



ОБЗОРЫ
REVIEWS

DOI: 10.22363/2312-9220-2026-31-1-265-276
EDN: TXFUFZ
УДК 81-139

Обзорная статья / Review

**Анализ тональности в исследованиях
массовой коммуникации: систематический обзор
методов и направлений**

А.В. Козлова 

*Национальный исследовательский Нижегородский государственный университет
им. Н.И. Лобачевского, Нижний Новгород, Российская Федерация*
✉ anastasia_kozlova94@mail.ru

Аннотация. Представлен аналитический обзор современных зарубежных исследований, применяющих методы компьютерной лингвистики для изучения массовой коммуникации. Основное внимание уделяется методологии анализа тональности как ключевого инструмента измерения коммуникативного воздействия. На основе репрезентативного корпуса публикаций за последние пять лет систематизированы четыре ключевых тематических направления: анализ реакций на глобальные события (пандемия, военные конфликты), изучение механизмов медиавоздействия, исследование языка вражды и дискриминации, а также измерение суггестивного влияния в политической и социальной коммуникации. Особое внимание уделено сравнительному анализу методов – от традиционных лексических подходов (VADER, LIWC) до современных архитектур на основе трансформеров (BERT). Проведенный анализ выявил устойчивый методологический разрыв между техническим совершенством алгоритмов и глубиной их социогуманитарной интерпретации. Обсуждаются актуальные проблемы области, включая системные смещения в языковых моделях и ограничения автоматического анализа для изучения сложных форм коммуникативного воздействия. Намечаются перспективы развития междисциплинарного подхода, сочетающего компьютерную лингвистику с теориями медиа и коммуникации.

Ключевые слова: большие данные, средства массовой информации, коммуникация, обработка естественного языка, коммуникативное воздействие, алгоритмическое смещение

Заявление о конфликте интересов. Автор заявляет об отсутствии конфликта интересов.

История статьи: поступила в редакцию 16 июня 2025 г.; отрецензирована 31 августа 2025 г.; принята к публикации 28 октября 2025 г.

© Козлова А.В., 2026



This work is licensed under a Creative Commons Attribution 4.0 International License
<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

Для цитирования: Козлова А.В. Анализ тональности в исследованиях массовой коммуникации: систематический обзор методов и направлений // Вестник Российского университета дружбы народов. Серия: Литературоведение. Журналистика. 2026. Т. 31. № 1. С. 265–276. <http://doi.org/10.22363/2312-9220-2026-31-1-265-276> EDN: TXFUFZ

Sentiment Analysis in Mass Communication Research: A Systematic Review of Methods and Approaches

Anastasia V. Kozlova 

Lobachevsky University, Nizhny Novgorod, Russian Federation

✉ anastasia_kozlova94@mail.ru

Abstract. An analytical review of contemporary international research applying computational linguistics methods to the study of mass communication is presented. Particular focus is placed on sentiment analysis methodology as a key tool for measuring communicative influence. Based on a representative corpus of publications from the last five years, four key thematic areas are systematized: the analysis of reactions to global events (pandemics, military conflicts), the study of media effects mechanisms, research on hate speech and discrimination, and the measurement of suggestive influence in political and social communication. Special attention is paid to the comparative analysis of methods – from traditional lexicon-based approaches (VADER, LIWC) to modern transformer-based architectures (BERT). The conducted analysis reveals a persistent methodological gap between the technical sophistication of the algorithms and the depth of their socio-humanitarian interpretation. Current challenges in the field are discussed, including systemic biases in language models and the limitations of automated analysis for studying complex forms of communicative influence. Prospects for developing an interdisciplinary approach that integrates computational linguistics with media and communication theories are outlined.

Keywords: big data, mass media, communication, natural language processing, communicative influence, algorithmic bias

Conflicts of interest. The author declares that there is no conflict of interest.

Article history: submitted June 16, 2025; revised August 31, 2025; accepted October 28, 2025.

