Ирина Ивановна КАРПЕНКО,

Валерий Юрьевич МЕРИНОВ

МИНИМИЗАЦИЯ ДЕСТРУКТИВНОГО ВЛИЯНИЯ AI-ТЕХНОЛОГИЙ НА СОВРЕМЕННУЮ ЖУРНАЛИСТИКУ

В статье рассматривается четыре группы проблем, с которыми сегодня сталкиваются современные медиа при работе с нейросетевыми моделями: качество и достоверность контента (публикация поверхностных или ошибочных медиаматериалов с элементами нейросетевых галлюцинаций, намеренное или ненамеренное распространение фейков и дипфейков), экономические и профессиональные риски (угроза сокращения рабочих мест), этические и правовые вопросы (неосведомленность аудитории о том, что СМИ публикуют сгенерированные ИИ материалы, нарушение авторских прав нейросетевыми моделями при генерации изображений, видео- и аудиоматериалов), потеря доверия аудитории к СМИ.

Ключевые слова: *искусственный интеллект; ИИ; генеративные технологии; традиционный ИИ; нейросеть; журналистика; медиа.*

Irina I. KARPENKO,

Valery Yu. MERINOV

MINIMIZING THE DESTRUCTIVE IMPACT OF AI TECHNOLOGIES ON MODERN JOURNALISM

The article considers four groups of problems that modern media face today when working with neural network models: quality and reliability of content (publication of superficial or erroneous media materials with elements of neural network hallucinations, intentional or unintentional distribution of fakes and deepfakes), economic and professional risks (threat of job cuts), ethical and legal issues (audience unawareness that the media publishes AI-generated materials, copyright infringement by neural network models when generating images, video and audio materials), loss of audience trust in the media.

Keywords: *artificial intelligence; AI; generative technologies; traditional AI; neural network; journalism; media.*

За последние четыре неполных года, с момента выхода в 2022 году в свет ChatGPT, современная журналистика столкнулась с рядом вызовов и беспрецедентных изменений, которые не испытывала за всю свою многовековую историю. Искусственный интеллект (ИИ), и прежде всего генеративный его тип, предоставил

современным медиа целый спектр перспективных направлений развития: генерация текстов и иллюстративного материала (как статичного, так и динамичного), системы анализа данных, создание виртуальных аватаров, верификация контента посредством анализа больших объемов данных и многое другое – все это изменило традиционные подходы к созданию и распространению медиаконтента. Однако, стоит отметить, что эйфория, которую испытал медиамир, только приняв в свою профессиональную деятельность технологии ИИ, сегодня уже сошла на нет. Теоретики и практики медиа активно обсуждают негативные последствия от повсеместного внедрения ИИ в деятельность современных СМИ (Болдина, 2024; Дзялошинский, 2022; Арсентьева, Морозова, 2021; Гурушкин, 2024 и др.). Наряду с повышением эффективности журналистики возникают и серьезные проблемы: снижение доверия к медиа, рост фейковых новостей, угроза замены человеческого труда машинным и многое другое. В связи с этим актуальной задачей становится разработка мер, позволяющих минимизировать деструктивное влияние ИИ на журналистику.

Проблемы, с которыми столкнулась современная журналистика при тотальном внедрении технологий искусственного интеллекта на разных уровнях профессии, можно разделить на четыре группы.

1. **Качество и достоверность контента** являются основными и наиболее важными проблемами внедрения генеративных технологий искусственного интеллекта в работу современных медиа. К сожалению, соблазн упростить работу с текстом, переложить обязанности по его написанию на машину и тем самым ускорить производственные процессы заставляет журналистов все чаще обращаться в своей работе к генеративным моделям. По результатам исследования крупнейшей в России медиакраудсорсинговой компании Pressfeed, «66% журналистов и 48% экспертов СМИ прибегают к помощи нейросетей при подготовке контента для СМИ» (48% экспертов СМИ… 2023). В этом контексте следует подчеркнуть, что бесконтрольное и тотальное использование генеративных моделей может спровоцировать лавинообразное распространение поверхностных или ошибочных медиаматериалов. Нейросети, и особенно бесплатные и плохо обученные, но более доступные модели, могут создавать тексты без глубокого анализа, что снижает ценность медиаконтента. Также исследователи отмечают такие проблемы, как «ограниченность тематики, шаблонность текстов, затратность внедрения и поддержки технологий» (Бейненсон 2020: 297).

