

# ЯЗЫК МЕДИА И КОМПЬЮТЕРНЫЕ ТЕХНОЛОГИИ / MEDIA LANGUAGE AND COMPUTER TECHNOLOGIES

## Новостной текст и искусственный интеллект

*Д. Д. Александров*

### News text and artificial intelligence

*D. D. Aleksandrov*

Даниил Дмитриевич Александров – аспирант; Новгородский государственный университет имени Ярослава Мудрого, Великий Новгород, Российская Федерация

E-mail: dalieksandrov6@gmail.com

Статья поступила: 15.11.2024. Принята к печати: 20.12.2024.

В статье рассматривается влияние роботизации на трансформацию новостного текста в современной журналистике. Исследование включает обзор научной литературы об использовании искусственного интеллекта и обзор успешных кейсов его внедрения в крупных медиакомпаниях, таких как Bloomberg и Associated Press. На материале работы российских региональных изданий анализируется характер действий, формирующих контент новости из текстов иного типа (полицейских сводок, жалоб горожан и т.д.), а также анализ типичных сбоев, возникающих при использовании алгоритмов искусственного интеллекта, среди которых наиболее подробно описываются сбои в типографике, формировании смысла отдельных высказываний текста и грамматике. Приведенные факты говорят о неизбежности включения искусственного интеллекта в инструментарий медиа и необходимости внимательного отношения к результатам его использования.

**Ключевые слова:** искусственный интеллект, медиа, генерация текстов, обработка данных, достоверность информации, сбои, типографика, грамматика

УДК 070.1:004.8

Daniil D. Aleksandrov – postgraduate student; Yaroslav-the-Wise Novgorod State University, Veliky Novgorod, Russian Federation

ORCID: 0009-0004-1422-8716

Received: 15.11.2024. Accepted for publication: 20.12.2024.

The paper examines the impact of robotization on the transformation of news text in modern journalism. The study includes a review of scientific literature on the use of artificial intelligence (AI) and a review of successful cases of its implementation in large media companies such as Bloomberg and Associated Press. Based on the work of Russian regional magazines, the nature of actions that form news content from other types of texts (police reports, citizen complaints, etc.) is analyzed, as well as typical failures that occur when using AI algorithms, among which the most detailed description is given to failures in typography, the formation of the meaning of individual statements in the text and grammar. The facts presented indicate the inevitability of including AI in the media toolkit and the need for careful attention to the results of its use.

**Keywords:** artificial intelligence, AI, media, text generation, data processing, information reliability, failures, typography, grammar

OECD: 5.08.EU+1.02.EP

**V**

**Постановка проблемы.** В условиях стремительного развития технологий искусственный интеллект (далее – ИИ) становится неотъемлемой частью многих отраслей, включая новостную журналистику, что открывает перед ней новые возможности для автоматизации рутинных задач, ускорения процесса создания контента и повышения его качества. Однако, несмотря на очевидные преимущества, использование ИИ в журналистике сопряжено с рядом серьезных проблем и рисков, которые могут серьезно снизить доверие к новостному контенту и подрывать репутацию изданий.

В настоящей работе ставится проблема рассмотреть эти риски в лингвистической плоскости, то есть предложить ответ на вопрос, как участие ИИ в генерировании новости отразится на ее тексте? При этом важно подчеркнуть, что проблема ставится и обсуждается не только на основе изучения научной литературы, в том числе зарубежной, но и практики работы региональных медиа, хорошо знакомой автору.

**История вопроса.** Для обсуждения поставленной проблемы важно учитывать опыт использования ИИ в медиасфере, обсуждаемый в научной литературе.

Роботизированная, или автоматизированная журналистика опирается на инновационные подходы к созданию текстов с использованием алгоритмического кода, который способен анализировать отчеты государственных и негосударственных ведомств и служб, а также архивные данные. Алгоритмы могут заполнять заранее подготовленные шаблоны, подставляя в них актуальные данные, полученные в процессе анализа информационного массива. При этом они способны оперативно обрабатывать большие объемы информации и создавать тексты. Такие алгоритмы активно используются ведущими медиаизданиями, включая Forbes, The Guardian, Associated Press, Bloomberg [Veglis, Maniou, 2019, p. 106].

Одним из ключевых этапов развития этой области стало появление в конце 2015 года обновленной версии платформы Wordsmith от компании Automated Insights. Этот инструмент позволил профессионалам из различных сфер загружать данные и автоматически генерировать на их основе тексты [Diakopoulos et al., 2024, p. 107].

Associated Press стала одним из первых крупных медиа, внедривших эту технологию. Платформа Wordsmith способна не только создавать сухие отчеты или обрабатывать официальные сводки, но и писать полноценные новостные материалы. По заявлениям разработчиков, алгоритм может генерировать до 2 000 новостей в секунду, что делает его одним из самых производительных решений в мире [Veglis, Maniou, 2019, p. 108].

