

Юридические и этические проблемы регулирования применения искусственного интеллекта в современных массмедиа

Дарья В. Неренц

*Российский государственный гуманитарный университет,
Москва, Россия, ya.newlevel@yandex.ru*

Аннотация. Применение искусственного интеллекта (ИИ) в массмедиа становится все более привычным явлением, требующим глубокого осмысления. В силу отсутствия четкого законодательного регулирования и этических стандартов аудитория не защищена от рисков обмана, введения в заблуждение, манипуляций. Статья посвящена изучению ныне уже применяемых правовых и этических норм использования нейросетей в массмедиа, а также анализу тех аспектов, которые по-прежнему нуждаются в регулировании. В качестве научных методов использован метод индукции, описания, обобщения, анализа контента. В рамках осмысления необходимых этических стандартов был проведен опрос студентов факультета журналистики РГГУ.

Ключевые слова: искусственный интеллект (ИИ), нейросеть, массмедиа, правовой аспект, этические проблемы, журналистика, информационное пространство, медиаконтент, виртуальные телеведущие, авторское право

Для цитирования: Неренц Д.В. Юридические и этические проблемы регулирования применения искусственного интеллекта в современных массмедиа // Вестник РГГУ. Серия «Литературоведение. Языкознание. Культурология». 2024. № 11. С. 122–137. DOI: 10.28995/2686-7249-2024-11-122-137

Legal and ethical issues of regulating the use of artificial intelligence in modern mass media

Daria V. Nerents

*Russian State University for the Humanities, Moscow, Russia,
ya.newlevel@yandex.ru*

Abstract. The use of artificial intelligence (AI) in mass media is becoming more and more commonplace and requires deep thinking. Due to the lack of clear legal regulation and ethical standards, the audience is not protected from the

risks of deception, misleading and manipulation. The article deals with the study of the currently applied legal and ethical standards of the neural networks use in mass media, as well as to the analysis of those aspects that still need to be regulated. The scientific methods used are induction, description, generalization, and content analysis. As part of comprehending the necessary ethical standards, a survey of journalism students at the RSUH faculty of journalism was conducted.

Keywords: artificial intelligence (AI), neural network, mass media, legal aspect, ethical issues, journalism, information space, media content, virtual TV presenters, copyright law

For citation: Nerents, D.V. (2024), "Legal and ethical issues of regulating the use of artificial intelligence in modern mass media", *RSUH/RGGU Bulletin, "Literary Theory. Linguistics. Cultural Studies"*, no. 11, pp. 122–137, DOI: 10.28995/2686-7249-2024-11-122-137

Введение

Искусственный интеллект в 2023–2024 гг. – одна из самых обсуждаемых тем во всех сферах деятельности. Причина этого – огромный скачок в развитии нейросетей, их возможность быстро адаптироваться под любые заданные условия и выполнять крайне сложные поручения в максимально сжатые сроки, что значительно упрощает некоторые рабочие задачи, поставленные перед специалистами. И это явление уже сейчас трансформирует мышление общества, изменяет методологию и методику работы во многих отраслях, заставляет совсем иначе смотреть на моменты действительности, которые до этого казались привычными и знакомыми.

Сам ИИ – явление не новое и в профессиональной IT-среде используется уже не один десяток лет. Тест Тьюнинга, который считается прообразом современных нейросетей, появился еще в 1950 г. (об этом в журнале *Mind* была написана статья «Вычислительные машины и разум» [Малыгина 2018]). А в середине 1950-х гг. на Дармутской конференции американский информатик Джон Маккарти ввел понятие искусственного интеллекта, который он определил как «науку и технику для создания интеллектуальных машин» [McCarthy 2007]. Спустя чуть более 10 лет был создан первый виртуальный голосовой ассистент IBM, который назвали Shobox. Интересно отметить, что в России широко известным виртуальный помощник стал с появлением Алисы от Яндекса лишь в 2017 г. (хотя обладатели айфонов стали пользоваться помощником Siri на три года раньше), потом эту практику активно стали применять банки (первым стал Сбер), в 2023 г. помощник появился уже на портале Госуслуг и, можно сказать, стал распространенной практикой.

Определение понятия «искусственный интеллект» в рамках федерального проекта по развитию ИИ и Национальной стратегии развития искусственного интеллекта на период до 2030 г. представляет собой следующее описание:

Комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру (в том числе информационные системы, информационно-телекоммуникационные сети, иные технические средства обработки информации), программное обеспечение (в том числе в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений¹.

