В., Столяров С. С., Гарипов Т. А., Смаль И. А., Губарькова Я. Н.<br><b>Национальный корпус русского языка 2.0: корпусная платформа, инструменты анализа,<br/>нейросетевые модели разметки данных (полная версия)</b> ..... | 1001 |
| Евдокимова А. А.<br><b>Логическое ударение и синонимия жестов в цефалическом канале</b> .....                                                                                                                                                                                                                                                                                                                                                                                                        | 1043 |
| Kurtukova A., Kozachenko A.<br><b>Backtranslation Invariance Boosts Effectiveness of Non-English Prompts</b> .....                                                                                                                                                                                                                                                                                                                                                                                   | 1055 |
| Levikin A., Khabutdinov I., Grabovoy A., Vorontsov K.<br><b>The methodology of multi-criteria evaluation of text markup models<br/>based on inconsistent expert markup</b> .....                                                                                                                                                                                                                                                                                                                     | 1066 |
| Лурия А. С., Котов А. А.<br><b>Система междометных реакций для поддержания коммуникации роботом-компаньоном</b> .....                                                                                                                                                                                                                                                                                                                                                                                | 1081 |
| Пересыпкина К. А., Богданова-Бегларян Н. В.<br><b>Прецедентные тексты корпуса «один речевой день»<br/>и комические паспорты повседневной коммуникации</b> .....                                                                                                                                                                                                                                                                                                                                      | 1091 |
| Petrulina U., Zdorova N.<br><b>Readability assessment of written Adyghe using a baseline approach</b> .....                                                                                                                                                                                                                                                                                                                                                                                          | 1100 |
| Сулейманова Е. А., Момот С. Р., Власова Н. А., Воздвиженский И. Н.<br><b>К задаче оценки субъективной достоверности</b> .....                                                                                                                                                                                                                                                                                                                                                                        | 1110 |
| Tatarinov M., Demidovsky A.<br><b>Interpretable approach to detecting semantic changes based on generated definitions</b> .....                                                                                                                                                                                                                                                                                                                                                                      | 1123 |
| <b>Abstracts</b> .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 1133 |
| <b>Авторский указатель</b> .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 1135 |
| <b>Author Index</b> .....                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 1135 |

---

\* The reports of each section are ordered by the surname of the first author in compliance with the English alphabet.