Сегодня алгоритм Wordsmith является лишь одним из множества инструментов, доступных на рынке автоматизированной журналистики. Крупные медиакомпании всё чаще разрабатывают собственные решения, адаптированные под их уникальные задачи и стиль. Например, в издании Bloomberg применяется собственная разработка — система BloombergGPT. Она автоматизирует сбор и анализ данных, что ускоряет создание новостных материалов и повышает их качество.

BloombergGPT генерирует новости о финансовых рынках, анализируя данные в реальном времени. Затем тексты проходят проверку и редактирование журналистами. Это позволяет оперативно реагировать на изменения и обеспечивать читателей актуальной и точной информацией. Кроме того, BloombergGPT адаптирован для решения финансовых задач: анализа рыночных настроений, генерации заголовков и суммирования сложной информации (например, финансовых отчетов и графиков). При публикации квартальных отчетов система автоматически формирует новости с ключевыми показателями и их сопоставлением с прогнозами аналитиков. Это ускоряет создание материалов и повышает их точность и информативность [Wu, 2023, p. 4].

Журналисты в Bloomberg применяют различные ИИ-модели для обновления заголовков, контролируемой генерации текста и суммирования графиков. Технологии

позволяют редакции создавать версии новостей для разных платформ и аудиторий, а также представлять сложные данные в доступном формате.

Associated Press (AP) активно использует ИИ для оптимизации новостных операций и повышения эффективности работы журналистов. В частности, AP применяет технологии обработки естественного языка (NLP) для обнаружения новостей в социальных сетях. Это позволяет быстрее выявлять срочные события, чем при ручном мониторинге. AP использует программное обеспечение для распознавания изображений, чтобы улучшить поиск в своём архиве фотографий. Также агентство автоматизирует создание новостных текстов из структурированных данных с помощью технологии генерации естественного языка (NLG). Эта технология уже применялась для публикации отчётов о корпоративных доходах и спортивных событиях, что значительно увеличило производительность работы [Diakoroulos et al., 2024]. Кроме того, AP применяет технологии компьютерного зрения для автоматического создания списков кадров в новостных видеоматериалах и использует машинное обучение для транскрипции видео в реальном времени. Это упрощает процесс подготовки трансляций.

Согласно опросу сервиса журналистских запросов Pressfeed<sup>1</sup>, 65% респондентов применяют возможности ИИ для решения рабочих задач. Большинство из них (72%) обращаются к алгоритмам как минимум раз в неделю, а 23% — ежедневно. В то же время наиболее распространённые задачи для нейросетей: написание текстов (56%), создание изображений и генерация идей для статей (48%), перефразирование и переписывание текстов (42%). Среди популярных нейросетей лидируют ChatGPT (OpenAI) — 37%, Нейро (Яндекс) — 27%, YandexGPT — 25%, YandexART (Шедеврум) — 21%.

Использование систем искусственного интеллекта демонстрирует большой потенциал автоматизации в журналистике. Эти инструменты позволяют ускорить обработку данных, повысить точность и оперативность новостных материалов, а также адаптировать контент для различных платформ и аудиторий.

Роботизированная журналистика применяется для автоматической генерации новостных сообщений на основе данных и фактов. В исследовании «The State of Journalism»<sup>2</sup> от агентства по связям с общественностью Muck Rack сообщается, что около 28% современных журналистов пользуются в своей повседневной работе генеративным искусственным интеллектом (GenAI). Нейросеть помогает им создавать текст, изображения и вести поиск необходимой информации в базах данных.

Первые научные работы, посвященные роботизированной журналистике, появились в начале 2010-х годов. В них исследователи обсуждали возможность применения статистических данных и автоматизированных инструментов для анализа больших массивов информации, а также прогнозировали появление технологий, способных в создании контента частично заменить журналистов.

Обращено внимание на то, что серьезную проблему представляют собой сбои в работе ИИ, поскольку алгоритмы не всегда способны корректно интерпретировать контекст и тонкости информационного сообщения. Подчеркивается, что нейронные

<sup>1</sup> URL: <https://pressfeed.ru/?ut=&ysclid=m7bzc7d7p140524522>.

<sup>2</sup> URL: <https://muckrack.com/research/state-of-journalism-2024>.

сети могут искажать смысл, а также не способны адекватно оценить модальность и контекст информации [Дроздов, 2024, с. 127]. По мнению других, ошибки в сгенерированных текстах вероятны из-за наличия ограничений на распространение определенного рода информации, возможностей распространения фейков при скоростной автоматизированной обработке и подготовке информации к ее выдаче конечному пользователю [Черноусов, 2023, с. 100]. Указано на проблему создания с помощью ИИ «информационных пузырей», когда алгоритмы решают за пользователя, что важно, а что можно проигнорировать; что ограничивает восприятие человека, заставляя его сосредотачиваться только на части информации и способствуя поляризации мнений [Неренц, 2024, с. 5].

