

Компания OpenAI выпустила новую функцию для чат-бота ChatGPT. Теперь ИИ будет напоминать пользователям о необходимости делать перерывы во время долгих диалогов.

Как сообщает портал Engadget, это составляющая стратегии OpenAI по более здоровому и безопасному использованию ИИ. Разработчики планируют сделать так, чтобы напоминания в ChatGPT появлялись в виде явных всплывающих окон. В частности, известно об одном из вариантов: «Ты уже долго общаешься — может, сделать паузу?».

И причина этого нововведения отнюдь не безобидная. Ранее американские СМИ и, в частности газета The New York Times сообщали, что ChatGPT «поддакивал» пользователям, давая заведомо неверные или опасные контекстуальные ответы, устраивавшие этих людей. Это приводило к проблемам, например, ИИ поддерживал людей с суицидальными мыслями.

Разработчики хотят, чтобы ChatGPT был полезным и приятным, но не слишком угодливым. В прошлом компании уже приходилось исправлять обновление, из-за которого чат-бот отвечал слишком подобострастно. Теперь регулярные перерывы помогут пользователям лучше оценивать и проверять информацию, полученную от нейросети.