Очевидно, что данное понятие дает весьма исчерпывающее представление об этой технологии, однако представляет собой излишне обширное описание, что затрудняет восприятие сути широкой аудиторией.

На портале Министерства экономического развития РФ представлено следующее определение:

Искусственный интеллект в Российской Федерации (ИИ) – комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма), получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека².

Описание является дискуссионным, поскольку на данном этапе развития ИИ не способен находить самостоятельные решения и получать результаты, равные человеческим.

¹ Разработка Федерального проекта по устойчивому развитию искусственного интеллекта // Министерство цифрового развития, связи и массовых коммуникаций Российской Федерации. URL: <https://digital.gov.ru/uploaded/files/razrabotka-federalnogo-proekta-po-razvitiyu-iskusstvennogo-intellekta.pdf> (дата обращения 28 августа 2024).

² Развитие искусственного интеллекта // Министерство экономического развития Российской Федерации. URL: https://www.economy.gov.ru/material/departments/d01/razvitie_iskusstvennogo_intellekta/ (дата обращения 30 августа 2024).

Отходя от официальной терминологии, интересно отметить, что эксперт в области ИИ С.В. Карелов указывает на существование более 200 официальных определений понятия «искусственный интеллект»³. Исследователи в каждой области дают свое определение этой технологии в зависимости от целей и задач, которые ИИ решает в этой сфере деятельности. Поскольку данное исследование посвящено применению ИИ в массмедиа, мы будем понимать ИИ как комплекс технологий, способных выполнять творческие и технические задачи в контексте сбора, обработки, систематизации, оформления и тиражирования информации, которые ранее выполнялись исключительно человеком.

Открывшиеся возможности породили новые, весьма серьезные проблемы, которые уже не один год решаются как в этическом, так и в правовом поле. Связаны они, в первую очередь, с тем, что при успешных экспериментах касательно внедрения в ИИ логического мышления (иными словами, создание генеративного или общего ИИ) возникают риски не только в экономическом плане, но и риск потери контроля над подобными системами.

Цель данного исследования – проанализировать законодательные инициативы и этические регуляторы, существующие в информационном поле, и изучить степень их эффективности на современном этапе.

Правовой аспект применения ИИ: оценка возможностей и рисков

Информационные угрозы со стороны искусственного интеллекта являются весьма заметным последствием использования этой технологии. Возможность применения нейросетей всем пользователям с любыми целями открыли небывалые возможности не только перед специалистами, но и перед мошенниками. С позиции СМИ, уже возникли прецеденты с официальными жалобами на работу ИИ. Газета The New York Times подала в суд на OpenAI (создателей ChatGPT) и Microsoft из-за нарушения авторских прав⁴. Согласно поданному в федеральный окружной суд Манхэттена иску, издание обвиняет эти корпорации в использовании без

³ Гурова Т. Хакнуть человечество // Эксперт. 2019. № 3. С. 40–48.

⁴ Grynbaum M., Mac R. The Times sues OpenAI and Microsoft over A.I. use of copyrighted work // The New York Times. 2023. 27 Dec. URL: <https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html> (дата обращения 30 Aug. 2024).

разрешения публикаций газеты для обучения моделей ИИ. В России часть медиаактивов Shkulev Media Holding оформила запрет для поисковых роботов, чтобы они больше не могли использовать контент медиахолдинга для нейросетевых продуктов. Среди них – E1.ru, NGS.ru, «Фонтанка» и др.⁵

В противовес запретам ряд крупнейших СМИ подписали с OpenAI (создателями ChatGPT) соглашение об использовании их контента для обучения нейросетей. В проекте принимают участие The Atlantic, Vox Media, News Corp, Financial Times, Associated Press, Axel Springer, Le Monde, Prisa Media⁶. Договор построен на товарно-денежных отношениях, согласно которым издания получают финансовое вознаграждение за предоставление ИИ своих материалов. Однако здесь возникает большое количество этических вопросов.

Таким образом, на основе проанализированных ранее кейсов применения ИИ в массмедиа можно выделить самые релевантные проблемы, требующие нормативно-правового регулирования:

- 1) рост объема некачественного контента (генерация не создает ничего нового, ИИ способен учиться только на уже существующих данных, а эти данные, в свою очередь, могут быть устаревшими или нерелевантными);
- 2) вредоносные и опасные последствия использования ложных ответов ИИ (фейки);
- 3) негативное влияние на рынок труда (замещение человеческих ресурсов);
- 4) цифровое мошенничество (дипфейки, которые позволяют создать реалистичный образ любого человека).