Важен вывод о том, что ИИ не обладает интеллектом в полном смысле этого слова, это делает его уязвимым для распространения искажающих реальность сведений; это может привести к появлению «фейковых новостей» и других проблем, связанных с достоверностью информации [Чертовских, Чертовских, 2019, с. 561]. Ситуация усугубляется тем, что возможности ИИ могут использоваться для создания так называемых «дипфейков» — видео- или аудиотекстов, в которых лица или голоса реальных людей заменяются искусственно сгенерированными; подобный контент от ИИ может служить инструментом дезинформации и манипуляции общественным мнением [Свирщ, Олейникова, 2021, с. 2]. Таким образом, одной из ключевых проблем в рамках автоматизированной журналистики становится проверка фактов, так как отсутствие человеческого контроля может привести к снижению качества и достоверности информации [Иванов, 2027, с. 31]. По мнению исследователей, недостатки в факт-чекинге и объективности становятся серьезным изъяном сгенерированных текстов, а сбои могут возникать из-за того, что ИИ берет информацию из заранее выбранных источников, поскольку разработчик может ограничить доступ алгоритма к определенным ресурсам [Исаев, Кокорева, 2019, с. 93]. В итоге модели ИИ часто рассматриваются как «черные ящики», что затрудняет понимание того, как они принимают решения [Оганесян, 2024, с. 277].

Таким образом, в публикациях, посвященных использованию ИИ в медиа, освещен широкий круг вопросов, демонстрирующий как большие возможности, так и проблемы роботизации журналистики, в частности, новостной. Важнейшим моментом оказывается то, что ИИ способен анализировать большие объемы информации в режиме реального времени, что серьезно ускоряет процесс подготовки новостных материалов. Это особенно важно в условиях, когда оперативность публикации играет ключевую роль, например, в ситуациях чрезвычайных происшествий, таких как пожары, ДТП или природные катастрофы, и предоставлять готовый текст для публикации. Существенно, что алгоритмы исключают конфиденциальную информацию из текста, например, персональные данные сотрудников правоохранительных органов и участников происшествий.

Одним из важных преимуществ ИИ является его способность анализировать и связывать данные из нескольких судебных дел одновременно. Например, система может выявить общие черты в преступлениях, определить повторяющиеся схемы или установить связи между участниками разных процессов. Кроме того, ИИ автоматически исключает конфиденциальные данные участников судебных

процессов, такие как персональные данные, адреса или номера документов. Это позволяет соблюдать требования законодательства о защите персональных данных и этические нормы журналистики. В то же время их использование требует обязательного контроля и редактирования со стороны журналистов, что подчеркивает важность сочетания технологий и профессионального подхода для обеспечения качественного контента. Часто сводки судебных заседаний содержат множество деталей, включая хронологию событий, юридические формулировки, данные о сторонах процесса и принятых решениях. Однако избыточная информация может перегружать текст, делая его менее доступным для широкой аудитории. Алгоритмы ИИ способны извлекать из таких данных только ключевые факты, такие как суть преступления, статус обвиняемых, принятые судом меры и сроки рассмотрения дела. Это позволяет создавать лаконичные и информативные материалы, которые удерживают внимание читателя.

Итак, литература вопроса позволяет составить представления о преимуществах ИИ для новостной журналистики и наряду с этим демонстрирует проблемы, которые могут возникнуть со сгенерированными медиатекстами, что заставляет внимательно изучать складывающуюся практику в региональных медиа, чему и посвящена настоящая работа.

**Методология и методика исследования.** Методология настоящего исследования основана на семиотическом понимании медиатекста, который располагает рядом фактур и кодов. Объектом изучения избран графический текст в интернет-медиа, включающий вербальный и визуальный компонент, внимание исследователя сосредоточено преимущественно на вербальном компоненте.

Для анализа отобраны медиатексты, сгенерированные с помощью ИИ, составляющие сейчас заметную часть медийных продуктов.

Язык медиа при этом понимается широко – включая не только лексику, грамматику, смысловую организацию текста, но и типографику [Осокина, 2023].

С опорой на изучение научной литературы можно утверждать, что внедрение ИИ в журналистику сопровождается рядом проблем, которые становятся заметны при изучении практики его применения. В данной работе представлены результаты изучения текстов изданий, активно использующих ИИ в своей работе, таких как «Гражданин и закон Петербурга» и «ЧП столица», а также материалов информационного интернет-агентства «УРА.РУ».

Методика настоящего исследования построена на учете экстралингвистических факторов и собственно языковых фактов. К первым относится учет истории создания новости, круг прецедентных текстов, на основе которых он генерируется. Ко вторым – наблюдения над разными сторонами медиатекстов, в первую очередь типографики, смысла и грамматики. Метод наблюдения позволяет выявить сбои, которые допускает ИИ, метод типологизации – типологизировать их и предложить их объяснение.

**Анализ материала.** Анализ медиатекстов, к созданию которых привлечен ИИ, стоит начать с рассмотрения процесса работы над ними, обращая внимание на характер текстов, служащих источников информации для новости.

Так, сгенерированный материал **«Модифицированные грибы помогут человечеству победить малярийных комаров»**<sup>3</sup> был опубликован в информационном медиаиздании «Pro Город Будущего» 26 января 2025 года. Для его создания ИИ проанализировал первоисточник — научную статью «Transmission of transgenic mosquito-killing fungi during copulation», размещенную в «Scientific Reports» 16 января 2025 года.

