Sztuczna inteligencja w złych rękach stanowi poważne zagrożenie

19 lipca 2024

Wykorzystanie sztucznej inteligencji (w skrócie AI od Artificial Intelligence) do manipulowania ludźmi przyjmuje namacalną postać, jak deepfake, ale może się odbywać także dużo subtelniej, na przykład poprzez podsuwanie odpowiednich treści w wyszukiwarce. „Odwoływanie się w pewnych aspektach do kwestii etycznego stosowania tej technologii to za mało, jeśli w ślad za tym nie idą odpowiednie regulacje" – oceniają eksperci. Amnesty International (w skrócie także AI) przewiduje, że sztuczna inteligencja wykorzystywana bez kontroli i stabilnych ram prawnych może doprowadzić do nasilenia naruszeń praw człowieka, zwłaszcza w obliczu konfliktów zbrojnych i przez reżimy kontrolujące życie obywateli.

„Sztuczna inteligencja, która się rozwija od 70 lat, nagle stała się taka popularna. Rodzi to wyzwania etyczne. Przede wszystkim ona powinna pomagać ludziom, a nie szkodzić. Są takie przypadki, zwłaszcza kiedy dostaje się w ręce złych ludzi, że może nam zagrozić, ale jest bardzo wiele rzeczy dla nas pomocnych. Etyka powinna formułować takie zasady, żeby to było sprzyjające ludziom, a eliminować te rzeczy, które mogą być szkodliwe, eliminować zagrożenia, których trochę jest – mówi w wywiadzie dla agencji Newseria prof. dr hab. Krzysztof Jajuga, ekonomista z Uniwersytetu Ekonomicznego we Wrocławiu.

W najnowszym raporcie o stanie praw człowieka Amnesty International ostrzega, że istnieje korelacja między szybkim rozwojem sztucznej inteligencji a załamaniem rządów prawa, co przekłada się na ryzyko spotęgowania naruszeń praw człowieka. Autorzy raportu podkreślają, że niekontrolowane korzystanie z

nowych technologii, takich jak oprogramowanie szpiegujące i narzędzia masowej inwigilacji, narusza podstawowe prawa i wolności. Niektóre rządy wdrażają zautomatyzowane narzędzia wymierzone w najbardziej zmarginalizowane grupy społeczne, w szczególności migrantów i uchodźców, co może umacniać politykę dyskryminacyjną i pogłębiać nierówności społeczne.

„Nie wszyscy twórcy sztucznej inteligencji rozumieją to, co robią. Oni po prostu tworzą programy, a niewiele się orientują, po co to ma być. Niestety jeszcze gorzej jest z użytkownikami końcowymi, którzy dostają narzędzie szumnie nazwane sztuczną inteligencją. Prawie nikt tego nie weryfikuje, więc użytkownicy stosują coś, czego nie rozumieją, a co może im zaszkodzić. Etyka powinna trochę w tym pomóc, ale proszę pamiętać, że etyka jest to coś, co przyjmujemy jako pewien kodeks zachowania, to nie jest to regulacja prawna, czyli to nie jest obowiązek, tylko dobra praktyka" – podkreśla prof. Krzysztof Jajuga.

W ocenie autorów raportu Amnesty International samo odwoływanie się do kwestii etycznych nie wystarczy. Technologie takie jak generatywna sztuczna inteligencja, rozpoznawanie twarzy i oprogramowanie szpiegujące wymagają więc regulacji. One jednak w dużej mierze wciąż pozostają w tyle. Na polu europejskiego prawodawstwa odpowiedzią na ten postulat ma być przyjęta w maju regulacja „AI Act".

„Pojawiły się regulacje w tym zakresie, pokazujące, które systemy sztucznej inteligencji powinny być zakazane, bo one są zbyt ryzykowne i mogą powodować szkody, takie jak kradzież danych, ale nie tylko. Mogą powodować na przykład to, co widzimy w Chinach, tak zwany rating społeczny, czyli kategoryzowanie ludzi w zależności od pewnych parametrów, które wynikają z ich zachowań. Pewne rzeczy są nieetyczne i ci, którzy tworzą prawo, uznają, że trzeba to zrobić regulacją, czyli obowiązkową rzeczą, regulowaną, a nie tylko dobrą praktyką, która wynika z naszych zachowań" – dodaje ekspert Uniwersytetu Ekonomicznego we Wrocławiu.