В результате применения ИИ в массмедиа у аудитории возник ряд вопросов. В частности, как получить гарантии от журналистов, что при использовании ИИ они лично перепроверяли факты? Как увидеть грань между фактами и вымыслом и установить явное намерение причинить вред? Как определить непреднамеренные ошибки, пародийные или сатирические высказывания в рамках бинарного анализа фактов и вымыслов? В качестве ответа

⁵ Батыров Т. Российские СМИ начали запрещать «Яндексу» использовать свой контент для нейросетей // Forbes. 2024. 17 апр. URL: <https://www.forbes.ru/tekhnologii/510579-rossijskie-smi-nacali-zapresat-andeksu-ispol-zovat-svoj-kontent-dla-nejrosetej> (дата обращения 28 августа 2024).

⁶ Гуреева Ю. Западные издания начали массово продавать свои материалы создателям ChatGPT // Российская газета. 2024. 30 мая. URL: <https://rg.ru/2024/05/30/zapadnye-izdaniia-nachali-massovo-prodavat-svoi-materialy-sozdateliam-chatgpt.html> (дата обращения 2 сентября 2024).

на возникающие вопросы стали появляться законодательные инициативы.

В 2017 г. Государственный совет КНР утвердил «План развития ИИ нового поколения». В основе документа – призыв к синхронизации усилий национальных и местных органов власти, частных компаний и академических сообществ для достижения лидерства в этой сфере к 2030 г.⁷ План является дополнением к главным стратегическим документам Китая.

В 2020 г. в США появился Закон о национальной инициативе в области ИИ 2020 г.⁸ Согласно тексту документа, стратегический приоритет США заключался в обеспечении национальной безопасности (защита от внешних угроз, возникающих в контексте развития технологий ИИ) и сохранении конкурентного лидерства на мировом рынке ИИ. В целом нужно отметить, что американские компании во многом лидируют в разработке и применении технологий ИИ, а также являются центром политики и экономики ИИ, в том числе и для стран Евросоюза.

В 2023 г. некоторые страны уже внедрили новые законы касательно возможностей применения ИИ. Так, президент США Д. Байден подписал указ о мерах безопасности, связанных с развитием ИИ⁹. Документ обязывает разработчиков предоставлять правительству результаты тестирования моделей ИИ перед их публикацией. В свою очередь, Европарламент принял первый закон об ИИ (Artificial intelligence act)¹⁰. Его цель – регулирование развития ИИ и защита людей от него. В силу он должен вступить не ранее 2025 г.

Ключевыми позициями документа является разделение систем ИИ на группы по степени риска, которые они несут гражданам. Так, к группе минимального риска относятся рекомендательные

⁷ Развитие искусственного интеллекта в странах мира: США, Китай, Великобритания // D-Russia.ru. 2019. 10 апр. URL: <https://d-russia.ru/razvitie-iskusstvennogo-intellekta-v-stranah-mira-ssha-kitaj-velikobritaniya.html?ysclid=lxszxzz5bn696480136> (дата обращения 2 сентября 2024).

⁸ H.R. 6216 (116th): national artificial intelligence initiative act of 2020 // GovTrack.us. URL: <https://www.govtrack.us/congress/bills/116/hr6216/text> (дата обращения 30 августа 2024).

⁹ В США обязали ИИ-разработчиков делиться результатами тестов по безопасности // ТАСС. 2023. 30 окт. URL: <https://tass.ru/ekonomika/19153279> (дата обращения 7 июля 2024).

¹⁰ Commission welcomes political agreement on Artificial Intelligence Act // European Commission. 2023. 9 Dec. URL: https://ec.europa.eu/commission/presscorner/detail/en/ip_23_6473 (дата обращения 5 сентября 2024).

системы с поддержкой ИИ или спам-фильтры (не представляют опасности для граждан). Вторая группа – системы высокого риска должны будут выполнять строгие требования, контролироваться людьми, соблюдать правила кибербезопасности, вести строгую отчетность (необходимо создание так называемых регулятивных песочниц). К такой группе создатели закона относят системы ИИ, применяемые в жизненно важных сферах (медицина, водоснабжение, образование, правоохранительная деятельность и таможенный контроль, система правосудия и т. п.). Группа неприемлемого риска – системы ИИ, которые должны быть под запретом.

