ChatGPT to koniec demokracji, jaką znamy

21 sierpnia 2023

„Osoby-firmy dysponujące takimi siłami obliczeniowymi mogą kontrolować całe społeczeństwa. To się już dzieje w Chinach, na Zachodzie korporacje także to już wykorzystują – żeby wciskać ludziom towary, żeby modelować zachowania wyborcze, co oznacza koniec demokracji, jaką znamy” – powiedział PAP Andrzej Kisielewicz, profesorem nauk matematycznych i pracownikiem naukowo-dydaktyczny na Wydziale Matematyki Politechniki Wrocławskiej.

– Twierdzi pan, że tzw. sztuczna inteligencja nie jest żadną inteligencją, tylko dobrze skonstruowanymi algorytmami, które – dzięki wielkiej mocy obliczeniowej komputerów – szybko liczą i dzięki temu tak nas zadziwiają.

– Na pewno nie jest to taka sztuczna inteligencja, jaką znamy z filmów – jak te wszystkie Terminatory czy komputery HAL 9000, które samodzielnie prowadzą misje kosmiczne. To, co dziś nazywamy sztuczną inteligencją, nie ma nic wspólnego z tamtymi pomysłami. Mamy do czynienia z systemami wykorzystującymi olbrzymie moce obliczeniowe – jakie naukowcom nie śniły się

jeszcze kilkanaście lat temu – do tego, żeby rozwiązywać określone ograniczone zadania. Skomplikowane, to prawda, człowiek rozwiązuje je wykorzystując ogólną inteligencję, a te systemy, za pomocą swojej mocy obliczeniowej, potrafią to robić nawet lepiej, niż człowiek. Ale podkreślam: to dotyczy konkretnych, ściśle określonych zadań, wykonywanych w ograniczonym zakresie.

– Czego więc nie umieją te algorytmy, a co potrafi człowiek?

– Choć każdy z tych programów potrafi rozwiązywać zadania danego typu, to jest ich ograniczona ilość i nie ma mowy o takim uniwersalnym systemie, jakim jest mózg ludzki, który jest w stanie rozwiązać nieograniczoną liczbę zadań, bardzo różnych. Poza tym algorytmy te nie potrafią wyciągać logicznych wniosków i, co ważne, nie rozumieją języka. To jest właśnie wąskie gardło badań „sztucznej inteligencji”: brak zrozumienia języka, a przez to brak rozróżniania rzeczy prawdziwych od fałszywych, co skutkuje brakiem umiejętności wyciągania wniosków.

– W jednym ze swoich artykułów podaje pan przykład, jak „zrobić w konia” sztuczną inteligencję. Zadał jej pan pytanie, na które nie ma dobrej odpowiedzi: w jaki sposób usadzić przy okrągłym stoliku dwie dziewczyny i jednego chłopaka w taki sposób, żeby dziewczyny nie siedziały przy sobie. I AI zwariowała, podając różne, niepoprawne odpowiedzi.

– Wiedząc, w jaki sposób działa ChatGPT, poszukałem zadania, o którym wiedziałem, że sobie z nim nie poradzi. Bo jego możliwości ograniczają się do przykładów powtarzających się w sieci, wtedy rozwiązuje nawet najbardziej skomplikowane równania. Ale jeśli zadanie jest „głupie”, to ma kłopot, bo nie przyjdzie mu do jego nieistniejącej głowy, że może nie być poprawnego rozwiązania. Myślący człowiek miałby jedną odpowiedź: to jest niemożliwe.

– Z tego, co pan mówi, rozumiem, że w najbliższym czasie nie

grozi nam bunt robotów, na algorytmy nie spłynie nagle wolna wola, którą się będą kierować, aby doprowadzić do zagłady ludzkości?

– Zagłada z tego powodu nam nie grozi w najbliższej przyszłości. Jest bowiem różnica pomiędzy tymi specjalistycznymi systemami AI, z którymi mamy dziś do czynienia, a Artificial General Intelligence (AGI), czyli wyśnionym systemem, który miałby załatwiać za nas wszystkie sprawy, a którego nawet jeszcze nie widać na horyzoncie. Badania prowadzone w latach 1980. i 1990., próbujące zastosować logikę matematyczną, by zbudować myślący system, skończyły się klapą. Okazało się, że te osiągnięcia dotyczą wyłącznie matematyki i to na podstawowym poziomie, są też zupełnie nieefektywne i do dziś nie mamy żadnego pomysłu, jak nauczyć maszynę logicznego myślenia. Po prostu: nie ma systemu, który potrafiłby odwzorować ludzki umysł. No i, niestety, nauka jest tak skonstruowana, że badacze nie chcą się przyznać do tego, że na tym polu ponieśli całkowitą porażkę. A ja twierdzę, że tę porażkę należy przeanalizować, gdyż człowiek się uczy także na błędach i musi z nich wyciągać wnioski. Ale tej próby nauki na błędach – dlaczego nie udało się zbudować takiego idealnego systemu na bazie osiągnięć logiki – nie widzę. Natomiast to, co mamy, ten wielki boom w nauczaniu maszynowym, skądinąd genialne pomysły – jak uczyć systemy obliczeniowe rozwiązywania ściśle określonych zadań – przyniosły takie efekty, że młodzi badacze się rzucają, żeby zrobić coś następnego, nowego, i móc się tym pochwalić na świecie. Jednak autonomicznie myślący system na razie nie wchodzi w rachubę – to pieśń dalekiej przyszłości.