Другой важной проблемой в контексте качества ИИ-журналистики является распространение фейков и дипфейков. Генеративные технологии, при всей своей уникальности и перспективности, активно используются для создания разного рода дезинформации. Широко известный фейк о пророческой цитате из отчета Европола, Агентства Европейского союза по сотрудничеству в правоохранительной сфере, о том, что «к 2026 году до 90 процентов онлайн-контента может быть создано с использованием искусственного интеллекта», активно подхваченный мировыми СМИ, демонстрирует кризис критического мышления современных журналистов на фоне чрезмерного увлечения сенсациями и трендами. Вместе с тем «Facing reality? Law enforcement and the challenge of deepfakes» дает ценные рекомендации о том, какими способами можно подобный контент идентифицировать. Для этого могут быть использованы три основных типа мер: ручное обнаружение, автоматизированное обнаружение и превентивные меры.

Ручное обнаружение подразумевает под собой анализ несоответствий и артефактов. «Это трудоёмкая задача, которую можно выполнить лишь для очень ограниченного числа файлов, и требует соответствующей подготовки для понимания всех соответствующих признаков» (Facing reality?... 2025). В ряду таких несоответствий или артефактов в документе упоминаются: размытость по краям лица, отсутствие моргания, отражение света в глазах, несоответствия в волосах, рисунке вен, шрамах и т. д., несоответствие фонем и визем на видео, отсутствие корреляции между движением лица и движением головы, несоответствия в фоне, в теме, а также в фокусе, глубине и т.д.

Автоматизированное обнаружение дезинформации подразумевает под собой использование программного обеспечения, способного сканировать любой цифровой контент и автоматически выдавать заключение о его подлинности. Однако подобный способ только на первый взгляд кажется идеальным. Подобные программные продукты, разрабатываемые сегодня, в том числе компаниями Facebook (Признана экстремистской организацией и запрещена на территории РФ) и McAfee, работают также на базе технологий ИИ. Отсюда следует проблема: «большинство моделей обнаружения дипфейков обучаются на основе баз данных изображений, содержащих дипфейки. Таким образом, выявленные признаки манипуляции основаны на данных известных дипфейков, что затрудняет оценку эффективности обнаружения дипфейков, созданных неизвестными или обновлёнными моделями» (Facing reality?... 2025).

И, наконец, превентивные меры, которые можно сформулировать для журналистов с целью недопущения использования в своей деятельности фейковой и дипфейковой информации деструктивного характера: использование аудиовизуальной аутентификации интервьюируемого, а не только аудио, получение данных путем живого видеоподключения; выполнение биометрической аутентификации интервьюера, к примеру, просьба совершить произвольное движение рукой в зоне лица.

Мы полагаем, что минимизации проблем, связанных с качеством и достоверностью контента, может поспособствовать реализация гибридной модели «человек + алгоритм», в которой роль человека является главенствующей. Модель построена на разделении обязанностей, сохраняя за человеком право на редактирование, дополнение, расширение и углубление информации, привнесение оценки, прогнозирования, добавление экспертного заключения, которое даст возможность сопрячь информацию с ньюсмейкером и верифицировать представляемые данные с реальными людьми. Искусственному интеллекту в этой связке отводится роль автоматизированной системы для сбора данных, выполнение рутинных задач дайджестирования источников для их дальнейшей верификации силами профессионального журналиста.

В ситуации преодоления опасностей ненамеренного распространения фейков и дипфейков через медиаканалы, к сожалению, автоматизация не может стать панацеей, а только использоваться как дополнительный инструмент. В числе таких продуктов на современном программном рынке – разработки компаний Factmata и Full Fact. Все генерируемые искусственным интеллектом тексты должны проходить строгую процедуру «ручного» фактчекинга и редакторскую проверку перед публикацией.