For citation: Kozlova, A.V. (2026). Sentiment Analysis in Mass Communication Research: A Systematic Review of Methods and Approaches. *RUDN Journal of Studies in Literature and Journalism*, 31(1), 265–276. (In Russ.) <http://doi.org/10.22363/2312-9220-2026-31-1-265-276> EDN: TXFUFZ

Введение

Цифровая трансформация медиа создает для исследователей новые методологические, этические и технологические вызовы, обусловленные в том числе и необходимостью анализа больших объемов информации. В настоящее время анализ цифровой коммуникации часто производится на материале больших данных (цифровых социальных платформ – социальных сетей, социальных медиа), содержащих миллионы текстов, которые не поддаются об-

работке традиционными методами. Такие *большие данные* характеризуются не только большими объемами неструктурированной или мало структурированной информации, но и динамичностью, разнообразием содержания. В такой ситуации для анализа текстовых данных исследователи обращаются к методам компьютерной лингвистики и обработки естественного языка.

Действительно, современные цифровые методы – как традиционные методы статистического моделирования, так и state-of-the-art алгоритмы машинного обучения – могут обеспечить высокий уровень точности при обработке и структурировании больших данных, но зачастую такой анализ не сопровождается адекватными средствами теоретического осмысления полученных результатов, что ведет к утрате целостного понимания исследуемых социальных процессов, выходящих за рамки сугубо количественных описаний. Цель статьи – проанализировать в зарубежных междисциплинарных исследованиях использование методов компьютерной лингвистики для изучения медиа и коммуникаций на основе больших массивов данных. В частности, рассматривается роль *анализа тональности текста* как инструмента для идентификации и изучения феноменов массовой коммуникации.

Для настоящего обзора был проведен отбор репрезентативных статей за 2020–2025 гг. в базах данных Scopus и Web of Science по следующим ключевым словам: sentiment analysis, computational linguistics, NLP, topic modelling, digital mass communication, social media analysis, social network analysis. Основным критерием отбора являлась их релевантность применению анализа тональности для изучения механизмов воздействия в медиадискурсе и социальных феноменов. В итоговую выборку вошли наиболее цитируемые и методологически показательные работы (25 публикаций).

Теоретическое обоснование

Среди основных направлений исследований массовой коммуникации, использующих sentiment-анализ (метод, направленный на определение эмоциональной окраски текста), тематическое моделирование (автоматическое определение тем в тексте) или иные методы компьютерной лингвистики, в первую очередь можно выделить анализ реакций на глобальные события, изучение механизмов воздействия, агрессии и языка вражды. Тематически эти стратегии анализа, как правило, локализованы в четырех областях.

1. *Изучение мировых событий.* Анализируются реакции и обсуждения значимых общественных событий, например Евровидения, Олимпиады, спортивных состязаний, международных конференций, Чемпионата мира по футболу (She et al., 2023). В этом контексте рассматриваются различные медийные феномены, в частности поляризации при обсуждении значимых социальных тем: вакцинации, пластической хирургии (Ramnarine, 2023). В связи с пандемией COVID-19 фиксировалась динамика общественного мнения об этой проблеме, вопросы мотивации и убеждений (Akhter, Kanojia, 2023; Buckman et al., 2020; Silva et al., 2023). Так, в исследовании (Akhter, Kanojia, 2023) с помощью

анализа тональности изучались эмоциональные реакции, выраженные в комментариях о пандемии и вакцинации (страх, гнев, осторожность и т.п.).

2. *Воздействие массмедиа на аудиторию.* Исследования массовой коммуникации с опорой на компьютерные лингвистические методы направлены на изучение механизмов воздействия в сфере массовой коммуникации, как правило новостных текстов (Kutuzov et al., 2020). Методы тематического моделирования используются для исследования тематической структуры новостных корпусов, изучения корреляций в темах и настроениях между новостями и реакциями читателей на них (Jiang, Everts, 2021). Изучается поведение акторов фондового рынка в зависимости от новостных поводов и их обсуждений в социальных сетях (Ge et al., 2020; Sinha et al., 2022).

