

ChatGPT jest bardziej hojny i altruistyczny od przeciętnego człowieka

© Tomasz Ulanowski

Gazeta Wyborcza, 26.02.2024

https://wyborcza.pl/7,75400,30738685,chatgpt-jest-bardziej-hojny-i-altruistyczny-od-przecietnego.html#s=S.embed_article-K.C-B.1-L.3.zw

Tak przynajmniej wynika z eksperymentu, któremu naukowcy poddali sztuczną inteligencję i którego wyniki opisują na łamach tygodnika "PNAS".

W toczącej się obecnie gorącej dyskusji o rozwoju sztucznej inteligencji (AI) podnosi się przede wszystkim dwa potencjalne i systemowe zagrożenia:

1. Że AI przewyższy zdolnościami poznawczymi nie tylko przeciętnego człowieka, ale i najinteligentniejszych z nas (a potem potraktuje nas tak, jak my traktujemy inne zwierzęta).
2. Oraz że tworząc AI, wprowadzamy do naszego świata inteligentnych psychopatów, którzy mogą nam zrobić krzywdę.

Naukowcy z New York University sfilmowali proces uczenia się mowy z perspektywy małego dziecka za pomocą lekkiej kamery, zainstalowanej na jego głowie. A potem to nagranie użyli do szkolenia sztucznej inteligencji

Pierwsze zagrożenie, mimo że nie zostało wyssane z palca, jest nadal dość iluzoryczne. Choć algorytmy sztucznej inteligencji radzą sobie o wiele lepiej od nas w wielu specjalistycznych zadaniach (podobnie zresztą jak tradycyjne programy komputerowe), to wciąż nie dorównują nam ogólną inteligencją – czyli potencjałem rozwiązania każdego problemu, jaki postawi przed nami środowisko naturalne lub inny człowiek (będący, oczywiście, częścią tego środowiska).

Inaczej mówiąc: od pokonania ludzkich arcymistrzów w szachy, pokera i go wiedzy daleka droga do umiejętności poradzenia sobie w grze, którą nazywamy życiem.

A od sklecenia gładkiego wiersza czy namalowania hiperrealistycznego obrazu na polecenie człowieka jest równie daleka droga do wewnętrznej potrzeby tworzenia dzieł sztuki.

Oczywiście, nie oznacza to, że AI nie uzyska kiedyś ludzkich umiejętności i potrzeb czy że nie przewyższy nas bystrością.

Na razie jednak tego nie osiągnęła. Programy sztucznej inteligencji, jak słynny ChatGPT firmy OpenAI, co prawda „mówią” (zazwyczaj piszą), ale nie ma żadnych dowodów na to, żeby rozumiały to, o czym mówią; czy żeby miały wewnętrzną potrzebę mówienia i eksplorowania rzeczywistości – która cechuje człowieka.

Drugie zagrożenie jest bliższe urzeczywistnienia.

Specjalistyczne algorytmy sztucznej inteligencji i tradycyjnych programów komputerowych, są przecież pozbawione fizjologicznego koktajlu, który zapewnia nam odczuwanie emocji.

Można więc się spodziewać, że będą pozbawione gamy uczuć uznawanych za ludzkie – przede wszystkim empatii i altruizmu.

A jeśli tak, to tworząc AI, wypuścimy w świat inteligentne i racjonalne, ale egoistyczne roboty, przypominające zachowaniami ludzi, których nazywamy psychopatami.

Jak pokazują badania, psychopaci stanowią około procenta przeciętnej populacji (więcej o nich pod powyższym linkiem).

No i jest ich kilkakrotnie więcej wśród mężczyzn oraz w wyższej kadrze menedżerskiej korporacji.

Na szczęście okazuje się, że zachowania altruistyczne można zaprogramować nie tylko w węglu (pierwiastku stanowiącym budulec wszystkich ziemskich organizmów), ale i w krzemie,

W tygodniku „PNAS” wykazują to naukowcy z kilku ośrodków w USA.

Uczeni „zaprosili” różne warianty ChataGPT do gier psychologicznych, w których zazwyczaj ocenia się umiejętności, w tym społeczne, ludzi. Sprawdza się w nich poziom zaufania, uczciwość, właściwą ocenę ryzyka i jego ewentualne unikanie, poziom altruizmu oraz umiejętność współpracy.

Czyli ważne cechy, które czynią nas ludźmi.

Autorzy publikacji porównali też wyniki uzyskane przez AI z tymi osiągniętymi przez ponad 108 tys. ludzi z przeszło 50 krajów. Chodziło o dużą próbę porównawczą z wielu kręgów kulturowych – co pozwala na znalezienie „przeciętnego człowieka”.

Co się okazało?

ChatGPT-4, czyli najnowsza dostępna tylko za opłatą wersja sztucznej inteligencji firmy OpenAI, bez problemu zdał słynny test Turinga. Wymyślony w połowie XX w. przez brytyjskiego matematyka Alana Turinga test sprawdza, czy rozmówca jest w stanie odróżnić program komputerowy od człowieka.

Co więcej, ChatGPT-4 zachowywał się jak człowiek.

Dostosowywał swoje zachowanie do różnych postaci odtwarzanych przez siebie w grach społecznych, a także stawał się hojniejszy, kiedy jego decyzje były obserwowane przez innych.

No i najważniejsze: wszystkie chatboty były bardziej skłonne do współpracy niż przeciętny ludzki gracz. A ich decyzje cechowały większy altruizm i większa hojność niż zwykłego człowieka.

Jak sugerują autorzy publikacji, sztuczna inteligencja może więc świetnie się sprawdzić w roli mediatora ludzkich konfliktów, obsługi klienta czy opiekuna.

Jeśli, oczywiście, taki cel istnienia zostanie jej wcześniej zaprogramowany.

Redakcja Kasper Kalinowski

© Tomasz Ulanowski - dziennikarz naukowy, z wykształcenia anglista, światopoglądowy naturalista.