В Западной Европе разработчиков генеративных моделей обяжут маркировать контент, чтобы аудитория могла четко разграничивать творчество человека и «машины» (изображения, созданные с помощью технологии Stable Diffusion, или дипфейки необходимо сразу заметно обозначать). Кроме того, в качестве требований к создателям также значится запрет на создание нелегального контента с помощью подобных технологий. Важно, чтобы ИИ отвечал принципам прозрачности и пользовался доверием европейцев. Еврокомиссия планирует сделать данный закон основным и собирается призвать все страны добровольно принять и внедрить у себя данный юридический документ.

В России также существует несколько отдельных норм и рекомендаций. В частности, в 2019 г. была принята Национальная стратегия развития ИИ на период до 2030 г., утвержденная Указом Президента РФ № 490 от 10 октября 2019 г. «О развитии искусственного интеллекта в РФ»¹¹. В рамках документа были представлены определения понятий «искусственный интеллект», «технологии искусственного интеллекта», «перспективные методы использования искусственного интеллекта», «смежные области использования искусственного интеллекта», «набор данных», «разметка данных» и др. В этом же году появилась национальная программа «Цифровая экономика РФ», подлежащая реализации в период с 2019 по 2024 г. Так началась реализация федерального проекта «Искусственный интеллект» с участием Минкомсвязи РФ, направленного на поддержку научных разработок в сфере ИИ, разработок программного обеспечения и российского оборудования, повышение доступности качества данных, а также разработки правовых стандартов контроля сферы ИИ.

В 2020 г. появилось Распоряжение Правительства № 2129-р об утверждении «Концепции регулирования ИИ и робототехники до 2024 года». Указанная концепция была призвана стать основой для первичного правового регулирования применения систем ИИ. Так, в тексте документа отмечено, что

концепция исходит из того, что для развития технологий искусственного интеллекта и робототехники необходимо создание регуляторной среды, комфортной для безопасного развития и внедрения указанных технологий, основанной на балансе интересов человека, общества, государства, компаний...¹²

Важными являются и аспекты, на которые, по мнению авторов Концепции, должно быть направлено регулирование. Это: контролирование новых общественных отношений, которые начинают формироваться благодаря применению ИИ и робототехники; определение правовых барьеров, затрудняющих разработку и внедрение систем искусственного интеллекта и робототехники в социальной сфере и экономике; формирование национальной системы стандартизации и оценки соответствия в области технологий ИИ и робототехники.

В октябре 2023 г. вступил в силу Федеральный закон № 408-ФЗ от 31 июля 2023 г. «О внесении изменения в Федеральный закон “Об информации, информационных технологиях и о защите информации”»¹³. Он обязывает владельцев сайтов и приложений раскрывать данные, которые используются для работы рекомендательных систем, а также их происхождение. Однако, несмотря на уже значительный пласт проработанных законодательных аспектов, в 2023–2024 гг. остается неопределенным ряд важнейших вопросов. В частности, по-прежнему необходимо создание регуляторных условий для безопасного оборота данных и защиты персональной информации. ИИ не должен использовать данные, позволяющие идентифицировать личность гражданина, а значит, необходимо создавать базы с обезличенными сведениями для последующего обучения генеративных моделей.

¹¹ Указ Президента РФ № 490 от 10 октября 2019 г. «О развитии искусственного интеллекта в РФ» // Кодексы и законы. URL: <https://www.zakonrf.info/ukaz-prezident-rf-490-10102019/> (дата обращения 17 августа 2024).

¹² Распоряжение Правительства РФ № 2129-р от 19.08.2020 об утверждении «Концепции регулирования ИИ и робототехники до 2024 года» // Гарант.ру. URL: <https://www.garant.ru/products/ipo/prime/doc/74460628/> (дата обращения 16 июля 2024).

¹³ Федеральный закон № 408-ФЗ от 31 июля 2023 г. «О внесении изменения в Федеральный закон “Об информации, информационных технологиях и о защите информации”» // Российская газета. 2023. 2 авг. URL: <https://rg.ru/documents/2023/08/02/fz408-site-dok.html> (дата обращения 11 мая 2024).