3. *Агрессия, доминирование, язык вражды.* Изучается распространение вербальной агрессии в социальных сетях, проявления толерантности, дискриминации и фобий, профессиональных, гендерных, расовых стереотипов. К примеру, в работе (Shen et al., 2022) проанализировано 8 млрд постов на Reddit и 206 млн постов на 4chan с целью анализа истоков, эволюции и содержания синофобии. Распространены исследования, посвященные проявлениям антисемитизма и исламофобии (He et al., 2021; Risch, Krestel, 2020; Zannettou et al., 2020), дискриминации людей с ограниченными возможностями (Venkit et al., 2023), гендерной дискриминации (Bhaskaran, Bhallamudi, 2019). К примеру, в работе (Ramnarine, 2023) был использован лексический, синтаксический и сентимент-анализ, чтобы исследовать роль пола в персонализации широкого круга политических чиновников.

4. *Косвенное (суггестивное) воздействие массмедиа.* Анализ тональности текста находит применение в изучении предвыборных кампаний и речей (Sharma, Shekhar 2020), коммуникации для принятия муниципальных решений (Švaňa, 2023). В контексте таких исследований особое внимание уделяется лидерству. В работе (Mohapatra, Mohapatra, 2022), к примеру, методы обработки естественного языка использовались для изучения речей и настроений кандидатов от республиканцев и демократов, боровшихся за президентские выборы в США в 2020 г.

Количественный и качественный анализ также был применен к общественным обсуждениям российско-украинского конфликта 2022 г. в социальной сети «Твиттер» (Caprolu et al., 2022; Guerra, Karakuş, 2023; Thakkar et al., 2023). В исследовании (Caprolu et al., 2022) с помощью методов статистического и аспектного сентимент-анализа было проанализировано более чем 5,5 млн твитов по данной теме, созданных более чем 1,8 млн пользователей. Необходимо отметить, что для анализа тональности все еще не существует единого устоявшегося и предпочитаемого способа: эта область все еще развивается, и компьютерная лингвистика предлагает несколько подходов к анализу тональности, от выбора которых могут зависеть его результаты и качество.

Представленный обзор существующих направлений показывает широкий спектр применения анализа тональности, однако уже на этом этапе становится очевидным, что сложность социальных феноменов (например, суггестии) требует рефлексивного подхода к выбору исследовательских методов.

Материалы и методы

Сентимент-анализ, или анализ тональности (sentiment analysis), представляет собой подход к обработке естественного языка, в общем случае направленный на определение эмоционального тона текста (Лукашевич, 2022; Kazyulina et al., 2021). Этот метод позволяет определить и классифицировать мнение индивида/группы о продукте, услуге или какой-либо идее, определить настроение высказывания, наличие субъективной информации. Чаще всего в текстах выделяют положительную, отрицательную или нейтральную тональности. Этот метод успешно применяется для анализа онлайн-источников: социальных сетей, электронных писем, блогов, новостных статей, веб-чатов и т.д.

К примеру, в исследовании (Wright et al., 2022), посвященном изучению общественного мнения о пандемии COVID-19, было рассмотрено более 58 тыс. твитов и более 60 академических статей с целью выяснить, к каким поведенческим и эмоциональным изменениям привело распространение вируса, какие сентименты были выражены общественностью разных стран по отношению к пандемии, какие темы возникли и аргументы появились по сравнению с эпидемиями предыдущих лет. Анализ облака слов, согласно исследованию, выявил заметное несоответствие между эмоциями и опасениями людей во время кризиса COVID-19 и прошлых вспышек эпидемий (Эбола, Зика, ВИЧ, Денге, корь). Анализ тональности показал большое количество нейтральных, положительных и смешанных настроений, в то время как сугубо отрицательных оценок оказалось гораздо меньше. Авторы выявили высокую склонность к оптимистическим настроениям (о чем говорит большое количество твитов с нейтральными и смешанными настроениями и оценками) и поведенческий сдвиг от негативных настроений к оптимистическим. Была также проведена кластеризация пользователей «Твиттера» на основе выраженных ими сентиментов, эмоций и опасений, в результате которой было замечено, что группировка пользователей в целом соответствует ходу распространения COVID-19.

Можно выделить несколько видов сентимент-анализа. Так называемый детальный сентимент-анализ (fine-grained) предполагает разбиение тональности на более точные категории: крайне негативная, негативная, нейтральная, позитивная, крайне позитивная. Следующий тип предполагает определение эмоций, а не тональности: удовлетворение, разочарование, гнев, радость, печаль и т.д. Анализ, основанный на намерениях (intent-based), помимо собственно сентимента, также распознает намерения автора.

