

в этой области остаются актуальными, так как рост трафика, появление новых сервисов и увеличение количества подключенных устройств продолжают повышать требования к эффективности использования радио ресурсов даже в условиях развития 5G.

УДК 004.83

## **РАЗВИТИЕ И ПРИМЕНЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СОЗДАНИИ МУЛЬТИМЕДИЙНОГО КОНТЕНТА**

**Гомзяков О. А., магистрант**

*Научный руководитель – канд. техн. наук, доцент Макареня С. Н.  
Белорусский национальный технический университет  
Минск, Республика Беларусь*

**Аннотация.** В статье рассматривается текущее состояние и перспективы интеграции ИИ в медиаиндустрию, анализируются ИИ-инструменты, которые используются в процессе создания видео, графическом дизайне. Целью работы является исследование влияния ИИ на эффективность и качество производства контента.

**Ключевые слова:** искусственный интеллект, мультимедийный контент, ИИ-инструменты, нейросеть.

Искусственный интеллект (ИИ) – это широкая область компьютерных наук, которая занимается созданием систем, способных имитировать когнитивные функции человека, такие как обучение, рассуждение, восприятие, решение проблем и понимание языка. Проще говоря, это технология, которая позволяет компьютерам «думать» и «учиться».

За все время развития ИИ значительно и комплексно изменялся в сфере производства видео и звука. Он прошел путь от узкоспециализированных инструментов до комплексных, генеративных систем. При этом возникали этические вопросы, технологические проблемы, вопросы безопасности и социальные проблемы.

В данной статье будет рассмотрен период за последние 10 лет, когда ИИ начал стремительно развиваться и совершенствоваться.

В середине 2010-х годов ИИ обладал функционалом для автоматизации рутинных задач, таких как распознавание речи, распознавание изображений, системы рекомендаций, обнаружение мошенничества, оптимизация логистики [1].

С 2015 года по 2018 год проходил этап автоматизации процесса производства видео, который включал:

1. Появление умного монтажа: появились первые алгоритмы, которые могли автоматически монтировать видео на основе музыки. Одним из таких приложений было Magisto, которое анализировало видео, подбирало лучшие моменты и монтировало его.

2. ИИ научился отслеживать движения: начали развиваться технологии, которые позволяли отслеживать движения, объекты и лицо в кадре. Развитие в этом направлении упростило работу с наложением визуальных эффектов и графики.

3. Улучшилось изображение: автоматическую цветокоррекцию и стабилизацию изображения получили программы для работы с видео, которые работали на основе машинного обучения.

Далее с 2019 года по 2022 год появлялись новые генеративные модели, а также происходило повышение качества генерируемого контента. Появились новые технологии и произошло улучшение более ранних. Все эти новые возможности искусственного интеллекта и стали причиной для появления вопросов, связанных с этичностью. Примеры новых улучшений и возможностей:

1. Создание дипфейков и замена лиц: технологии искусственного интеллекта позволили создавать качественные подмены лиц, с помощью чего появилась возможность для омоложения актеров в кинопроизводстве или создания дублеров.

2. Произошло улучшение и масштабирование видео: инструменты искусственного интеллекта позволили масштабировать видео низкого разрешения до более высокого, при этом улучшая детализацию и убирая посторонние шумы. Примером такого инструмента является Topaz Video AI.

3. ИИ научился генерировать видео по тексту: начали появляться новые генераторы на основе искусственного интеллекта, которые генерируют видео по текстовому запросу, только все они были короткими и не всегда с высоким качеством и логикой.

Затем с 2023 года по 2025 год генерация видео улучшилась, и стало сложнее отличать контент, сделанный искусственным интеллектом, от того, который сделан человеком. Примером прорыва в развитии генеративного искусственного интеллекта является:

1. Улучшение качества видео по текстовому запросу: инструменты для генерации видео научились создавать реалистичные и качественные видео по текстовому запросу, при этом логика и длительность таких видео значительно выросли.

2. Полная автоматизация процесса: современные инструменты искусственного интеллекта позволяют выполнять весь цикл производства, начиная от генерации сценария раскадровки и заканчивая созданием финального видео и наложением спецэффектов.

