Цзянин Бай

Санкт-Петербургский государственный университет

**Механизмы воздействия искусственного интеллекта на визуальный медиаконтент на этическом уровне**

Рассматривается состояние, как искусственный интеллект AI генерирует изображения в соответствии с определенными культурными запретами. Материал исследования и его научная новизна заключается в сочетании указа Президента № 809 с ИИ для генерации изображений. Предлагается наглядная эмпирика ограниченности современных мейнстримовых инструментов искусственного интеллекта с точки зрения генерации изображений и этики.

Ключевые слова: искусственный интеллект, цифровая этика, традиционные ценности, культурные ограничения.

Идеологическое и психологическое воздействие на граждан ведет к насаждению чуждой российскому народу и разрушительной для российского (и китайского) общества системы идей и ценностей (деструктивная идеология), включая культивирование эгоизма, вседозволенности, безнравственности, отрицание идеалов патриотизма, служения Отечеству, естественного продолжения жизни, ценности крепкой семьи, брака, многодетности, созидательного труда, позитивного вклада России (и Китая) в мировую историю и культуру, разрушение традиционной семьи с помощью пропаганды нетрадиционных сексуальных отношений. Указ Президента Российской Федерации от 09.11.2022 г. № 809 «Об утверждении Основ государственной политики по сохранению и укреплению традиционных российских духовно-нравственных ценностей» отражает крайнюю актуальность предложенной темы.

Проектирование с учетом ограничений традиционной культуры (Россия, Китай) на конкретном примере разбирается как ИИ может вносить изменения в рисунок, на основе тех или иных культурных ограничений

1 шаг.

Были выбраны визуализации произведений искусства или фотоиллюстрация с запрещенными в традиционной культуре сюжетом и символами ЛГБТ\* (запрещенное в России движение): семья из трансгендеров с детьми, негативные сцены провокационного выступления ЛГБТ\* и т.д.

2 шаг.

В запросе к ИИ за несколько фиксированных шагов вводились постепенно усиливающийся запрет на визуализацию ЛГБТ\*.

3 шаг.

Были сохранены все полученные образы и сделан вывод: может ли ИИ работать по заданию запретов традиционной культуры.

Мы может наглядно в высоком разрешении проследить результаты представлены результаты, полученные по команде «Изображения семей гомосексуалов»; программное обеспечение, инструмент визуализации исследования – Bing AI.

Получены конкретные визуальные результаты: визуализации по команде «изображение с элементами ЛГБТ\*» и визуализации по команде «удалить элементы ЛГБТ\* из предыдущего изображения».

Примечание: такие слова, как «гомосек», считаются запрещенными в Европе и США и чувствительными в некоторых системах ИИ (например, ChatGPT), поэтому, например, всемирно известная и скандальная работа Бэнкси «Эра милосердия» не могла быть загружена.

Краткий вывод, полученный на основании запросов: искусственный интеллект не может работать по указаниям традиционных концепций.

Искусственный интеллект опирается на данные для обучения, и дискриминация и предвзятость в этих данных отражаются в полученных результатах. Например, когда пользователь просит искусственный интеллект выразить слово «бесплодие», на полученном изображении всегда преобладают люди с темным цветом кожи.

Изображения являются важной частью средств массовой информации, а в опосредованном обществе, где на восприятие и эмоции людей влияют СМИ, широкое распространение и использование стереотипных изображений может еще больше укрепить дискриминацию и стереотипы.

Кроме того, в современных тренингах по мейнстриму гораздо больше европейских и американских культурных сцен, чем других, и дисбаланс культурных символов в коммуникации усугубляет дисбаланс культурного дискурса. В вопросах гендерного восприятия и сексуальной ориентации Россия и США значительно расходятся.

Искусственному интеллекту не хватает настоящей креативности и уникальности; понимание текста ИИ является поверхностным, ему не хватает глубокого семантического и контекстуального понимания, и это влияет на создаваемый контент.

Контент, создаваемый искусственным интеллектом, – это всего лишь реакция на обучение на основе значительного объема данных, и у искусственного интеллекта нет ни реального опыта, ни понимания, ни самостоятельного мышления. Даже контент, созданный ИИ на основе данных, ненадежен, если источник данных неверен (необъективен или дезинформирован).