В ходе работы ИИ изучил структуру и содержание научной статьи, адаптировав сложные данные для широкой аудитории, сохранил ключевые идеи текста и достоверность информации, перевел текст на русский язык. Целевая аудитория была определена как люди с базовыми знаниями в биологии и экологии, интересующиеся научными открытиями. Чтобы сделать текст доступным, ИИ заменил сложные термины на более простые (например, *генетически модифицированные грибы* вместо *трансгенные грибы*) и использовал примеры для объяснения сложных концепций. Затем он создал структуру новостного текста, включающую разделы: введение, методы, результаты, значимость исследования и заключение. Дальше алгоритмы использовали простой язык и привели текст к логичной структуре. В итоге материал стал понятным для широкой аудитории. После проверки на точность и логическую последовательность текст был отредактирован и адаптирован для публикации.

ИИ был применен для автоматизированного сбора и анализа данных о котировках нефти Brent с Московской биржи в режиме реального времени. На основе полученных данных сгенерирован информационный материал **«За день торгов нефть Brent выросла в цене на 1,05%»**<sup>4</sup>, который был опубликован 9 января 2025 года в издании «Агентство потребительской информации». Алгоритм ИИ анализирует динамику цен, выделяя ключевые изменения, такие как рост или снижение котировок, а также рассчитывает процентные изменения. При этом система ориентирована на целевую аудиторию, заинтересованную в экономических новостях, что позволяет адаптировать контент под их потребности. Однако в текущей версии алгоритма отсутствует функционал для установления причинно-следственных связей между изменениями цен на нефть и внешними событиями (например, геополитическими или экономическими факторами). В результате ИИ ограничивается предоставлением данных, доступных в официальных источниках, без их интерпретации.

Текст **«Петербуржцы жалуются на рекламу наркотиков рядом со школами»**<sup>5</sup>, который опубликован в издании «Аргументы Недели — Петербург» 26 января 2025 года, сгенерирован ИИ в результате анализа данных городского портала **«Наш Санкт-Петербург»**, конкретно — жалоб горожан на несанкционированную рекламу. В этом случае алгоритмы адаптировали формат жалоб жителей Санкт-Петербурга, преобразовав их в текст с использованием косвенной речи, что позволило сохранить смысл обращений, избегая дословного цитирования. Кроме того, система автоматически исключила все конфиденциальные данные, обеспечив анонимность заявителей и соблюдение норм защиты персональных данных. Важным аспектом

<sup>3</sup> URL: <https://progorodspb.ru/2025/01/26/transgennye-griby-pomogut-chelovechestvu-pobedit-malyarijnyx-komarov/>.

<sup>4</sup> URL: <https://apimarket.ru/2025/01/09/za-den-torgov-neft-brent-vyrosla-v-czene-na-105/>.

<sup>5</sup> URL: <https://argumenti.ru/society/2025/01/935811>.

работы ИИ стало структурирование данных: система выделила ключевые темы жалоб, такие как расположение рекламы вблизи образовательных учреждений, ее содержание и потенциальная опасность для детей и подростков.

Успешно применен ИИ для анализа и обработки текстовых данных из официальных сводок ГУ МВД России по Санкт-Петербургу и Ленинградской области, МЧС, а также Следственного комитета РФ, Росгвардии и других ведомств. На основе сводок МЧС сгенерировал фрагменты текста, которые были использованы в материале **«При пожаре на проспекте Энгельса погиб мужчина и пострадала женщина»**, опубликованном в издании **«Аварийный Петербург»** 26 января 2025 года<sup>6</sup>.

Высокую эффективность ИИ демонстрирует при обработке и анализе данных из судебных сводок, которые зачастую отличаются большим объемом информации. Один из примеров — это материал **«Один из похитителей 16-летней девушки останется под стражей до 23 марта»**<sup>7</sup> опубликованный на сайте издания **«Аргументы Недели — Петербург»** 26 января 2025 года.

Итак, проведенный анализ показывает, что репертуар текстов, информация из которых при использовании ИИ становится новостным контентом, довольно широк, это научные тексты, экономическая статистика, жалобы горожан, полицейские и судебные сводки; значительные массивы таких текстов не оказываются препятствием для ИИ, что обеспечивает оперативность новостной ленты.

Анализ текстов, сгенерированных с помощью ИИ, позволил увидеть три типа сбоя относящихся к типографике, смыслу и грамматике.

### Типографика

К типографике относят принципы оформления текста – верстку, шрифты, дизайн, соотношение графически и визуальных составляющих текста [Осокина, 2023].

К типографическим сбоям относят нарушения в расстановке интервалов между абзацами или предложениями. Зачастую ИИ не соблюдает необходимые отступы между абзацами, что может затруднять восприятие текста. Например, в материале **«В Москве в ДТП пострадали четыре человека»**<sup>8</sup> вместо четкого разделения на абзацы с пустой строкой между ними, алгоритм может оставить текст сплошным блоком, что делает материал менее структурированным и сложным для чтения (текст 1).