Что касается звукового производства, то изменения начались в 2015 году и длились они до 2018 года. К ним относятся:

1. Умное микширование и мастеринг: появились первые плагины искусственного интеллекта, которые могли автоматически мастерить и микшировать треки, анализируя музыкальный материал. Одним из примеров такого плагина является LANDR.

2. Улучшение голоса и шумоподавление: развивались технологии для удаления фонового шума и улучшения качества записи голоса.

3. Синтез голоса: улучшались системы синтеза речи, что позволило создавать более натурально звучащие голосовые дорожки для озвучивания и голосовых ассистентов.

С 2019 года по 2022 год появились инструменты искусственного интеллекта, которые могли генерировать и клонировать голоса:

1. Клонирование голоса: технологии, которые позволяли клонировать голоса людей достигли высокого уровня реализма. Эти технологии позволили создавать новые голосовые озвучки и дубляжи.

2. TTS (перевод текста в речь): с развитием системы TTS начали звучать более естественно, что положило начало использования их для создания аудиокниг, подкастов и видеоконтента.

3. Разделение аудиодорожек: ИИ научился автоматически разделять музыку на отдельные элементы (голос, инструменты), что упростило ремикширование и обработку.

С 2023 года по 2025 год искусственный интеллект получил возможность генерировать музыку и выполнять аудиодизайн:

1. Генерация музыки: сервисы на основе искусственного интеллекта начали генерировать полноценные музыкальные треки в разных жанрах по текстовому описанию.

2. Перевод текста в аудио: появились модели, способные генерировать реалистичные звуковые эффекты (например, шум дождя, шаги, звуки города) по текстовому запросу.

3. Анализ аудио: искусственный интеллект используется для анализа аудиопотоков в реальном времени, например, для обнаружения проблем на трансляциях или адаптации звука в зависимости от ситуации.

В настоящее время имеется достаточное количество систем на основе искусственного интеллекта и их количество продолжает стремительно расти. Самыми современными системами являются: Sora 2 от компании OpenAI, Veo 3.1 от компании Google, Luma AI, Kling, Pika Labs, Vidu AI, Runway (Gen-2, Gen-4), Runway Motion Brush и другие. Все они оказывают огромное влияние на количество и качество контента в интернете. По статистике в настоящее время приблизительно 57 % всего контента в интернете сгенерировано с помощью ИИ [2].

Лучшими ИИ для генерации мультимедийного контента являются Sora и Veo. Но они обладают рядом отличий друг от друга. Данные отличия приведены в табл. 1.

Таблица 1 – Сравнительный анализ систем генерации мультимедийного контента

<b>Критерий</b>	<b>OpenAI Sora 2</b>	<b>Google Veo 3.1</b>
1	2	3
Реализм	«Симулятор мира». Благодаря улучшенной физике объектов, модель лучше передает взаимодействие объектов в кадре, движения и последовательность действий, что делает ее идеальной для сложных, повествовательных видео	«Коммерческий реализм». Создание высококачественных, полированных кадров с невероятно реалистичными освещением, текстурами и материалами. Идеально для профессиональной рекламы и кинематографии

Окончание таблицы 1

1	2	3
Длительность видео	Позволяет создавать более длинные видеоклипы, до 15–25 секунд, что лучше подходит для создания коротких фильмов или контента для социальных сетей	Ориентирован на короткие, до 8 секунд кинематографические клипы. Для более длинных видео нужно использовать специальные инструменты для сшивания или расширения кадров
Аудио	Генерирует синхронизированный звук, включая диалоги и звуковые эффекты, которые соответствуют действиям в кадре. Особенно хорошо подходит для озвучивания историй	Также генерирует синхронизированный звук и диалоги, включая высококачественные саундскейпы и фоновую музыку, идеально подходящие для кинематографических сцен
Инструменты редактирования	Включает встроенные инструменты (например, Remix и Recut) для быстрого внесения корректировок на уровне сцены до экспорта. Также появилась функция Sameo, позволяющая встраивать себя или других людей в сгенерированные видео	Меньше встроенных инструментов для редактирования. Veo, скорее, фокусируется на генерации, а для последующей доработки видео предполагается использование сторонних программ
Разрешение	Поддерживает разрешение до 1080p. Хотя этого достаточно для большинства задач, это уступает максимальному разрешению Veo	Поддерживает генерацию видео до 4K, что идеально для профессиональной обработки и производства
Водяные знаки	Использует сочетание видимого водяного знака и метаданных C2PA для маркировки контента	Использует SynthID – невидимый водяной знак для идентификации ИИ-контента