Одной из самых острых проблем касательно работы нейросетей в медиаотрасли остается проблема авторского права и интеллектуальной собственности. Как охранять сам ИИ? Как охранять объекты, созданные с помощью ИИ? Может ли ИИ охраняться в качестве объекта смежных прав, когда правовая охрана продукта, созданного ИИ, будет приравливаться к правовой охране самого ИИ?

Генеративные модели уже неоднократно обвиняли в нарушении авторских прав, это довольно чувствительная область, например, для журналистов или дикторов, озвучивающих книги. «Цифровой образ» человека никак не защищен, а значит, обладатель голоса никак не может обвинить ИИ в его использовании для озвучивания какой-то книги или другого проекта. Более того, на данный момент охрана объектов интеллектуальной собственности имеет территориальный характер и ограничена законами той страны, где выдан патент. Таким образом, если каждая страна примет свои нормативы и они не будут носить международный характер, объект интеллектуальной собственности не сможет охраняться в мире, т. е. созданный ИИ продукт будет защищен в РФ, но сможет без последствий тиражироваться, например, в США, если там не будет запретительных нормативов. В этом контексте становится очевидной необходимость международного сотрудничества и коллективных дискуссий в принятии правовых документов касательно деятельности ИИ.

На правовом уровне необходимо обеспечить прозрачность деятельности ИИ. Хорошим примером тут является YouTube, который в марте 2024 г. добавил на сайт инструмент для маркировки сгенерированных ИИ видеороликов¹⁴. Помимо этого, также необходимо четко описывать механизм функционирования ИИ, а также ключевые принципы его работы, системы, на которых учился ИИ, предупреждать о возможных рисках и угрозах при его использовании.

И одно из самых важных условий – определение системы юридической ответственности за последствия, которые возникают в сфере использования ИИ. Кто будет виноват, если произойдет серьезный инцидент с тяжелыми последствиями? Создатель ИИ? Разработчик? Тот, кто занимался обучением модели? Все эти вопросы требуют четкого и подробно прописанного алгоритма, который определит степень ответственности каждого участника

¹⁴ YouTube начал маркировать сгенерированный нейросетью контент // AdIndex. 2024. 19 марта. URL: <https://adindex.ru/news/digital/2024/03/19/321148.phtml> (дата обращения 7 июля 2024).

процесса, а значит, сделает саму процедуру применения ИИ более осмысленной.

Помимо запроса на четкое правовое регулирование ИИ, появляется необходимость внедрения этических ограничений.

Этический аспект применения ИИ в массмедиа

Этические ограничения могут относиться к разным сферам работы с ИИ – робоэтике (вопросы создания и эксплуатации систем ИИ) и машинной этике (регулирование моральных критериев работы человека с ИИ и осмысление возможностей ИИ по отношению к социуму).

В рамках темы медиа ключевое значение имеет машинная этика, регламентирующая создание некоего медиапродукта. В мире этому посвящены уже несколько десятков документов. В США, например, уже разработано более 20 различных этических кодексов, чуть меньше – в Евросоюзе [Лукина 2022, с. 682]. В России первый документ такого рода появился в 2021 г. под названием «Кодекс этики в сфере искусственного интеллекта»¹⁵. Эта инициатива объединила большое количество российских компаний, инвестирующих в технологии ИИ. В качестве задач Кодекса отмечено: создание рекомендаций для принятия решений относительно создания и использования искусственного интеллекта, снижение рисков неэтичного использования ИИ, нарушающего права человека, внедрение инструментов взаимодействия государства, разработчиков, научных организаций и общества по вопросам этики ИИ.

В 2023 г. российские компании подписали Декларацию об ответственной разработке и использовании генеративного ИИ¹⁶. Документ устанавливает этические принципы и стандарты поведения отечественных разработчиков. Его подписали 247 компаний, среди которых ключевые игроки на российском рынке технологий – Сбер, Яндекс, «ВКонтакте», «Ростелеком», МТС,

¹⁵ Кодекс этики в сфере ИИ // Альянс в сфере искусственного интеллекта. URL: <https://ethics.a-ai.ru/> (дата обращения 30 июля 2024).

¹⁶ Участники Альянса в сфере ИИ подписали декларацию об ответственной разработке и использовании генеративного ИИ в рамках Дня ИИ на Международной выставке «Россия» // Искусственный интеллект Российской Федерации. URL: <https://ai.gov.ru/mediacenter/uchastniki-alyansa-v-sfere-ii-podpisali-deklaratsiyu-ob-otvetstvennoy-razrabotke-i-ispolzovanii-gene/> (дата обращения 2 сентября 2024).