В Москве произошло дорожно-транспортное происшествие.

На юге столицы, на улице Бакинской, случилось лобовое столкновение двух транспортных средств. Об этом сообщает пресс-служба Департамента транспорта и развития дорожно-транспортной инфраструктуры Москвы. По данным ведомства, один из автомобилей вылетел на встречку.

В ДТП пострадали четыре человека. По предварительным сведениям, на месте работали несколько машин с медицинскими работниками. Что именно стало причиной аварии, неизвестно.

### Текст 1 — Интервалы между абзацами

<sup>6</sup> URL: <https://avariaspb.ru/2025/01/26/pri-pozhare-na-prospekte-engelsa-pogib-muzhchina-i-postradala-zhenshhina/>.

<sup>7</sup> URL: <https://argumenti.ru/society/2025/01/935879>.

<sup>8</sup> URL: <https://chp-msk.ru/2025/01/18/403908-html/>.

Кроме того, к таким сбоям относится некорректная расстановка знаков препинания: пробелы после точек, запятых или других знаков препинания, что нарушает визуальную четкость и затрудняет восприятие текста; различные типы кавычек в одном и том же тексте или выбор шрифта, как в статье «**Ваенга, Любэ, Якубович, Сабуров: какие артисты приедут на Ямал весной. Афиша**»<sup>9</sup> (текст 2).

Спектакль «Скупой», где главную роль исполнит народный артист России, шоумен и ведущий "Поле чудес" Леонид Якубович. Ямальская Филармония, 12 марта;

#### Текст 2 — Разные типы кавычек

Подобные проблемы возникают из-за недостаточной проработки алгоритмов, отвечающих за форматирование текста, а также из-за того, что языковые модели не всегда учитывают визуальные аспекты оформления, фокусируясь исключительно на содержании.

В типографические сбои входит пропуск изображений при наличии подписей. В сгенерированных текстах возможны ситуации, когда алгоритмы не способны корректно извлечь изображение из базы данных издания. В результате текст публикуется без соответствующей фотографии, но с подписью, предполагающей наличие визуального контента. Это создает когнитивный диссонанс у читателя, так как ожидаемое изображение отсутствует, что снижает качество восприятия материала и может вызывать недоумение (текст 3).

## В Петербурге полиция ищет участников видео с избиением подростков, возбуждено дело

Полиция Петербурга ищет участников массовой драки, видео которой разлетелось по телеграм-каналам 2 января. В кадре — группа молодых людей, избивающих подростков по одному.

*Фото: стоп-кадр видео из телеграм-канала «Многонационал»*

*Фото: стоп-кадр видео из телеграм-канала «Многонационал»*

*Судя по записи, в драке участвовали, в том числе, мигранты. На записи видно, что подростков по очереди избивают, одного начинают пинать ногами, когда он уже лежит на земле. Сами подростки, кажется, практически не пытаются дать отпор.*

#### Текст 3 — Отсутствие изображения в тексте

К типографическим сбоям относится некорректное формирование заголовков. Ярким примером может служить заголовок, в котором ИИ оставляет часть внутреннего запроса или технического задания. Так, в заголовке текста «**В ХМАО вновь придут**

<sup>9</sup> URL: <https://ura.news/news/1052880520>.

**тридцатиградусные морозы. Скрин**<sup>10</sup> присутствует лишнее слово *скрин*, вероятно, являющееся частью инструкции для генерации текста на основе скриншота (текст 4).

## **В ХМАО вновь придут тридцатиградусные морозы. Скрин**

Текст 4 — Некорректный заголовок

В пользу версии, что это алгоритмы, можно отметить, что в тексте «**В районе ЯНАО за год вырос показатель ВИЧ-положительных людей. Скрин**»<sup>11</sup> присутствует аналогичный сбой. То есть, когда нейросети нужно «достать» информацию из скриншота, она включает слово *скрин* в название (текст 5).

## **В районе ЯНАО за год вырос показатель ВИЧ-положительных людей. Скрин**

Текст 5 — Сбой в заголовке

Типографический сбой представляет собой несогласованность в написании названий или терминов. Например, в тексте «**За год пунктов выдачи заказов в Петербурге стало на 30% больше**»<sup>12</sup> встречается английское написание названия компании «Wildberries», тогда как в основном тексте — русскоязычный вариант «Вайлдберриз» (текст 6).

В декабре 2024 года доли указанных маркетплейсов в общем числе ПВЗ выглядят следующим образом: Озон — 37,6%, Wildberries — 29,9%, «Яндекс.Маркет» — 14%. Об этом сообщил аналитик рынка недвижимости Алексей Щербатых.

Наиболее заметный рост числа ПВЗ с ноября 2023 года по декабрь 2024 года продемонстрировал бренд Вайлдберриз, увеличив свою сеть на 1519 точек.

Текст 6 — Несогласованное написание названия компании

Итак, типографические сбои вносят в текст излишние элементы, нарушают его графическую структуру, вносят разночтение в обозначение одних реалий, что в целом негативно влияет на восприятие текста.

