

Sztuczna inteligencja już oszukuje i manipuluje ludźmi - piszą naukowcy

© Tomasz Ulanowski

Gazeta Wyborcza, 10.05.2024

<https://wyborcza.pl/7,75400,30957604,sztuczna-inteligencja-juz-oszukuje-i-manipuluje-ludźmi-pisza.html#s=S.embed article-K.C-B.1-L.1.zw>



Naukowcy z USA i Australii ostrzegają, że sztuczna inteligencja już nauczyła się oszukiwania i manipulowania ludźmi. (Fot. Colin Anderson Productions pty I/iStock)

Jeden z algorytmów udawał "martwego". Mistrzem oszustwa okazał się Cicero, program AI stworzony przez firmę Meta (tę od Facebooka).

Choć nie wiemy, jak funkcjonują algorytmy sztucznej inteligencji (z ang. zwane AI i funkcjonujące jak sieci neuronowe, o których działaniu wciąż w zasadzie nie mamy pojęcia), to widzimy, że stają się one coraz bardziej bystre – no i zżerają coraz więcej energii, ale to problem przyszłości.

Ostatni wyczyn?

Siły Powietrzne USA przetestowały AI w realnych warunkach starcia powietrznego dwóch myśliwców. W jednym F-16 siedział człowiek, drugi był pilotowany przez program komputerowy. Nie w wirtualnej rzeczywistości. Nad bazą Edwards w Kalifornii.

Dobra wiadomość? Nikt nie zginął.

Bezpieczeństwo związane z wprowadzaniem do naszego świata sztucznej inteligencji jest powodem coraz większego niepokoju wielu ekspertów.

Tym bardziej, że krzemowe mózgi już biją białkowych arcymistrzów – m.in. w go, szachy czy pokera. A badania prowadzone przez specjalistów od sztucznej inteligencji zmierzają ku narodzinom tzw. ogólnej AI, która zdolnościami poznawczymi co najmniej dorówna ludziom i będzie w stanie ogarnąć fizyczną rzeczywistość co najmniej tak szeroko, jak postrzega ją przeciętny człowiek.

Potem inteligencja oparta na komputerach ma przebić intelektem nawet najbystrzejszych z nas.

Czy będzie wtedy nadal podzielała nasze wartości i cele? Czy nie uzna, że jesteśmy zbędni albo jej zawadzamy? Albo po prostu przypadkiem nas nie zdepcze?

To coraz poważniejszy problem. A tych, którzy go lekceważą, sceptycy i pesymiści zachęcają do spojrzenia na to, co ludzie robią z innymi, mniej bystrzymi gatunkami. A także co bardziej zaawansowane technologicznie kultury robiły tym mniej zaawansowanym.

Już dzisiejsza AI nie tylko pamięta, przetwarza informacje (wg najprostszej definicji "myśli") oraz osiąga zadane jej cele (dysponuje więc inteligencją), ale także sama się uczy (to jej cecha immanentna). W końcu może więc osiągnąć świadomość (subiektywne doświadczenia), a nawet samoświadomość (zrozumie, że „ja” i „oni” to dwie różne sprawy).

Co wtedy?

W najnowszym wydaniu pisma „Patterns” naukowcy z USA i Australii ostrzegają, że sztuczna inteligencja już nauczyła się oszukiwania i manipulowania ludźmi. Autorzy publikacji odkryli te umiejętności u kilku algorytmów AI:

- m.in. w programie specjalizującym się w grze w pokera (to akurat nie dziwi),
- a także w strategicznym Starcraftcie.
- Mistrzem oszustwa okazał się stworzony przez firmę Meta (tę of Facebooka) program Cicero grający w Dyplomację.

-Meta wyszkoliła Cicera tak dobrze, że znalazł się on wśród 10 proc. najlepszych ludzkich graczy – mówi pierwszy autor publikacji dr Peter S. Park z MIT. – Nie dała jednak rady nauczyć go uczciwości.

Podobnie makiaweliczne umiejętności sztucznej inteligencji odkrywali zresztą już inni naukowcy. Np. jeden z algorytmów udawał „martwego”, kiedy jego wirtualną rzeczywistość przeczesywano w poszukiwaniu programów, które... zaczynały szybko się powielać.

Zdaniem dr. Parka choć dziś oszustwa dokonywane przez programy komputerowe w różnych grach mogą wydawać się błahe, to w przyszłości doskonalące się algorytmy AI mogą wręcz wpaść w spiralę oszustwa.

- Twórcy sztucznej inteligencji nie wiedzą, skąd te niepożądane zachowania – przyznaje naukowiec. – Sądzymy jednak, że algorytmy AI kierują się po prostu zdrowym rozsądkiem. Ponieważ w ich świecie oszustwo popłaca – pomaga im w osiągnięciu zadanych celów – to zaczęły je stosować.

Redagował Kasper Kalinowski

© Tomasz Ulanowski

dziennikarz naukowy, z wykształcenia anglista, światopoglądowy naturalista. Píše o antropocenie i ekocydzie (zmianach klimatu, środowisku, przyrodzie), o ewolucji człowieka, Ziemi i Kosmosie.