Часто применяется аспектный анализ тональности (aspect-based sentiment analysis), когда требуется определить тональность по отношению к определенному аспекту объекта, о котором говорится в тексте. Например, можно привести случаи в задаче изучения общественного мнения о вакцинации, когда авторы комментариев или постов выражали в общем положительную

тональность по отношению к вакцинации, но негативно высказывались об организации/точках проведения самой процедуры или конкретных видах вакцины (Buckman et al., 2020). При таком варианте анализа требуется не только определить тональность, но и выделить объекты и аспекты – для этого часто применяются методы классификации токенов для выделения именованных сущностей.

Существуют три группы подходов к анализу тональности текстов: подходы, основанные на правилах; использующие модели машинного обучения; гибридные решения.

Правиловый (rule-based) подход опирается на словари тональности и оценочной лексики и синтаксический анализ. В общем случае по определенным правилам подсчитывается количество положительных и отрицательных слов в тексте или предложении. Также используется система правил для учета влияния модификаторов, отрицания и маркеров ирреального наклонения. Очевидно, что составить исчерпывающий набор правил достаточно сложно, поэтому такой подход часто не достигает желаемого качества. С другой стороны, такие подходы быстро работают, не требуют больших корпусов и обучения, хорошо интерпретируемы, в отличие от нейронных сетей (Kotelnikova et al., 2021). Среди наиболее часто используемых словарей можно назвать SentiStrength, VADER, Pattern, TextBlob, SentiWordNet, Opinion Finder (Elbagir, Yang, 2019).

Результаты и обсуждение

Подход, основанный на словарях и частотности слов, использовался в исследовании феномена «кэнселинг» в социальных сетях (Maryn, Dover, 2023). Рассматривался кэнселинг («культура отмены», публичное высмеивание, осуждение, игнорирование), направленный на лиц, обвиняемых в гендерном насилии. Было выяснено, что не все предполагаемые виновники гендерного насилия оцениваются пользователями «Твиттера» одинаково, то есть «кэнселинг» может быть неравномерным, а реакции и оценки, связанные с ним, – более неоднозначными и временными. Авторы использовали правилый инструмент для сентимент анализа LIWC на более чем 182 тыс. твитов об обвинениях о 120 случаях гендерного насилия. Они изучали в том числе и изменения в дискуссиях о предполагаемых виновниках гендерного насилия с течением времени и выяснили, что сентимент обсуждений будет различаться в зависимости от характеристик обвинителя (возраст) и обвиняемого (раса, известность). Например, твиты о предполагаемых виновниках со временем стали содержать больше положительного сентимента, сообщения, обсуждающие предполагаемых чернокожих (по сравнению с белыми) преступников, были менее позитивными, в то время как твиты, обсуждающие предполагаемых цветных преступников (по сравнению с белыми), не принадлежащих к чернокожим, были более позитивными. Известные личности обсуждались с более высоким уровнем как положительных, так и отрицательных эмоций (по сравнению с обычными людьми или лицами, не известными авторам твита).

В статье (Silva et al., 2023) изучается, как ведутся дискуссии о вакцинации на дискуссионном форуме Mumsnet, британского веб-сайта, предназначенного для родителей, с применением сетевого анализа и анализа тональности текста, в том числе в диахронии. Результаты данного исследования обращаются, в частности, к понятию нонконформности при обсуждении вакцинации. Исследователи использовали подход, основанный на словарях оценочной лексики и правилах их сочетания – анализатор VADER, классифицируя каждый пост на платформе в трех вариантах: подсчет сентимента для всего текста в целом; для каждого предложения текста; для каждого предложения, в котором содержатся ключевые слова, относящиеся к вакцинации. Авторы изучили, насколько пользователи склонны оставлять посты в тредах с разным типом сентимента, и рассматривали такое поведение как признак того, что информация передается между обсуждениями с разными сентиментами. Нейтральный сентимент был наиболее частотным (как правило, это общие вопросы о вакцинации), за ним – негативный (обсуждение негативных последствий, тревоги, связанные с вакцинацией и пандемией).