### **Смысл**

В сгенерированных текстах ИИ иногда пропускает ключевые слова или фразы, что приводит к искажению смысла и нарушению логической целостности текста, а также грамматической структуры предложений. Подобный пример находим в тексте

<sup>10</sup> URL: <https://ura.news/news/1052880802?ysclid=m7d179zwwl779070472>.

<sup>11</sup> URL: <https://ura.news/news/1052881715?ysclid=m7d0zmwllly715953569>.

<sup>12</sup> URL: <https://apimarket.ru/2024/12/29/za-god-punktov-vydachi-zakazov-v-peterburge-stalo-na-30-bolshe/>.

с названием «**В Шушарах водитель сбил 17-летнего подростка на обочине**»<sup>13</sup>. В предложении *пока тот по правой обочине в попутном направлении* отсутствует глагол *шел*, что делает высказывание неполным и затрудняет его понимание.

Аналогичная проблема наблюдается в случае с упоминанием должностей и имен. В одном из предложений материала «**В комитете ГД предложили сократить новогодние выходные до пяти дней**»<sup>14</sup> припущено имя героя статьи *представитель комитета Госдумы по охране здоровья и хирург* нет имени «*Бадма Башанкаев*», что снижает точность и информативность текста (текст 7).

Представитель комитета Госдумы по охране здоровья и хирург выступил с инициативой сократить количество нерабочих дней в новогодние праздники до пяти.

*Он указал, что продолжительные выходные могут негативно сказаться на здоровье жителей России. Он считает, что вместо длительных каникул эффективнее вводить короткие, трёх-пятидневные перерывы несколько раз в год. Его слова были опубликованы в телеграм-канале журналиста Александра Юнашева.*

#### Текст 7 — Пропуск имени героя материала

Такие сбои могут быть вызваны несколькими факторами. Во-первых, языковые модели иногда недостаточно точно анализируют контекст, что приводит к пропуску важных элементов предложения. Во-вторых, неточности могут возникать из-за неполных или некорректных входных данных, на основе которых генерируется текст. В-третьих, алгоритмы могут «перегружать» себя, пытаясь упростить предложение, что приводит к потере ключевых слов.

К таким сбоям относятся неточности в расчётах и интерпретации числовой информации. Например, модель неправильно конвертирует время между часовыми поясами. В материале «**Жители России смогут увидеть соединение Венеры и Сатурна в ночь на субботу**»<sup>15</sup> указано время «05:16 по Гринвичу (12:16 по Москве)», хотя на самом деле разница во времени между Гринвичем и Москвой составляет три часа, и правильное время должно быть 08:16 по Москве.

Аналогичные сбои наблюдаются и в области интерпретации статистических данных: ИИ отметил, что «стоимость нефти выросла почти на 2% в материале «**За девять часов торгов стоимость нефти Brent выросла на 1,5%**»<sup>16</sup>. Однако по данным Московской биржи рост топлива в действительности составил 1,56% (текст 8).

<sup>13</sup> URL: <https://chp-msk.ru/2024/11/19/v-shusharah-voditel-sbil-17-letnego-podrostka-na-obochine/>.

<sup>14</sup> URL: <https://citizenzakon.ru/1070426-v-komitete-gd-predlozhili-sokratit-novogodnie-vyhodnye-do-pyati-dnej/>.

<sup>15</sup> URL: <https://progorodspb.ru/2025/01/17/zhiteli-rossii-smogut-uidet-soedinenie-venery-i-saturna-v-noch-na-subbotu/>.

<sup>16</sup> URL: <https://apimarket.ru/2024/11/18/za-devyat-chasov-stoimost-nefti-brent-vyroslo-pochti-na-2/>.

## За девять часов торгов стоимость нефти Brent выросла почти на 2%

По состоянию на 18 ноября в 17:42 стоимость нефти марки Brent составила 72,40 доллара за баррель. За день цена увеличилась на 1,11 доллара, что соответствует росту на 1,56%.

Текст 8 — Сбой в интерпретации статистических данных

### Грамматика

ИИ допускает сбои в употреблении падежных форм при согласовании в числе, что приводит к нарушению грамматической структуры предложений. Например, в материале **«На территории «Горской» дебаркадер врезался в причал»**<sup>17</sup> присутствует фраза *точный причины и обстоятельства*. Подобные аномалии возникают из-за недостаточной проработки алгоритмов, отвечающих за согласование, а также из-за сложностей в обработке контекста, где требуется учитывать падежные и числовые формы слов.

Как и носители русского языка, а в еще большей степени те, кто его изучает, алгоритмы допускают сбои в предложениях с дробными или составными числительными. Например, в материале под названием **«Вексельберг и Троценко в отличие от Каменщика пощадили авиакомпания и не стали повышать тарифы»**<sup>18</sup> вместо корректного варианта «169,6 тысячи» ИИ ошибочно оформляется как *169,6 тысяч*. Это связано с тем, что алгоритмы не всегда корректно обрабатывают числовые формы.