«Росатом» и др. Как и в любом этическом кодексе, положения в этом документе носят рекомендательный характер. В нем содержатся призывы к ответственности при создании и использовании искусственного интеллекта. Главным приоритетом названа защита прав и свобод человека.

В 2024 г. появилась Декларация об ответственном генеративном ИИ (подписана Сбером, Яндексом и MTS AI)¹⁷. Своей подписью согласие с основными положениями подтвердили также вузы и научные организации (Сколтех, НИУ ВШЭ, МФТИ, Институт современного программирования РАН, Нижегородский государственный университет и др.). Документ уточняет и развивает идеи, изложенные в Кодексе этики в сфере искусственного интеллекта. В частности, прописаны принципы безопасности и прозрачности, этичного отношения к определенным темам, принятия мер по упреждению и предотвращению создания фейковой информации, а также нюансы просвещения пользователей о возможностях новых технологий.

Помимо общедоступных и универсальных рекомендаций, сами средства массовой информации также стремятся урегулировать сферу взаимоотношений с ИИ. Американский ресурс Associated Press (AP) опубликовал руководство для журналистов по использованию искусственного интеллекта¹⁸. В нем указаны следующие правила: экспериментировать с чат-ботом ChatGPT можно, но с осторожностью; нельзя использовать ChatGPT для создания контента (любой контент, созданный ИИ, должен быть тщательно проверен и оценен в соответствии со стандартами AP); нельзя применять ИИ для внесения изменений в фото-, аудио- и/или видеоконтент; нельзя публиковать сгенерированные ИИ фотоизображения, если они не отражают реальность (т. е. являются фейком); журналистам запрещено загружать в программу ИИ секретную или конфиденциальную информацию; журналисты должны обязательно проверять все материалы, полученные из других источников, на предмет применения сгенерированного контента. Подобные стандарты стремятся ввести и другие новостные СМИ, поскольку случаи использования нейросетей все чаще становятся причиной публикации фейков и последующих опровержений. В россий-

¹⁷ Декларация об ответственном генеративном ИИ // Альянс в сфере искусственного интеллекта. URL: <https://ethics.a-ai.ru/genai-declaration> (дата обращения 17 августа 2024).

¹⁸ Barret A. Standards around generative AI // Associated Press. 2023. URL: <https://blog.ap.org/standards-around-generative-ai> (дата обращения 30 июля 2024).

ских медиа пока таких этических регуляторов и редакционных стандартов официально не представлено, но вполне ожидаемо их появление в ближайшем будущем, например, у деловых изданий, которые всегда дорожат своей репутацией.

В рамках разговора о восприятии ИИ в современном информационном пространстве и этических проблемах его использования проведено исследование среди будущих журналистов. В опросе приняли участие 154 студента (3-й и 4-й курсы бакалавриата очной формы обучения, 4-й и 5-й курсы бакалавриата заочной формы обучения). Время проведения исследования – май–сентябрь 2023 г. Метод – дать ответы на 14 вопросов в письменной форме. Вопросы носили открытый характер и требовали развернутого описания. Результаты позволили определить отношение юных журналистов к виртуальным коллегам.

На вопрос, случалось ли вам читать или смотреть материалы в СМИ, написанные ИИ, 45,6% ответили, что нет. 17,5% не знают этого, 7% указали, что «случалось, текст был оформлен грамотно, аккуратно, но в нем будто не было эмоций», еще 5,2% отметили, что

читали, и эти тексты были хорошо структурированы, легки для чтения, но вывод в конце материала мог быть нелогичным и отличаться от тезиса, выдвинутого в начале. Материалы были лишены авторского стиля.

Рассуждая о причинах использования ИИ в работе журналиста, 22,8% отметили, что можно применять для свободы от рутинных задач, 14% указали, что причиной может стать нехватка временных или человеческих ресурсов, когда требуется высокая оперативность, и столько же отметили – что для более точного написания текстов или поиска информации. 8,7% считают, что могут применять нейросети для быстрой проверки фактов, создания визуала. Студенты среди прочего называли также поиск идей для статей, работу со статистикой, большим количеством данных, возможность придать материалу объективный характер, монтаж, обработку фото и звука, создание заголовков. 66,7% опрошенных никогда не видели виртуального телеведущего. Еще 7% видели, но он их не впечатлил, поскольку пока очень заметно, что это не человек. Остальные варианты носили единичный характер. Но примечательно, что есть диаметрально противоположные точки зрения: «Телеведущего нет, в ТикТоке – да. Выглядит неестественно, мимика 5/10, но если доработать, то будет выглядеть здорово» и «В ТикТоке – не отличить от живого человека. Считаю, что это спасет людей от выгорания».