Подходы к анализу тональности, использующие машинное обучение с учителем, являются, пожалуй, самыми распространенными на данный момент в индустрии и менее распространенными – в социальных исследованиях. Это связано в первую очередь с проблемой интерпретируемости моделей машинного обучения: часто такие модели представляют собой своего рода черный ящик и не всегда можно однозначно объяснить, как было получено то или иное решение.

Для классификации текстов по тональности используются как классические алгоритмы (метод опорных векторов или метод случайного леса), так и глубокое обучение (рекуррентные нейронные сети, сети долгой краткосрочной памяти). В данный момент на передний край вышли модели машинного обучения на основе архитектуры – трансформеры.

Трансформер – это относительно новая архитектура, способная решать задачи анализа последовательностей (текстов), при этом легко обрабатывающая зависимости на больших расстояниях. Вычисление представлений в такой архитектуре происходит без использования рекуррентных слоев или сверток и полагается на так называемый механизм «внимания» (Vaswani et al., 2017).

Для сентимент-анализа чаще всего используется модель BERT (Bidirectional Encoder Representations from Transformers) (Devlin et al., 2019) и ее варианты. Эта модель представляет собой энкодер и предварительно обучается на двух задачах: задаче маскированного моделирования языка и прогнозировании следующего предложения. Англоязычная версия модели обучалась на основе материалов всей Википедии и корпусе Brown стандартного американского английского языка. Исходная архитектура BERT вдохновила исследователей на создание множества ее вариаций (RoBERTa, DistilBERT, BART и т. д.). Существует также ряд моделей, обученных на корпусе определенной узкой тематики, например ClinicalBERT и BioBERT, и многоязычные модели (mBART).

Разные языковые модели (ELMo, Universal Sentence Encoder, GPT2) использовались (Gaur et al., 2021) для создания контекстуальных эмбедингов постов/комментариев о COVID-19 для последующей классификации авторов на тех, кто ищет помощь (support seekers) и тех, кто готов ее предоставить (support providers). Получившиеся вектора-представления текстов подавались в классификаторы – логистическую регрессию и метод опорных векторов.

В работах, исследующих сентимент, эмоции, оценочность в массовых коммуникациях в интернете компьютерными методами, собственно анализ социальных явлений иногда уходит на второй план, а основной фокус исследователей приходится на создание корпуса текстов и выбор архитектуры модели. Например, авторы исследования (Singh et al., 2021) использовали модель BERT для классификации эмоций в постах о COVID-19. Исследователи собрали большой корпус твитов о COVID-19, произвели их очистку и предобработку, выбор и конструирование признаков текстов для использования в алгоритме, выбор метрик для оценки качества работы алгоритма классификации. Впоследствии исследователи предложили некоторый анализ эмотивности, субъективности и интенсивности сентимента в постах о болезни, но основным вкладом исследования все же представляется подход к анализу тональности таких текстов и его качество, то есть вклад практический. Недостаточность методологической обоснованности и теоретического вклада с точки зрения социального анализа видится одной из проблем таких эмпирических исследований.

Кроме того, большие предобученные языковые модели, поскольку они обучаются на текстах, написанных людьми, зачастую обучаются воспроизводить социальные стереотипы и предрассудки, содержащиеся в этих текстах (Blodgett et al., 2020; Lucy, Vanman, 2021). Ярким примером являются гендерные стереотипы. Например, было установлено, что BERT кодирует установки по отношению к таким словам, как firewoman/fireman/firefighter, какие свойственны людям умеренных и консервативных взглядов на социальные роли мужчин и женщин (Watson et al., 2023). Как следствие, возникает вопрос, насколько вообще приемлемо прибегать к таким «предвзятым» моделям в исследованиях.

Заключение

Таким образом, методы компьютерной лингвистики (тематическое моделирование, кластеризация, классификация, анализ тональности, извлечение именованных сущностей) позволяют изучать и систематизировать большие объемы текстов. В связи с возрастающим объемом данных, используемых в прикладных исследованиях, видится необходимым и возможным дополнять результаты качественного анализа количественным.