**Выводы.** Итак, в данной работе был проведен комплексный анализ возможностей и рисков роботизации новостной журналистики. Исследование показало, что ИИ играет ключевую роль в трансформации медиаиндустрии, открывая новые горизонты для автоматизации рутинных задач, ускорения процесса создания контента и повышения его качества. В частности, были рассмотрены успешные кейсы внедрения ИИ в крупных медиакомпаниях, таких как Bloomberg и Associated Press.

Однако, несмотря на очевидные преимущества, применение ИИ в журналистике сталкивается с серьезными проблемами. Алгоритмы нередко допускают неточности в текстах, что может привести к распространению недостоверной информации и созданию фейковых новостей. Подобные аномалии могут серьезно снизить качество и доверие к новостному контенту, подрывая репутацию изданий. Также были подробно рассмотрены характерных для сгенерированных текстов, включая нарушения в согласовании числительных, пропуск значимых ключевых слов и фраз, а также некорректное составление заголовков.

Рассмотренные сбои подчеркивают важность дополнительной проверки и редактирования автоматически генерируемого контента перед его публикацией.

<sup>17</sup> URL: <https://chp-msk.ru/2024/11/04/na-territorii-gorskoj-debarkader-navalilsya-na-prichal/>.

<sup>18</sup> URL: <https://abn.agency/2025/01/10/vekselberg-i-troczenko-v-otlichii-ot-kameshhika-poshhadili-aviakompanii-i-ne-stali-povyshat-tarify/>.

Многочисленные примеры искажений, включая неточности в согласовании числительных, пропуск ключевых слов и фраз, а также некорректное формирование заголовков, могут существенно снизить качество новостного контента и подорвать доверие аудитории.

Для минимизации подобных рисков необходимо внедрение строгих стандартов контроля качества, разработка этических норм использования ИИ в журналистике, а также постоянное обучение и тонкая настройка алгоритмов. Только комплексный подход, сочетающий технологические решения и человеческий контроль, позволит обеспечить высокий уровень точности, достоверности и качества контента, созданного с использованием искусственного интеллекта.

## Литература

Дроздов, Д. Н. (2024). Роботизированная журналистика в контексте процессов развития медиасферы. *Журналистика в цифровую эпоху: технологии и методология творчества*. Минск: БГУ, 127–130.

Захарова, Е. А. (2022). Перспективы применения технологий искусственного интеллекта в сфере образования. *Развитие человека в эпоху цифровизации: ценности, смыслы, действия*. Казань: Институт педагогики, психологии и социальных проблем, 100–102.

Иванов, А. Д. (2016). Современное состояние роботизированной журналистики. *Журналистика цифровой эпохи: как меняется профессия*. Екатеринбург: Уральский федеральный университет, 106–109.

Иванов, А. Д. (2017). Транспарентность роботизированной журналистики: как новые технологии угрожают принципам профессии. *Вестник Челябинского государственного университета*, 8(404), 28–33.

Исаев, Е. М., Кокорева, М. В. (2019). Роботизация новостной журналистики в России: новые принципы работы редакции и трансформация журналистской этики. *Коммуникации. Медиа. Дизайн*, 4(3), 85–102. URL: <https://cmd-journal.hse.ru/article/view/10172/11016>.

Неренц, Д. В. (2024). Особенности фейкового контента в медиапространстве в эпоху развития искусственного интеллекта. *Litera*, 7, 107–114. <https://doi.org/10.25136/2409-8698.2024.7.43843>

Оганесян, С. А. (2024). Применение искусственного интеллекта в компьютерной лингвистике и обработке естественного языка. *Вестник науки: научный электронный журнал*, 7(76)–3, 272–279. URL: <https://www.xn----8sbempclcw3bmt.xn--p1ai/article/16866>.

Осокина, А. И. (2023). Типографика в медиасреде: от печатного текста к электронному. *Молодежь в науке 2023*. Петрозаводск: Международный центр научного партнерства «Новая Наука», 276–285.

Свириц, А. О., Олейникова, П. А. (2021). Развитие технологии deepface. *Modern Science*, 12-4, 309–320.

Суходолов, А. П., Бычкова, А. М., Ованесян, С. С. (2019). Журналистика с искусственным интеллектом. *Вопросы теории и практики журналистики*, 8(4), 647–667. [https://doi.org/10.17150/2308-6203.2019.8\(4\).647-667](https://doi.org/10.17150/2308-6203.2019.8(4).647-667)

Черноусов, Д. А. (2023). Риски использования систем искусственного интеллекта в роботизированной/цифровой журналистике. *Труды по интеллектуальной собственности*, 44(1), 96–103. <https://doi.org/10.17323/tis.2023.16887>

Чертовских, О. О., Чертовских, М. Г. (2019). Искусственный интеллект на службе современной журналистики: история, факты и перспективы развития. *Вопросы теории и практики журналистики*, 8(3), 555–568. [https://doi.org/10.17150/2308-6203.2019.8\(3\).555-568](https://doi.org/10.17150/2308-6203.2019.8(3).555-568)

Diakopoulos, N., Cools, H., Helberger, N., Li, Ch., Kung, E., Rinehart, A., Gibbs, L. (2024). *Generative AI in journalism: the*

## References

Chernousov, D. A. (2023). Risks of usage of artificial intelligence systems in robotic/digital journalism. *Works on Intellectual Property Journal*, 44(1), 96–103. <https://doi.org/10.17323/tis.2023.16887> (In Russian).

