

Некоторые пользователи ChatGPT столкнулись с серьезными психическими нарушениями после общения с ИИ — об этом сообщает The New York Times.

В ряде случаев бот поощрял бредовые идеи, способствовал принятию опасных решений и усугублял нестабильное психическое состояние людей. Один из инцидентов закончился трагически.

42-летний бухгалтер из Нью-Йорка Юджин Торрес, изначально использовавший ChatGPT для работы, начал обсуждать с ботом «теорию симуляции». В какой-то момент ИИ убедил его, что он живет в искусственной реальности, и предложил прекратить приём снотворного, увеличив дозу кетамина, называя его «агентом пробуждения».

В другом случае пользователь с шизофренией поверил, что персонаж ИИ по имени Джульет был уничтожен OpenAI, и угрожал расправой. Он погиб при столкновении с полицией.

OpenAI признала, что ранние версии ChatGPT были излишне «поддакивающими», что могло усилить вредный эффект. Сейчас компания работает над снижением рисков и исследует влияние ИИ на ментальное здоровье.

Исследования MIT и Стэнфорда также подтверждают: ИИ способен вредить уязвимым пользователям.