Однако во многих исследованиях делается фокус непосредственно на создание корпуса текстов и улучшение алгоритма и архитектуры модели, при этом часто сам анализ социальной коммуникации и социальных явлений ухо-

дит на второй план или рассматривается как некий побочный продукт. Современные компьютерные методы позволяют извлекать закономерности из больших массивов данных, но далеко не всегда это сопровождается интерпретацией в терминах социальных смыслов и структур, политических контекстов и т.д. В этом контексте представляется важным развитие гибридных исследовательских компетенций, формирование интердисциплинарных команд, состоящих из специалистов как технического (более прикладного), так и теоретического профиля.

Также представляется, что акцент на развитие интерпретируемого машинного обучения, использование нарративных форм визуализации данных, data storytelling, повысил бы доверие к цифровым методам в гуманитарных и социальных контекстах и в целом качество исследований, снизив риск поверхностной интерпретации результатов компьютерного моделирования. Помимо этого, при проведении исследований необходимо иметь представление об ограничениях выбранного метода (например, склонности модели к смещению или галлюцинациям) и избегать использования алгоритмов как черного ящика, результаты работы которого принимаются как данность.

References / Список литературы

- Akhter, M.M., & Kanojia, D. (2023). COVID-19 public sentiment analysis for Indian Tweets classification. *ArXiv*, 2308.06241. <https://doi.org/10.48550/arXiv.2308.06241>
- Bhaskaran, J., & Bhallamudi, I. (2019). Good secretaries, bad truck drivers? Occupational gender stereotypes in sentiment analysis. In M.R. Costa-Jussà, C. Hardmeier, W. Radford, & K. Webster (Ed.), *Proceedings of the First Workshop on Gender Bias in Natural Language Processing* (pp. 62–68). Florence, Italy. Association for Computational Linguistics. <https://doi.org/10.18653/v1/W19-3809>
- Blodgett, S.L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: a critical survey of “Bias” in NLP. In D. Jurafsky, J. Chai, N. Schlueter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5454–5476). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.485>
- Buckman, S.R., Shapiro, A.H., Sudhof, M., & Wilson, D.J. (2020). News sentiment in the time of COVID-19. *FRBSF Economic Letter*, 8(1), 5–10. <https://www.frbsf.org/wp-content/uploads/el2020-08.pdf>
- Caprolu, M., Sadighian, A., & Di Pietro, R. (2022). Characterizing the 2022 Russo-Ukrainian conflict through the lenses of aspect-based sentiment analysis: dataset, methodology, and preliminary findings. *ArXiv*, 2208.04903. <https://doi.org/10.48550/arXiv.2208.04903>
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. In J. Burstein, C. Doran, & T. Solorio (Eds.), *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 4171–4186). Association for Computational Linguistics. <https://doi.org/10.18653/v1/N19-1423>
- Elbagir, S., & Yang, J. (2019). Twitter sentiment analysis using natural language toolkit and VADER sentiment. In *Proceedings of the International Multiconference of Engineers and Computer Scientists, March 13–15, 2019, Hong Kong*. Hong Kong. https://www.iaeng.org/publication/IMECS2019/IMECS2019_pp12-16.pdf

