

Международный опыт регулирования маркировки контента, созданного с помощью систем ИИ, в СМИ

Научный руководитель – Левченко Татьяна Григорьевна

Валуева Ксения Андреевна

Студент (бакалавр)

Московский государственный университет имени М.В.Ломоносова, Факультет мировой политики, Кафедра информационного обеспечения внешней политики, Москва, Россия

E-mail: valueva.ksen@yandex.ru

С распространением ИИ и его внедрением в повседневную жизнь возникает вопрос о частоте применения этих технологий в СМИ. Предложение о маркировке синтезированного контента стало важным шагом в обеспечении прозрачности и укреплении доверия аудитории. В России вопрос о законодательном регулировании маркировки контента, сгенерированного ИИ активно обсуждается уже на протяжении двух лет. Еще в 2023 году академическое сообщество обращалось в профильное министерство с таким предложением [3], а позднее этот вопрос был поднят в законодательных органах РФ – Государственной Думе [1] и Совете Федерации [2]. Однако, на данный момент этот процесс находится на этапе инициативы и еще не оформлен в законодательный акт.

Между тем, в международной практике уже существуют такие инициативы. Одним из ключевых примеров является AI Act [10], принятый в марте 2024 года Европейским союзом. Этот документ стал первым правовым актом, охватывающим все вопросы применения ИИ и разделившим такие системы на несколько категорий по степени риска. Так, генеративные технологии были отнесены к средней категории, но при условии, что все сгенерированные материалы обязаны быть маркированы цифровыми водяными знаками и метаданными, которые обеспечивают аудиторию необходимой информацией о происхождении контента и внесенных в него изменениях. В сфере маркировки материалов США опередили Европейский союз на год, приняв AI Labeling Act в 2023 году [9], обязывающий ставить водяные знаки и раскрывать информацию о контенте, созданном при помощи ИИ, особенно в контексте политической рекламы и новостей. Однако, этот законодательный акт не носит такого всеобъемлющего характера, как AI Act ЕС.

Данные законодательные инициативы были основаны, в том числе, на анализе общественного мнения, который государства запрашивают у исследователей для соответствия политики потребностям граждан. В 2023 году Reuters Institute [5] провел исследование в 6 странах мира, которое показало, что большинство респондентов из разных возрастных категорий поддерживают обязательную маркировку контента, созданного с помощью ИИ, особенно в новостных материалах, а также в статьях, связанных с политикой и социальной сферой. В 2024 году институт изучения журналистики Reuters выпустил ежегодный отчет, основанный на опросах почти 100 000 человек в 47 странах мира [4]. Он продемонстрировал, что тенденции, выделенные в исследовании 2023 года, развиваются - в мире растет обеспокоенность, связанная с использованием ИИ в производстве новостей, а самый большой процент респондентов недоволен использованием ИИ для создания контента и уверен в необходимости маркировки. Однако, несмотря на жесткую позицию общественности, исследование Б. Тоффа и Ф. Саймона 2023 года [7] говорит о том, что в научных кругах существует опасение, что маркировка контента может негативно сказаться на доверии аудитории. Правдивые, но сгенерированные ИИ материалы, могут быть проигнорированы из-за общего негативного отношения к синтезированному контенту.

На международном уровне до сих пор не принято актов, которые бы регулировали применение систем ИИ в средствах массовой информации, из-за этого вопрос стандартизации в области маркировки контента, созданного с помощью ИИ, остается острым. Отсутствие международной координации в сфере правового регулирования ИИ и принятие в некоторых странах собственных подходов, затрудняют выработку общих документов. При этом, технологическое развитие значительно опережает законодательный процесс, что может привести к снижению эффективности введенной маркировки, нарушению прав, а также к дестабилизации общества из-за снижения доверия правдивости информации в СМИ [8]. В отсутствии четких стандартов возникают не только риски, но и трудности с установлением ответственности за фальсификацию и другие нарушения.

Маркировка контента, созданного с использованием систем ИИ, является неотъемлемой частью обеспечения контроля за прозрачностью и достоверностью материалов, публикуемых СМИ. Приведенный анализ общественного мнения и существующих инициатив подтвердил необходимость дифференцированного подхода, учитывающего степень опасности, которую несут разные системы искусственного интеллекта. Такие меры будут призваны не только минимизировать риски, но и укрепить доверие аудитории к системам СМИ в цифровую эпоху.

Различные технические механизмы, включая водяные знаки и метаданные, доказали свою эффективность, однако их внедрение сталкивается с проблемами надежности и унификации [6]. Правовые акты, которые уже были приняты в США и ЕС создают основу для разработки международных стандартов регулирования для усиления информационной глобальной безопасности.

Для достижения значимого прогресса необходимы скоординированные усилия, которые могут привести к созданию стабильного и безопасного медиапространства.

Источники и литература

- 1) В Госдуме начали работу над концепцией закона о маркировке контента нейросетей. – URL: <https://tass.ru/ekonomika/17930785> (дата обращения: 09.03.2025)
- 2) Контент, созданный искусственным интеллектом, предложили маркировать. – URL: <https://ria.ru/20230424/ii-1867306770.html> (дата обращения: 09.03.2025)
- 3) Минцифры предложили ввести маркировку контента, созданного с помощью нейросетей. – URL: <https://tass.ru/ekonomika/17746919> (дата обращения: 09.03.2025)
- 4) Arguedas, A. R. Public Attitudes towards the Use of AI in Journalism // Reuters Institute Digital News Report 2024. – 2024 – С. 39-43
- 5) Fletcher, R., Nielsen, R. K. What Does the Public in Six Countries Think of Generative AI in News? // Reuters Institute for the Study of Journalism. – 2024 – С. 10-28
- 6) Molavi Vasse'i, R., Udoh, G. In Transparency We Trust? Evaluating the Effectiveness of Watermarking and Labeling AI-Generated Content, Mozilla – 2024 – С.20-35
- 7) Toff, B., Simon, F. M. "Or They Could Just Not Use It?": The Paradox of AI Disclosure for Audience Trust in News // The International Journal of Press/Politics- 2023 – С. 1-19
- 8) Wittenberg, C., Epstein, Z., Berinsky, A. J., Rand, D. G. Labeling AI-Generated Content: Promises, Perils, and Future Directions // An MIT Exploration of Generative AI. – 2024 – С. 5 - 10
- 9) AI Labeling Act of 2023. URL: <https://www.congress.gov/bill/118th-congress/senate-bill/2691/text> (дата обращения: 09.03.2025)

- 10) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).
– URL: <http://data.europa.eu/eli/reg/2024/1689/oj> (дата обращения: 09.03.2025)