Полагаем, что эффективным инструментом оптимизации понимания как медиакомпаниям работать в условиях тотального распространения ИИ и минимизировать риски распространения недостоверной или заведомо ложной информации, стала бы практика цехового саморегулирования, выразившаяся в разработке этических медиагайдлайнов, редакционных стандартов, которые бы регламентировали работу с ИИ-генерациями, к примеру, обязывали бы маркировать сгенерированные тексты, которые, в отличие от иллюстративного материала, пока никак не маркируются. Маркировочные знаки должны находиться на видном месте и содержать

«как минимум фразу “сгенерировано искусственным интеллектом”, “сгенерировано”, “сгенерировано ИИ”» (Карпенко, 2024: 52–53). Визуальная маркировка текстовых материалов необходима еще и потому, что к подобному типу контента можно применить только один ее тип – визуальный. В отличие от текста, с графическими объектами, видео и аудиоматериалами возможно проводить двухуровневую маркировку: посредством расстановки визуальных водяных знаков и посредством включения в файлы метаинформации, считывание которой возможно автоматизировано, что упрощает программный контроль за подобным контентом.

2. **Экономические и профессиональные риски** – это еще одна группа проблем, которые вызывают опасения у современного медиасообщества. Одним из направлений в этой области является угроза сокращения рабочих мест, связанных с делегированием рутинных обязанностей или простых алгоритмизированных рабочих процессов генеративным и традиционным технологиям искусственного интеллекта. Подобная автоматизация может привести к уменьшению спроса на журналистов, особенно в новостных агентствах, медиа, занимающихся дайджестированием новостей, составляющих обзоры СМИ и обработкой иной вторичной информации. Информация о первых прецедентах написания нейросетевыми моделями обзорных статей для медиа на основе анализа больших данных достаточно разрозненная. Наиболее часто транслируемая гипотеза состоит в том, что первый ИИ-текст был написан в 2020 г. GPT-3 для The Guardian, в котором самая известная на сегодня ИИ-модель пытается убедить читателей британского издания в том, что искусственного интеллекта бояться не стоит: «Стивен Хокинг предупреждал, что ИИ может “означать конец человеческой расы”. Я же здесь для того, чтобы убедить вас не беспокоиться. Искусственный интеллект не уничтожит людей. Поверьте» (Цит. по: Робот полностью… 2020). Согласно другой точке зрения, «первая статья, написанная программой с искусственным интеллектом, (была) опубликована на сайте Los Angeles Times в 2008 году» (Акентьев, Максимов, Лысенко 2018: 691). В материале компьютер, на основе анализа числовых данных сейсмографа представил данные о землетрясении в Калифорнии.

Сегодня делегирование написания обзоров, рекламных сообщений для социальных медиа, сводок, дайджестов и эффективных заголовков и лидов стало тем делом, которое журналисты делегировали искусственному интеллекту. Однако эта практика может повлечь за

собой сокращение рабочих мест для журналистов, пишущих короткие заметки для новостных лент СМИ, сотрудников, занимающихся рерайтом и дайджестированием. Творческих авторов, работающих с эксклюзивным и аналитическим контентом, эта проблема не коснется. Другими мерами, которые необходимо уже сегодня предпринимать журналистам, находящимся в зонах риска потери работы, является расширение своей квалификации на аналитику, расследования, авторские форматы, то есть те сферы, которые пока недоступны ИИ. С другой стороны, искусственный интеллект – это программируемый продукт, который не способен принимать решения самостоятельно; во многом качество производимых им объектов зависит от грамотного управления со стороны промт-инженера. Промтинжиниринг – понимание принципов работы нейромоделей, структуры промтов, весов токенов, особенностей галлюцинирования моделей, правовых и этических аспектов взаимодействия с ИИ-генерациями, особенностей деструктивных коммуникаций и конструктивных практик при создании дипфейков и др. – является для современных журналистов необходимым в обретении навыков работы с нейросетевыми моделями. В ситуации же бесконтрольного и неграмотного использования в медиадеятельности нейросетевых генераций, изобилующих артефактами и галлюцинациями, журналистику может постигнуть деградация профессиональных стандартов, которая выразится в уменьшении разнообразия тем, жанров и форм репрезентации контента, снижении уровня аналитики, нивелировании авторского стиля.