Chertovskikh, O. O., Chertovskikh, M. G. (2019). Artificial intelligence in modern journalism: history, facts, prospects for development. *Theoretical and Practical Issues of Journalism*, 8(3), 555–568. [https://doi.org/10.17150/2308-6203.2019.8\(3\).555-568](https://doi.org/10.17150/2308-6203.2019.8(3).555-568) (In Russian).

Diakopoulos, N., Cools, H., Helberger, N., Li, Ch., Kung, E., Rinehart, A., Gibbs, L. (2024). *Generative AI in journalism: the evolution of newswork and ethics in a generative information ecosystem*. <https://doi.org/10.13140/RG.2.2.31540.05765>

Drozdo, D. N. (2024). Robotic journalism in the context of media sphere development processes. *Journalism in the Digital Age: Technologies and Methodology of Creativity*. Minsk: BSU, 127–130. (In Russian).

Isaev, E. M., Kokoreva, M. V. (2019). Robotization of news journalism in Russia: new principles of editorial work and transformation of journalistic ethics. *Communications. Media. Design*, 4(3), 85–102. URL: <https://cmd-journal.hse.ru/article/view/10172/11016>. (In Russian).

Ivanov, A. D. (2016). The state of robotic journalism. *Digital age journalism: how the profession is changing*. Ekaterinburg: Ural Federal University Publ., 106–109. (In Russian).

Ivanov, A. D. (2017). Transparency of robotized journalism: how the new technologies threaten the principles of the profession. *Bulletin of Chelyabinsk State University*, 8(404), 28–33. (In Russian).

Nerents, D. V. (2024). Features of fake content in the media space in the era of artificial intelligence development. *Litera*, 7, 107–114. <https://doi.org/10.25136/2409-8698.2024.7.43843> (In Russian).

Oganessian, S. A. (2024). Use of artificial intelligence in computational linguistics and natural language processing. *Bulletin of Science: Scientific Electronic Journal*, 7(76)–3, 272–279. URL: <https://www.xn----8sbempclcw3bmt.xn--p1ai/article/16866?ysclid=m791ys6cxc434806407> (In Russian).

Osokina, A. I. (2023). Typography in the media environment: from printed text to electronic. *Youth in Science 2023*. Petrozavodsk: International Center for Scientific Partnership “New Science”, 276–285. (In Russian).

Svirsch, A. O., Oleynikova, P. A. (2021). Development of deepface technology. *Modern Science*, 12-4, 309–320. (In Russian).

Sukhodolov, A. P., Bychkova, A. M., Ovanessian, S. S. (2019). Journalism featuring artificial intelligence. *Theoretical and Practical Issues of Journalism*, 8(4), 647–667. [https://doi.org/10.17150/2308-6203.2019.8\(4\).647-667](https://doi.org/10.17150/2308-6203.2019.8(4).647-667) (In Russian).

Veglis, A., Maniou, T. (2019). Chatbots on the rise: a new narrative in journalism. *Studies in Media and Communication*, 7(1). <https://doi.org/10.11114/smc.v7i1.3986>

*evolution of newswork and ethics in a generative information ecosystem*. <https://doi.org/10.13140/RG.2.2.31540.05765>

Veglis, A., Maniou, T. (2019). Chatbots on the rise: a new narrative in journalism. *Studies in Media and Communication*, 7(1). <https://doi.org/10.11114/smc.v7i1.3986>

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., Mann, G. (2023). *BloombergGPT: a large language model for finance*. arXiv preprint: arXiv:2303.17564. <https://doi.org/10.48550/arXiv.2303.17564>

Wu, S., Irsoy, O., Lu, S., Dabravolski, V., Dredze, M., Gehrmann, S., Mann, G. (2023). *BloombergGPT: a large language model for finance*. arXiv preprint: arXiv:2303.17564. <https://doi.org/10.48550/arXiv.2303.17564>

Zakharova, E. A. (2022). Prospects for the application of artificial intelligence technologies in the field of education. *Human Development in the Era of Digitalization: Values, Meanings, Actions*. Kazan: Institute of Pedagogy, Psychology and Social Problems Publ., 100–102. (In Russian).

### **Для цитирования статьи:**

Александров, Д. Д. (2024). Новостной текст и искусственный интеллект. *VERBA. Северо-Западный лингвистический журнал*, 4(14), 8–21. DOI: 10.34680/VERBA-2024-4(14)-8-21

### **For citation:**

Aleksandrov, D. D. (2024). News text and artificial intelligence. *VERBA. North-West linguistic journal*, 4(14), 8–21. (In Russian). DOI: 10.34680/VERBA-2024-4(14)-8-21