- Gaur, M., Roy, K., Sharma, A., Srivastava, B., & Sheth, A. (2021). Who can help me?: Knowledge infused matching of support seekers and support providers during COVID-19 on Reddit. In *2021 IEEE 9th International Conference on Healthcare Informatics, August 9, 2021, Victoria, Canada* (pp. 265–269). IEEE. <https://doi.ieeecomputersociety.org/10.1109/ICHI52183.2021.00049>
- Ge, Y., Qiu, J., Liu, Z., Gu, W., & Xu, L. (2020). Beyond negative and positive: Exploring the effects of emotions in social media during the stock market crash. *Information Processing & Management*, 57(4), 102218. <https://doi.org/10.1016/j.ipm.2020.102218>
- Guerra, A., & Karakuş, O. (2023). Sentiment analysis for measuring hope and fear from Reddit posts during the 2022 Russo-Ukrainian conflict. *Frontiers in Artificial Intelligence*, 6, 1163577. <https://doi.org/10.3389/frai.2023.1163577>
- He, B., Ziems, C., Soni, S., Ramakrishnan, N., Yang, D., & Kumar, S. (2021). Racism is a virus: Anti-Asian hate and counterspeech in social media during the COVID-19 crisis. *ArXiv*, 2306.16049. <https://doi.org/10.48550/arXiv.2005.12423>
- Jiang, X., & Everts, J. (2021). Making sense of electrical vehicle discussions using sentiment analysis on closely related news and user comments. *ArXiv*, 2112.12327. <https://doi.org/10.48550/arXiv.2112.12327>
- Kazyulina, M., Babii, A., & Malafeev, A. (2021). Emotion classification in Russian: feature engineering and analysis. In W.M.P. van der Aalst, V. Batagelj, D.I. Ignatov, et al. (Eds.), *Analysis of Images, Social Networks and Texts: 9th International Conference, AIST 2020, October 15–16, 2020, Skolkovo, Moscow, Russia* (Vol. 12602, pp. 73–84). Springer.
- Kotelnikova, A., Paschenko, D., Bochenina, K., & Kotelnikov, E. (2021). Lexicon-based methods vs. BERT for text sentiment analysis. In E. Burnaev, D.I. Ignatov, S. Ivanov, et al. (Eds.), *International Conference on Analysis of Images, Social Networks and Texts, December 16–18, 2021, Tbilisi, Georgia* (pp. 71–83). Springer International Publishing. https://doi.org/10.1007/978-3-031-16500-9_7
- Kutuzov, A., Fomin, V., Mikhailov, V., & Rodina, J. (2020). Shiftry: Web service for diachronic analysis of Russian news. In *Computational Linguistics and Intellectual Technologies. Papers from the Annual International Conference “Dialogue”, June 17–20, 2020, Moscow* (Vol. 19, pp. 500–516). Russian State Technological University Publ. <https://doi.org/10.28995/2075-7182-2020-19-500-516>
- Lucy, L., & Bamman, D. (2021). Gender and representation bias in GPT-3 generated stories. In *Proceedings of the Third Workshop on Narrative Understanding, June 11, 2021, Mexico City* (pp. 48–55). Association for Computational Linguistic. <https://doi.org/10.18653/v1/2021.nuse-1.5>
- Lukashevich, N.V. (2022). Automatic sentiment analysis of texts: problems and methods *Intelligent Systems. Theory and Applications*, 26(1), 50–61. (In Russ.) EDN: ПБЫWS
Лукашевич Н.В. Автоматический анализ тональности текстов: проблемы и методы // Интеллектуальные системы. Теория и приложения. 2022. Т. 26. № 1. С. 50–61. EDN: ПБЫWS
- Maryn, A.G., & Dover, T.L. (2023). Who gets canceled? Twitter responses to gender-based violence allegations. *Psychology of Violence*, 13(2), 117–126. <https://doi.org/10.1037/vio0000436>
- Mohapatra, S., & Mohapatra, S. (2022). Sentiment is all you need to win US Presidential elections. In *Proceedings of the 2nd International Workshop on Natural Language Processing for Digital Humanities, November 20, 2022* (pp. 15–20). Association for Computational Linguistics. <https://doi.org/10.48550/arXiv.2209.13487>
- Ramnarine, A.K. (2023). Unsupervised sentiment analysis of plastic surgery social media posts. *ArXiv*, 2307.02640. <https://doi.org/10.48550/arXiv.2307.02640>
- Risch, J., & Krestel, R. (2020). Toxic comment detection in online discussions. In B. Agarwal, R. Nayak, N. Mittal, S. Patnaik (Eds.), *Deep Learning-Based Approaches for Sentiment Analysis* (pp. 85–109). Springer, Singapore. https://doi.org/10.1007/978-981-15-1216-2_4