3. **Этические и правовые вопросы.** Одним из важных вопросов в контексте использования продукта нейросетевых генераций в современных медиа является вопрос осведомленности аудитории и отсутствия прозрачности, так как читатели не всегда знают, создан материал человеком или программой. В отношении графических объектов в отдельных современных медиа уже выработались определенные стандарты маркирования изображений. Сегодня на законодательном уровне обсуждаются «возможности маркирования и текстового контента: “развитие ИИ-сервисов несет опасность появления огромных массивов не только никем не проверенных текстов, но и даже полностью выдуманных фактов, цифр и данных”. Этот вопрос, однако, остается пока непроясненным, так как ни визуальной, ни программной возможности маркировать текстовые данные на уровне мета-информации сегодня нет» (Карпенко 2024: 53–54).

Еще одним важным аспектом является проблема нарушения авторских прав при генерации изображений, видео- и аудиоматериалов. Нейросетевые модели, в отличие от алгоритмизированного традиционного искусственного интеллекта, – самообучаемые системы, использующие для своих нужд триллионы единиц данных, большая часть из которых – это охраняемые объекты чьей-то интеллектуальной собственности, у которых ИИ заимствует образы, стиль, сюжет и пр. На фоне этого в мире оформилась практика проведения охранных мероприятий по защите объектов интеллектуальной собственности от их индексации генеративными моделями. Одним из самых эффективных инструментов, защищающих контент от ИИ, является «глазирование» изображений – использование программных продуктов типа «Nightshade и Glaze, задача которых состоит в защите стиля произведения и авторских прав в целом художников и фотографов при использовании их работ для обучения различных моделей искусственного интеллекта» (Карпенко, 2024: 47). Практика названных компаний доказала, что нейросетям не удается заимствовать стиль изображения в 92% случаев при использовании «глазури», в то время как без нее этот процент равен семи (Nightshade и Glaze… 2024).

Другим действенным инструментом наложения ограничений на ИИ-индексации может стать использование блокчейн-технологий для подтверждения авторства, что даст в дальнейшем основание для судопроизводства против незаконного использования авторского контента.

Об использовании ИИ как помощника для журналиста, а не как автора графического контента и автора текстов сегодня говорят не только теоретики, но и практики. Руководитель образовательных программ в Prague School Media Злата Понировская предлагает современным медиа идеальное решение для снижения рисков и сохранения своей репутации: «Проверять информацию — это задача журналиста, а не AI» (Гайсина 2024). Наиболее адекватным решением проблемы этического регулирования практик использования генеративных нейросетей в редакциях СМИ является разработка внутренних этических нормативов для работы с ИИ. На федеральном уровне такая практика существует с 26 октября 2021 г., когда был принят Кодекс этики в сфере искусственного интеллекта. Сегодня в числе подписантов Кодекса 422 организации, в числе которых, в основном, компании, работающие в сфере IT, крупные экосистемы типа Сбера, Яндекса, VK, МТС и пр. и крупнейшие вузы России – МГИМО, НИУ

ВШЭ, МГТУ им. Н.Э. Баумана и др. (Подписанты кодекса этики в сфере ИИ 2025). Средств массовой информации в этом списке нет. Это обусловлено, прежде всего, тем, что на законодательном уровне ИИ ассоциируются именно с технологической стороной вопроса. Сам альянс позиционирует себя как организацию, объединяющую «ведущие технологические (выделение – И.К.) компании для совместного развития их компетенций и ускоренного внедрения искусственного интеллекта в образовании, научных исследованиях и в практической деятельности бизнеса» (Альянс в сфере искусственного 2025), игнорируя огромный пласт медиакомпаний.