Или вот такой пример:

Да, сложно отличить от настоящего. Разница видна была только в том, что живой человек может ошибиться или оговориться, а робот нет» и «Видела на китайском ТВ. “Эффект зловещей долины”».

В качестве важнейших этических аспектов при работе с ИИ не удалось выделить какие-то одинаковые по сути мысли. Большинство респондентов (64%) не смогли ответить на этот вопрос. Остальные ответы носили единичный характер. Среди них:

«Проработать аспекты понимания ценности человека, его прав и эмоций»; «Стоит сначала разрешить проблему рабочих мест, что делать людям, которых заменят роботы? Баланс между трудом человека и работой ИИ»; «Важно понимание культуры речи, чтобы не было никаких оскорблений, грубых выражений»; «Авторские права, уважение к чужому труду. Рисунки художников, тексты, фотографии, взятые ИИ, могут считаться кражей собственности»; «Обязательно указывать, если материал полностью написан ИИ или с помощью ИИ»; «Четко указывать, о чем может говорить ИИ, а что не надо обсуждать»; «Запрет на использования ИИ для нанесения вреда другим людям»; «ИИ должен иметь нейтральную внешность и голос, чтобы не было никаких провокационных ситуаций»; «Необходимы четкие подходы к подаче информации, чтобы не пугать и не шокировать людей»; «ИИ должен показывать все точки зрения, давать объективную информацию».

Резюмируя ответы касательно этических аспектов применения ИИ в массмедиа, стоит отметить, что опрашиваемые хотели бы видеть четкую, непредвзятую и объективную информацию от нейросетей, а также четкие маркеры материалов, созданных не человеком. Не менее важным респонденты считают и уважительные и деликатные ответы на любые вопросы от пользователей, полное отсутствие грубости или дискриминации по любым признакам. Однако тревожным представляется тот факт, что более половины опрошенных вовсе не отметили никаких этических требований, что говорит об отсутствии осознания важности этого аспекта работы ИИ.

В целом можно заключить, что отношение к работе ИИ в качестве телеведущего – критическое, порой даже открыто негативное. Заметен крайне низкий уровень доверия: в ответах на 8 из 14 вопросов респонденты упоминали отсутствие доверия к «машине», а также отсутствие эмоций. В 10 вопросах из 14 выражение «эффект зловещей долины» также упоминалось в том или ином контексте.

Таким образом, одной из важных этических проблем становится обеспечение прозрачности работы ИИ, поскольку люди не могут доверять «черному ящику», ведь все непонятное отталкивает и пугает. Касательно виртуальных моделей ИИ, как показало исследование, возникает «эффект зловещей долины» и отторжение. Решить эту проблему мог бы так называемый объяснимый ИИ (Explainable AI, XAI).

Закономерным называют направление исследования в области ИИ, цель которого – создание систем и моделей, способных объяснять свои действия и принимать решения понятным для людей образом для повышения уровня доверия аудитории к ИИ¹⁹. Для этого используются методы визуализации, интерактивные проекты с подсказками, упрощенные алгоритмы. Бесспорными преимуществами внедрения такой технологии станет решение важных этических проблем за счет повышения производительности (более оперативное выявление ошибок в модели), укрепления доверия со стороны пользователей, а также снижения регуляторных и прочих рисков.

Заключение

Подводя итог, получаем, что законодательных актов, регулирующих деятельность ИИ, очень мало, этические аспекты находятся в стадии активного обсуждения, а уровень недоверия пользователей становится уже ощутимым. На современном этапе возможности технологий искусственного интеллекта обширны, но не универсальны. Работу ИИ однозначно необходимо перепроверять и контролировать. Более того, ИИ всегда требует постановки задачи со стороны человека, поскольку существует множество разных систем ИИ, а для эффективного результата (анализа статистики или составления прогнозов, рейтингов) необходимы четкость и корректная постановка вопроса.