В конечном итоге выходит, что ИИ прошел путь от обычного инструмента к полноценному соавторству во всем, что мы делаем

и производим сегодня. В настоящее время он обладает следующими характеристиками:

1. Автоматизация и эффективность: ИИ значительно сокращает временные и финансовые затраты на производство контента, автоматизируя такие процессы, как монтаж видео, создание анимации, обработка изображений и звуковое сопровождение.

2. Творческий потенциал и персонализация: модели генеративного ИИ, такие как Veo и Sora 2, предоставляют создателям контента новые творческие инструменты. Они позволяют визуализировать сложные идеи, генерировать высокореалистичные или стилизованные видео из текста и изображений, а также создавать уникальные визуальные эффекты. ИИ также помогает персонализировать контент, анализируя предпочтения аудитории для создания более релевантных материалов.

3. Увеличение доступности: ИИ снижает барьеры для входа в креативные индустрии, предоставляя даже небольшим студиям и отдельным авторам доступ к технологиям, которые ранее были доступны только крупным компаниям. Это способствует созданию равных конкурентных возможностей.

4. Гиперреализм и повествование: модели, такие как Veo, ориентированы на создание реалистичных, кинематографичных кадров, в то время как Sora 2 делает акцент на более сложном взаимодействии объектов и повествовании. Это позволяет авторам выбирать инструменты в зависимости от своих целей: для реалистичной рекламы или для креативных, сюжетных роликов.

5. Встроенные функции: современные ИИ-модели, такие как Veo и Sora 2, все чаще включают встроенные функции для генерации синхронизированного аудио, редактирования и стилизации. Это упрощает рабочий процесс, делая его более интуитивным.

На ряду с преимуществами использования ИИ существует ряд проблем и вызовов для человечества:

1. Этические дилеммы: использование ИИ поднимает серьезные этические вопросы, включая создание дипфейков, распространение дезинформации и предвзятость алгоритмов.

2. Вопросы авторского права: возникают сложности с определением авторства и защитой прав на контент, созданный ИИ, а также с несанкционированным использованием чужих данных для обучения нейросети.

3. Социальные последствия: внедрение ИИ может привести к изменению рынка труда в креативных сферах, но одновременно создает новые профессии, связанные с управлением ИИ-инструментами [3].

В заключение хотелось бы отметить, что ИИ в настоящее время быстро развивается и проникает во все сферы нашей жизни. Ниже приведем некоторые возможные перспективы его развития:

1. Коллаборация человека и ИИ: будущее индустрии – в эффективном сотрудничестве между человеком и ИИ, где ИИ берет на себя техническую работу, а человек отвечает за творческую составляющую, сюжет и эмоции.

2. Непрерывное развитие технологий: скорость развития ИИ-моделей не снижается, и их возможности будут только расширяться, делая контент еще более персонализированным и доступным.

3. Регулирование и медиаграмотность: для решения этических и правовых проблем потребуются создание новых стандартов и правил, а также повышение медиаграмотности аудитории, чтобы отличать контент, созданный ИИ, от человеческого.

### **Список использованных источников**

1. The Evolution and Future Perspectives of Artificial Intelligence Generated Content / Y. Hu [et al.] // arXiv.org. – 2024. – URL: <https://arxiv.org/abs/2412.01948> (date of access: 05.11.2025).

2. A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT / Y. Cao [et al.] // arXiv.org. – 2023. – URL: <https://arxiv.org/abs/2303.04226> (date of access: 03.11.2025).

3. Imagen Video: High Definition Video Generation with Diffusion Models / J. Ho [et al.] // arXiv.org. – 2022. – URL: <https://arxiv.org/abs/2210.02303> (date of access: 03.11.2025).