Зарубежная практика доказывает, что разработка этических стандартов по работе с ИИ – также не общая практика, а скорее исключение. Этические правила разработали Associated Press (2023), The New York Times (2024), информационное агентство Deutsche Presse-Agentur (2023), The Guardian (2023) и многие другие. В ряду основных тезисов: ИИ не может писать полноценные новостные материалы, ИИ не может обрабатывать конфиденциальные источники, ИИ не может генерировать изображения и видео без явной маркировки, контент, созданный ИИ, всегда проверяется на точность и достоверность человеком. Зарубежные медиа сегодня делегируют ИИ только небольшую часть своих производственных процессов: составление дайджестов, кратких сводок, генерирование заголовков, оптимизированных под SEO, создание рекламного контента для социальных сетей, перефразирование предложений для повышения ясности, создание информационных карточек для повышения вовлеченности читателей, поиск идей для вопросов к интервьюируемым (Аббас, 2025).

В России широкой практики создания внутриредакционных правил работы с ИИ пока нет.

4. **Потеря доверия аудитории** также может стать одной из серьезных проблем в условиях повсеместного использования ИИ-технологий. И в данном случае мы можем вести речь не только о генеративных практиках, из-за которых вероятен рост скептицизма и недоверия к современным медиа по причине большого количества фейков и дипфейков, а также автоматизированного новостного и информационного контента, но и о традиционных алгоритмизированных нейропрактиках, используемых в современных медиа – персонализированные алгоритмы, рекомендательные технологии, которые только на первый взгляд гибко подстраиваются

под интересы аудитории, на деле же создают своего рода «информационный пузырь», сужая кругозор аудитории, ограничивая ее в темах на основании поисковых запросов, просмотренных материалов и заполненных при регистрации анкетных данных. Исследователи отмечают, «что в сфере рутинизированных операций технологии могут лишить человека выбора – уже сейчас ИИ предлагает нам музыку и фильмы, предоставляя конкретной аудитории определенный контент, за рамки которого она может не выйти. Предполагается, что алгоритм будет принимать все больше решений за человека, в чем и состоит опасность…» (Пузанова, 2024: 751–752).

Со стороны медиа наиболее адекватным решением ситуации будет уклон в объяснительную журналистику и сторителлинг, который может разнообразить сухой, алгоритмизированный «быстрый» контент увлекательными историями и эмоциями.

Решение этих проблем, к сожалению, не ограничивается только мерами со стороны медиа, которые должны предупреждать аудиторию об использовании рекомендательных ИИ-технологий на своих ресурсах и объяснять, как с ними работать, создавая тем самым обстановку прозрачности и публичности медиаресурсов, но и требует от самой аудитории желания включаться в медиаобразовательные проекты.

Подводя итог, стоит отметить, что при всем обилии проблем, которые связаны с использованием генеративного и традиционного ИИ в современных медиа, нахождение пути их преодоления возможно – это нахождение правильного баланса между инновациями и рисками. Любое внедрение ИИ в социально значимые коммуникационные процессы необходимо при наличии строгого и ответственного контроля над качеством контента, этическими и правовыми нормами. Решение возникающих проблем требует совместных усилий журналистов, регуляторов, технологических компаний-разработчиков нейросетевых моделей и аудитории современных медиа.

Практика работы с искусственным интеллектом в журналистике требует ответственного подхода. С целью минимизации риска потери качества при использовании генеративных технологий необходимо внедрять качественный и многоступенчатый контроль качества медиапродукта, так как скорость, низкозатратность и высокая производительность нейросетей никак не могут заменить высокую достоверность, верифицируемость и оригинальность контента, созданного человеком. По этой причине необходимо осознавать и принимать риски работы с ИИ и подходить к их использованию в

массовой коммуникации ответственно: сохранять человеческий контроль, регулировать использование ИИ на правовом и этическом уровнях и инвестировать в качество, а не скорость. Это позволит сохранить доверие аудитории и профессиональные журналистские стандарты.

Список литературы:

48% экспертов СМИ и 66% журналистов используют нейросети для подготовки материалов для СМИ // cnews. URL: https://www.cnews.ru/news/line/2023-07-18_48_ekspertov_smi_i_66_zhurnalistov (дата обращения 06.08.2025).

Аббас А. Как ИИ трансформирует журналистику: подход The New York Times к Echo // Unite.Ai. URL: https://www.unite.ai/ru/how-ai-is-transforming-journalism-the-new-york-times-approach-with-echo/ (дата обращения: 31.07.2025).

