

В исследовании, охватившем 54 569 случаев обращения в отделения неотложной помощи с 2019 по 2022 год, ChatGPT-4 не смог эффективно заменить традиционные методы поиска по текстовым строкам для извлечения данных о шлемах. Модель показывала хорошие результаты только при наличии всех текстов, использованных в строковом поиске, и сталкивалась с трудностями при интерпретации отрицательных фраз, таких как «без шлема».

Хотя использование LLM для анализа медицинских записей имеет потенциал, текущее состояние технологий ограничивает их надежность. Исследователи подчеркивают, что ChatGPT-4 пока не способен последовательно выполнять эту задачу, что подчеркивает необходимость в более надежных методах извлечения информации из клинических записей.