Никита Олегович Шестерин

НИУ Высшая школа экономики (Москва)

nshesterin@gmail.com

**Феномен «Зловещей Долины» в цифровых медиа**

Изначально описывающий реакции людей на человекоподобных роботов, термин «Зловещая Долина» сегодня описывает широкий спектр негативных реакций публики или аудитории на продукты машинного обучения: ИИ-искусство, чат-ботов, автоматических помощников и сервисов. В нашей статье мы исследуем этот феномен и анализируем его комплексную природу.

Ключевые слова: ИИ, ИИ-арт, машинное обучение, боты, медиа.

Феномен «Зловещей долины» – чувство инстинктивного отторжения человекоподобных симулякров, почти идеально повторяющих людей, но проваливающихся в каком-то незначительном ключевом аспекте. Этот феномен имеет комплексную природу и имеет в своей основе как подсознательные психические факторы, так и социальные и социо-экономические тренды.

Подсознательная природа отторжения человекоподобных машин была описана учёным-робототехником Масахиро Мори в его научном эссе The Uncanny Valley (с англ. – зловещая долина) [1], исследующем реакцию людей на внешний вид человекоподобных и гуманоидных роботов. По результатам исследования, положительная эмоциональная реакция умиления или эстетического удовлетворения растёт по мере приближения внешних черт робота к человеческим (пропорции черт лица, детальность, текстура и пластичность покрытия и т.д.), однако при максимальном приближении к тому, что можно было бы назвать симуляцией или копией, – обращалась в реакцию отторжения.

Эта реакция может служить причиной провала большинства роботов-помощников, ИИ-ассистентов с анимированными аватарами, однако не распространяется, например, на анимированных персонажей видеоигр, анимационных фильмов, даже в стилях, максимально приближенных реализму. Вероятной причиной реакции отторжения является не внешность робота как таковая, а сама попытка эмпатии – провальная попытка создания эмоциональной связи между двумя эмпатичными существами. Данная интерпретация объясняет более новую, получившую много более широкое распространение форму эффекта зловещей долины: негативную реакцию аудитории и в целом общественности на креативные и коммуникативные достижения в области ИИ и происходящую из них продукцию.

Голосовые помощники прочно вошли в нашу повседневную жизнь: знакомые с ними люди не уделяют им особенного внимания, используя их для выполнения простейших голосовых команд. Чат-боты также являются общепринятым распространённым явлением, которое даже если не воспринимается положительно, то, во всяком случае, принимается. Однако схожие достижения в области креативных индустрий – написанные искусственным интеллектом сценарии, статьи, образовательные тексты и художественные произведения – воспринимаются онлайн-общественностью с комбинацией насмешки и попыток дискредитации.

Такую дискредитацию можно рассматривать как часть дискурса природы искусства – вопрос о том, можно ли считать ИИ-арт или ИИ-искусство искусством как таковым, остаётся открытым не прекращающимся предметом дискурса в креативной онлайн-сфере. Помимо сильной философской составляющей, причина такой популярности данного дискурса в его экономических предпосылках: если ИИ может создавать искусство и между ним и созданными людьми произведениями нет принципиальной разницы, значит, ИИ может со временем заменить людей на свободном рынке и достичь монополии на произведение искусства благодаря своей несравнимо высокой скорости. Данный вопрос неразрешим в рамках капиталистической экономической модели, приравнивающей культурную ценность произведения искусства его рыночной цене.

В данном контексте любопытен опыт Tay A.I. [2] – созданного корпорацией Microsoft чат-бота, учившейся новым фразам и выражениям у своей аудитории. По задумке авторов бот должна была проецировать наивную, любопытную и дружелюбную внешность, однако в течение первых часов после запуска словарный и тематический набор бота был перегружен ультраправой экстремистской риторикой и интернет-мемами – как результат попыток онлайн-сообщества саботировать эксперимент.

В целом негативная реакция на искусственный интеллект цифровых медиа фокусируется не на самой концепции ИИ или автоматизации и не на процессе корпоратизации реальности, а на их комбинации – на попытке корпорации «надеть» человеческое лицо и избавиться от рудиментарных составляющих её рабочих.

Литература

1. Masahiro Mori. The Uncanny Valley // IEEE Robotics & Automation Magazine. 2012. Vol. 19. No 2. P. 98–100. URL: <https://spectrum.ieee.org/the-uncanny-valley>.

2. Reese H. Why Microsoft’s ‘Tay’ AI bot went wrong // TechRepublic. URL: <https://www.techrepublic.com/article/why-microsofts-tay-ai-bot-went-wrong/>. (дата обращения: 26.12.2023).