Акентьев Д. Г., Максимов Д. И., Лысенко И. С. Проблема искусственного интеллекта и искусственного разума в жизни современного человека / Д. Г. Акентьев, // Актуальные проблемы авиации и космонавтики. – 2018. Т. 3, № 4(14). С. 690–692.

Альянс в сфере искусственного интеллекта // Альянс в сфере искусственного интеллекта. URL: https://a-ai.ru/ (дата обращения: 29.07.2025).

Арсентьева А. Д., Морозова А. А. Проблемы внедрения алгоритмов искусственного интеллекта в российскую журналистику // Огарёв-Online. – 2021. № 2(155).

Бейненсон В. А. Использование технологий искусственного интеллекта при создании развлекательного контента СМИ // MEDIAОбразование: медиа как тотальная повседневность: Материалы V Международной научной конференции, Челябинск, 24–25 ноября 2020 года / Под редакцией А.А. Морозовой. Том Часть 1. – Челябинск: Челябинский государственный университет, 2020. С. 296–301.

Болдина К. А. Риски автоматизации новостной журналистики на основе ИИ // Успехи гуманитарных наук. 2024. № 3. С. 7–14.

Гайсина Л. AI и журналистская этика: как и когда медиа использовать нейросети без риска потерять репутацию // New Reporter. URL: https://newreporter.org/2024/06/07/ai-i-zhurnalistskaya-etika-kak-i-kogda-media-ispolzovat-nejroseti-bez-riska-poteryat-reputatsiyu/ (дата обращения 06.07.2025).

Гурушкин П. Ю., Корнеева К. В. Искусственный интеллект в медиакоммуникациях и творческих профессиях: угрозы и возможности // Litera. 2024. № 9. С. 91–101.

Дзялошинский И. М. Когнитивные процессы человека и искусственный интеллект в контексте цифровой цивилизации: монография. – Москва: Ай Пи Ар Медиа, 2022. 583 с.

Карпенко И. И. Генеративные технологии искусственного интеллекта в

современных медиа: проблемы авторского права и маркирования нейросетевого продукта // Современный дискурс-анализ. – 2024. – № 2(35). – С. 45–55.

Подписанты кодекса этики в сфере ИИ // Альянс в сфере искусственного интеллекта. URL: https://ethics.a-ai.ru/assets/ethics_files/2025/08/19/%D0%A1%D0%BF%D0%B8%D1%81%D0%BE%D0%BA_%D0%BA%D0%BE%D0%BC%D0%BF%D0%B0%D0%BD%D0%B8%D0%B9_2.pdf (дата обращения 29.07.2025).

Пузанова Ж. В., Тертышникова А. Г., Павлова У. О. Технологический дискурс в российских СМИ: основные стратегии в репрезентации искусственного интеллекта // Вестник Российского университета дружбы народов. Серия: Социология. 2024. Т. 24, № 3. С. 747–763. EDN FJVAXU.

Робот полностью написал эту статью. Человек, тебе уже страшно? // Коммерсантъ. URL: https://www.kommersant.ru/doc/4485415 (дата обращения: 01.08.2025).

Facing reality? Law enforcement and the challenge of deepfakes // Europol. URL: https://www.europol.europa.eu/publications-events/publications/facing-reality-law-enforcement-and-challenge-of-deepfakes (date of access: 12.08.2025).

Nightshade и Glaze — как работает защита авторского права художников отAI? URL: https://club.dns-shop.ru/blog/t-328-prilojeniya/118167-nightshade-i-glaze-kak-rabotaet-zaschita-avtorskogo-prava-hudojn/ (date of access: 08.08.2022).

КАРПЕНКО Ирина Ивановна – доктор филологических наук, профессор кафедры журналистики, Белгородский государственный национальный исследовательский университет (г. Белгород), *karpenkoirina@bsuedu.ru*

МЕРИНОВ Валерий Юрьевич – кандидат философских наук, доцент кафедры социологии и медиакоммуникаций, Российский государственный университет социальных технологий (г. Москва).