Istotnym zagrożeniem związanym z rozwojem sztucznej inteligencji jest technologia deepfake, która umożliwia tworzenie zmanipulowanych komunikatów poprzez generowanie obrazu i dźwięku. W ten sposób można na przykład opracować film, na którym człowiek cieszący się dużym autorytetem społecznym prezentuje treści, które chce zaszczepić w społeczeństwie osoba posługująca się tą techniką manipulacji. „Twarzą" deepfake może być zarówno wpływowy polityk, jak i znany lekarz. Cyberprzestępcy posłużyli się w ten sposób niedawno wizerunkiem chirurga, prof. Krzysztofa Bieleckiego, by namawiać ludzi do zakupu pseudoleku o rzekomym działaniu kardiologicznym.

„Deepfake to już jest bardzo namacalna rzecz, to jest po prostu już powszechne. Od razu widać, że przekaz to deepfake. Nie po twarzy, nie po głosie, ale po tym, co ta osoba mówi. Ta osoba normalnie takich rzeczy by nie powiedziała i ktokolwiek rozumny nie patrzy na twarz, nie słucha głosu, tylko słucha, co ta osoba powiedziała, mówi: nie, ten pan czy ta pani tego nie mogły powiedzieć, to jest sztucznie zrobione. Chodzi mi też o bardziej skomplikowane rzeczy, które powodują, że podejmiemy jakieś decyzje finansowe czy medyczne" – wskazuje prof. Krzysztof Jajuga.

W tym kontekście warto wspomnieć o tzw. stronniczości algorytmów, które np. na poziomie wyszukiwarki mogą podsuwać odbiorcy nieobiektywnie dobrane treści. Fundacja Digital Poland zwraca uwagę na to, że jeżeli na przykład narzędzie do filtrowania podań o pracę jest trenowane na decyzjach podjętych przez ludzi, algorytm uczenia się maszyn może się nauczyć dyskryminacji kobiet lub osób o określonym pochodzeniu etnicznym.

„Etyka nie będzie stopować rozwoju sztucznej inteligencji jako takiej, natomiast w pewnych obszarach może przyjść zrozumienie i będą ludzie to stosowali rozważniej, natomiast na pewno sztuczna inteligencja jest bardzo potrzebna. Zresztą trzeba podkreślić, że nie ma czegoś takiego jak sztuczna

inteligencja. To jest ludzka inteligencja, która tworzy pewne bardzo przydatne rozwiązania. To się powinno rozwijać. Natomiast to, co jest szkodliwe, to kwestie etyczne, ale prędzej czy później regulacje powinny pewne rzeczy przystopować i to już się dzieje" – podkreśla ekspert. „Na poziomie Unii Europejskiej regulacje podzieliły wszystkie systemy na cztery poziomy. Jest poziom o najwyższym stopniu ryzyka, właśnie na przykład tzw. rating społeczny, który w UE byłby niedozwolony. Są też pewne rzeczy, które są obarczone bardzo dużym ryzykiem, ale jest też wiele systemów, które są o małym ryzyku i które są przydatne".

Jak wynika z przeprowadzonego wśród pracowników w Stanach Zjednoczonych badania EY „Human Risk in Cybersecurity", rozwój sztucznej inteligencji zwiększył skalę zagrożeń, na które pracownicy nie są wystarczająco przygotowani. 80 proc. respondentów wyraża zaniepokojenie możliwością wykorzystania tej technologii do przeprowadzania cyberataków. 39 proc. przyznaje, że nie wie, jak odpowiedzialnie korzystać z SI. Wysoki poziom niepewności i brak odpowiedniej wiedzy sprawiają, że co trzeci badany martwi się, że może nieświadomie narazić organizację na cyberataki.

Źródło: Newseria.pl