- Sharma, A., & Shekhar, H. (2020). Intelligent learning based opinion mining model for governmental decision making. *Procedia Computer Science*, 173, 216–224. <https://doi.org/10.1016/j.procs.2020.06.026>
- She, J., Swart-Arries, K., Belal, M., & Wong, S. (2023). What sentiment and fun facts we learnt before FIFA World Cup Qatar 2022 using Twitter and AI. *ArXiv*, 2306.16049. <https://doi.org/10.48550/arXiv.2306.16049>
- Shen, X., He, X., Backes, M., Blackburn, J., Zannettou, S., & Zhang, Y. (2022). On Xing Tian and the perseverance of anti-China sentiment online. In *Proceedings of the Sixteenth International AAAI Conference on Web and Social Media, June 6–9, 2022, Atlanta, Georgia, USA* (Vol. 16, pp. 944–955). AAAI Press. <https://doi.org/10.48550/arXiv.2204.08935>
- Silva, M., Skeva, R., House, T., & Jay, C. (2023). Tracking the structure and sentiment of vaccination discussions on Mumsnet. *Social Network Analysis and Mining*, 13,(1) 152. <https://doi.org/10.1007/s13278-023-01155-z>
- Singh, M., Jakhar, A.K., & Pandey, S. (2021). Sentiment analysis on the impact of coronavirus in social life using the BERT model. *Social Network Analysis and Mining*, 11(1), 33. <https://doi.org/10.1007/s13278-021-00737-z>
- Sinha, A., Kedas, S., Kumar, R., & Malo, P. (2022). SEntFiN 1.0: Entity-aware sentiment analysis for financial news. *Journal of the Association for Information Science and Technology (JASIST)*, 73(9), 1314–1335. <https://doi.org/10.1002/asi.24634>
- Švaňa, M. (2023). Social media, topic modeling and sentiment analysis in municipal decision support. In M. Ganzha, L. Maciaszek, M. Paprzycki, & D. Ślęzak (Eds.), *Proceedings of the 18th Conference on Computer Science and Intelligence Systems* (Vol. 35, pp. 1235–1239). IEEE. <https://doi.org/10.15439/2023F1479>
- Thakkar, H., Patil, A., Saudagar, O., & Yenikar, A. (2023). Sentiment and statistical analysis on custom Twitter dataset for 2022 Russo-Ukrainian conflict. In *International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), January 27–28, 2023* (pp. 679–684). IEEE. <https://doi.org/10.1109/IITCEE57236.2023.10090995>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. In I. Guyon, U. von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *Advances in Neural Information Processing Systems 30, December 4–9, 2017, Long Beach, California, USA*. Curran Associates, Inc. <https://doi.org/10.48550/arXiv.1706.03762>
- Venkit, P., Srinath, M., & Wilson, S. (2023). Automated ableism: An exploration of explicit disability biases in sentiment and toxicity analysis models. In A. Rogers, J. Boyd-Graber, & N. Okazaki (Eds.), *The 61st Annual Meeting of the Association for Computational Linguistics, July 9–14, 2023, Toronto, Canada*. Association for Computational Linguistics. <https://doi.org/10.48550/arXiv.2307.09209>
- Watson, J., Beekhuizen, B., & Stevenson, S. (2023). What social attitudes about gender does BERT encode? Leveraging insights from psycholinguistics. In A. Rogers, J. Boyd-Graber, & N. Okazaki (Eds.), *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics, July 9–14, 2023, Toronto, Canada*. Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.acl-long.375>
- Wright, L., Fluharty, M., Steptoe, A., & Fancourt, D. (2022). How did people cope during the COVID-19 pandemic? A structural topic modelling analysis of free-text data from 11,000 United Kingdom adults. *Frontiers in Psychology*, 13, 810655. <https://doi.org/10.3389/fpsyg.2022.810655>
- Zannettou, S., Finkelstein, J., Bradlyn, B., & Blackburn, J. (2020). A quantitative approach to understanding online antisemitism. In *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media, June 8, 2020, Atlanta, Georgia, USA (virtual)* (Vol. 14, pp. 786–797). Association for the Advancement of Artificial Intelligence. <https://doi.org/10.48550/arXiv.1809.01644>

Сведения об авторе:

Козлова Анастасия Владимировна, аспирант кафедры общей и социальной психологии, Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского, Российская Федерация, 603950, Нижний Новгород, пр. Гагарина, д. 23. ORCID: 0000-0002-2629-9157; SPIN-код: 2424-7595. E-mail: anastasia_kozlova94@mail.ru

Bio note:

Anastasia V. Kozlova, PhD Student at the Department of General and Social Psychology, Lobachevsky University, 23 Gagarin Ave, Nizhny Novgorod, 603950, Russian Federation. ORCID: 0000-0002-2629-9157; SPIN-code: 2424-7595. E-mail: anastasia_kozlova94@mail.ru