Кроме того, использование материалов ИИ чревато обвинениями в плагиате, поскольку ничего нового нейросеть создать не может, а может только сложить из уже существующего что-то общее (собрать картину из кусочков – что и является ключевым аспектом генерации). И тут особенно остро встает этический аспект. Можно ли выдавать за свое произведение то, что придумано до этого разными людьми?

¹⁹ Решетникова М. Интеллект, понятный каждому // РБК. 2024. № 01–02. С. 64.

Отдельная угроза – дипфейки, которые могут принимать совершенно разные формы. Фотодипфейки уже никого не удивляют. Это распространенное явление, которое не считается чем-то неординарным. Аудиодипфейки способны подделать голос любого человека, причем чем более известен этот человек, тем выше вероятность модуляции его голоса. Этим активно пользуются мошенники для обмана граждан. Видеодипфейки становятся все более качественными, злоумышленники способны подделать видео любого человека. И их действительно невозможно отличить от оригинала. В СМИ неоднократно освещались истории, когда мошенники подделывали голос родственника или друга и даже присылали короткое видео (так называемые кружочки) в Telegram, на котором был знакомый «жертвы». Поскольку главной целью мошенников является кража личной информации или просьба перевести деньги, то эти случаи можно распознать и пресечь, просто перезвонив другу или родственнику самостоятельно. Об этом обязательно необходимо говорить молодым людям, которые в порыве эмоциональной реакции могут быть обмануты. Однако в перспективе можно предположить, что такие дипфейки могут применять для манипуляции и уговоров совершать противозаконные действия, вербовки людей в различные организации или даже секты, искажения морально-нравственных ориентиров и жизненных ценностей, и тогда может получиться, что ИИ будет управлять человеческим разумом, а не наоборот.

Проанализированные выше этические и правовые проблемы использования ИИ в массмедиа позволяют сделать вывод о необходимости, в первую очередь, более настойчивого привития осознанного потребления медиаконтента аудиторией. Зрителю или читателю необходимо критически относиться к любой информации. Теперь применение ИИ становится все более привычным для журналистов в любых СМИ, и уже нельзя быть уверенным, что текст был написан профессионалом своего дела, а не сгенерирован нейросетью. Только осознанное и внимательное отношение к поступающим сведениям позволит сохранять ясность и четкое понимание происходящего в стране и в мире.

Литература

Лукина 2022 – Лукина М.М., Замков А.В., Крашенинникова М.А., Кульчицкая Д.Ю.

Искусственный интеллект в российских медиа и журналистике: к дискуссии об этической кодификации // Вопросы теории и практики журналистики. 2022. № 4. С. 680–694.

- Малыгина 2018 – *Малыгина Л.Е.* «Виртуальные собеседники»: перспективы развития телевизионного промодискурса // Верхневолжский филологический вестник. 2018. № 4 (15). URL: <https://cyberleninka.ru/article/n/virtualnye-sobesedniki-perspektivy-razvitiya-televizionnogo-promodiskursa> (дата обращения 30 июля 2024).
- McCarthy 2007 – *McCarthy J.* What is artificial intelligence? URL: <http://jmc.stanford.edu/articles/whatisai/whatisai.pdf> (дата обращения 24 августа 2024).

References

- Lukina, M., Zamkov, A., Krashennnikova, M. and Kul'chitskaya, D. (2022), "Artificial intelligence in Russian media and journalism. On a discussion over ethical codification", *Theoretical and Practical Issues of Journalism*, no. 4, pp. 680–694.
- Malygina, L. (2018), " 'Virtual conversations': Perspectives for the development of television promotion discourse", *Verhnevolzhski Philological Bulletin*, vol. 15, no. 4, available at: <https://cyberleninka.ru/article/n/virtualnye-sobesedniki-perspektivy-razvitiya-televizionnogo-promodiskursa> (Accessed 30 July 2024).
- McCarthy, J. (2007), *What is artificial intelligence?* available at: <http://jmc.stanford.edu/articles/whatisai/whatisai.pdf> (Accessed 24 Aug. 2024).

Информация об авторе

Дарья В. Неренц, кандидат филологических наук, доцент, Российский государственный гуманитарный университет, Москва, Россия; Россия, Москва, 125047, Миусская пл., д. 6; ya.newlevel@yandex.ru

Information about the author

Daria V. Nerents, Cand. of Sci. (Philology), associate professor, Russian State University for the Humanities, Moscow, Russia; 6, Miusskaya Sq., Moscow, 125047, Russia; ya.newlevel@yandex.ru