AI jako zagrożenie dla ludzkości? Przyspieszają prace nad regulowaniem sztucznej inteligencji

Gazeta Wyborcza, 13.04.2023

https://wyborcza.pl/7,75399,29658409,ai-jako-zagrozenie-dla-ludzkosci-przyspieszaja-prace-nad-regulowaniem.html#s=S.schowek_side-K.P-B.1-L.1.zw

UE, USA, a nawet Chiny - wyścigowi o prymat na rynku generatywnej AI towarzyszą coraz bardziej zdecydowane próby nałożenia na niego (i samą AI) kontroli.

Sam Altman, prezes OpenAI, odkąd jego firma stworzyła ChatGPT, musi odpowiadać na zarzuty o zbyt lekkomyślne podejście do sztucznej inteligencji (AI) i sprowokowanie wyścigu technologicznego o trudnych do przewidzenia skutkach.

Na te zarzuty odpowiada w dwojaki sposób. Po pierwsze: lepiej było udostępnić technologię opartą na dużych modelach językowych (LLM) światu teraz, gdy jest ona jeszcze mocno niedoskonała, i zyskać czas na otwartą dyskusję o niej. Łatwiej bowiem rozmawiać o czymś, co nie jest jeszcze inteligentniejsze od człowieka ani tym bardziej świadome. A tym być może kiedyś w przyszłości będzie zaawansowana AI.

Z tym wiąże się druga część riposty Altmana: trzeba jakoś sztuczną inteligencję uregulować, na co też swoimi działaniami chciał zwrócić uwagę rządów i społeczeństwa.

OpenAI dostał listę zadań

Pierwsze zareagowały Włochy - i nałożyły ban na ChatGPT za przetwarzanie danych osobowych bez stosownych uprawnień i zgód. Włosi zauważyli też, że aplikacja, która może generować zarówno nieprawdy, jak i, przy odpowiednio sformułowanych zapytaniach, treści rasistowskie, seksualne czy przemocowe, nie wprowadziła żadnego mechanizmu weryfikacji wieku swoich użytkowników.

W efekcie OpenAI zapowiedziała współpracę z włoskim odpowiednikiem polskiego UODO - Garante per la Protezione dei Dati Personali.

W środę włoski urząd postawił konkretne wymogi, które OpenAI musi spełnić do 30 kwietnia, aby ban na ChatGPT został zdjęty. Firma będzie obarczona obowiązkiem informowania o sposobie przetwarzania danych osobowych użytkowników czatbota, których będzie mogła używać wyłącznie za ich zgodą. Użytkownicy mają też dostać prawo wglądu do tych danych, ich zmiany i usunięcia. OpenAI ma także przeprowadzić kampanię informacyjną we włoskich mediach do 14 maja, w trakcie której ma wyjaśnić, jak używa danych osobowych użytkowników do trenowania swoich algorytmów. A do końca maja OpenAI musi przedstawić szczegółowy plan systemu weryfikacji wieku, który ma być wdrożony najpóźniej do końca sierpnia.

W Polsce, jak przekazał "Wyborczej" rzecznik UODO Adam Sanocki, "jak dotąd do UODO nie wpłynęły żadne skargi związane z przetwarzaniem danych osobowych w ramach korzystania z ChatGPT". "Ponadto pragnę poinformować, że nie była prowadzona kontrola w związku z jego funkcjonowaniem" - dodał rzecznik.

Lawina warunków ze strony Europy?

OpenAI nie ma siedziby w UE, więc może być celem regulatorów każdego z państw unijnych z osobna. ChatGPT zainteresowały się już Niemcy i Irlandia, ale także Francja i Hiszpania, które, jak donosi Politico, w trakcie czwartkowego spotkania Europejskiej Rady Ochrony Danych zrzeszającej regulatorów ochrony danych osobowych państw UE poruszą ten temat.

AI Act w obecnej formie już jest przestarzały

To rodzące się zainteresowanie regulatorów technologią opartą na LLM wywołało już zaniepokojenie największych graczy z Doliny Krzemowej.

W grudniu Rada UE przyjęła projekt regulacji dotyczącej sztucznej inteligencji - tzw. akt o sztucznej inteligencji (AI Act), i przekazała do dalszych prac.

Projekt jednak nie uwzględnia kwestii generatywnej sztucznej inteligencji, do której zaliczają się m.in. rozwiązania oparte na LLM. Organizacja Corporate Europe Observatory (CEO) zajmująca się monitorowaniem lobbingu w UE w swoim raporcie z lutego sugeruje, że to może nie być przypadek.

Komisja Europejska, która sygnuje AI Act, zaproponowała w dokumencie podział technologii AI ze względu na poziom przedstawianego przez nie zagrożenia.

Najwyższy, czyli "nieakceptowalny", a więc zakazany, ma objąć technologie takie, jak np. znany z Chin social scoring czy zabawki mogące głosem manipulować dziećmi.

Stopień niżej są technologie wysokiego ryzyka, czyli dopuszczalne, ale ściśle regulowane (np. autonomiczne pojazdy). Do tej kategorii miała też trafić AI "ogólnego przeznaczenia", czyli m.in. produkty oparte na LLM w rodzaju czatbotów, cyfrowych asystentów, translatorów, generatorów obrazu czy dźwięku, ale też platformy do trenowania modeli językowych.

Jak piszą autorzy raportu, big techy (m.in. Google i Microsoft) przy wsparciu rządu USA mocno lobbowały (i zapewne nadal lobbują) za tym, aby surowe regulacje wynikające z wysokiego ryzyka nie dotyczyły wytwórców ogólnego AI, lecz wyłącznie firm wdrażających dostarczone im rozwiązania. Microsoft jesienią 2022 wysłał nawet otwarty list (podpisały go też inne firmy) do czeskiej prezydencji w Radzie UE. "Nie ma potrzeby, aby akt o AI zawierał specjalną sekcję dotycząca AI ogólnego przeznaczenia" - napisano w nim.

W kwietniu w ramach rozmów trójstronnych pomiędzy KE, PE i RE mają się rozpocząć prace nad końcowym kształtem AI Act.

Za Europą mocarstwa sznurem

Kwestia regulacji AI coraz mocniej wybrzmiewa również w USA. W zeszłym tygodniu prezydent Biden rozmawiał ze swoimi doradcami o zagrożeniach związanych z

generatywnym AI. Został też zapytany przez dziennikarzy, czy ta technologia jest niebezpieczna. - Zobaczymy. Może być - odpowiedział Joe Biden.

We wtorek amerykańska administracja ogłosiła, że rozpoczyna zbieranie komentarzy odnośnie do wpływu AI na bezpieczeństwo narodowe i edukację.

A 11 kwietnia chiński organ nadzorujący internet (Chińska Administracja Cyberprzestrzeni – CAC) przedstawił własny projekt regulacji dotyczący rozwoju sztucznej inteligencji.