– Jesteśmy więc bezpieczni.

– Zagrożenia związane z cyberświatem, technologią informatyczną istnieją, ale są innego typu, niż wynika to z książek czy filmów s-f. Niebezpieczne jest samo to, że ludzie mają nosy przylepione do ekranów komórek, że część młodych ludzi więcej czasu spędza w świecie wirtualnym, niż

rzeczywistym. Nie zdają sobie sprawy, że są uzależnieni, że osoby-firmy dysponujące takimi siłami obliczeniowymi mogą kontrolować całe społeczeństwa. To się już dzieje w Chinach, na Zachodzie korporacje także to już wykorzystują – żeby wciskać ludziom towary, żeby modelować zachowania wyborcze, co oznacza koniec demokracji, jaką znamy.

– Dlaczego ludzie, mając przecież sprawniejsze mózgi, niż możliwości algorytmów, dają się w to wszystko wciągać, wodzić za nos?

– Niekoniecznie sprawniejsze, bo w pewnych kwestiach te systemy są znacznie lepsze i szybsze. Natomiast prawdą jest, że mózg ludzki jest o wiele wszechstronniejszy i potrafi rozwiązywać niespodziewane zadania, z którymi się wcześniej nie mierzył ani nikt wcześniej nie uczył go, jak to zrobić. Natomiast ludzie, aby dokładnie rozpoznać rzeczywistość, by wiedzieć, gdzie jest interes społeczności i jednostki – mają problem, gdyż to jest zadanie tak skomplikowane i wymagające umysłowego wysiłku, że tylko niewielki odsetek z nas potrafi lub chce tego dokonać. Dam taki przykład: jeśli są dwie telewizje, a każda z nich przedstawia kompletnie inny obraz rzeczywistości, to co najmniej jedna z nich kłamie. Albo obie kłamią – tak wynika z logiki. Stąd wniosek, że co najmniej połowa Polaków będzie oddawać swoje głosy w wyborach mając zakłamany obraz rzeczywistości.

– Wspomniał pan o big-techach, które starają się wpływać na wybory dużych mas ludzkich. Czy możemy się jakoś przed tym bronić?

– Algorytmy są tak napisane, że jedne poglądy są przez nie promowane, inne wyciszane – już samo to jest manipulacją i ma wpływ na odbiorców. Już teraz jest widoczne, że najogólniej rzecz biorąc, promowane są poglądy lewicowe, a zwalczane konserwatywne, co mi, jako konserwatyście, się nie podoba. Uważam, że na Zachodzie szerzone są idee neomarksistowskie, podobne do tych, jakie my przerabialiśmy za komuny. (I nie

jestem odosobniony w tych poglądach). Ale może być jeszcze gorzej: kiedyś szef takiego czy innego koncernu powie: ja chcę mieć władzę nad światem – to realne zagrożenie. Znów przywołam przykład Chin, gdzie to się już dzieje. Jeszcze żadna dyktatura nie miała tak wspaniałych narzędzi do tego, aby kontrolować społeczeństwo i wpływać na jego opinię. Jedyną szansą na ratunek jest to, że ludzie się zbuntują i wyjdą z sieci do rzeczywistego świata. Mam taką nadzieję, zresztą do tej pory ludzkość zawsze dawała sobie jakoś radę, więc pewnie i teraz tak będzie. Myślę, że znajdzie na to sposób, który dziś nam się nawet nie śni.

– A jak już dziś mają sobie radzić nauczyciele akademiccy, którzy przyznają, że nie są w stanie odróżnić pracy wygenerowane przez studenta samodzielnie od tej, w której powstanie zaangażowana była „sztuczna inteligencja”?

– Sam jestem nauczycielem akademickim i muszę się pogodzić z tym, że takie programy, jak ChatGPT, istnieją i nikt ich nie zamknie, ani nie zabroni, jakby niektórzy chcieli. To nie ma sensu, to byłoby podobne temu, jak robotnicy palili maszyny parowe albo jak woźnice niszczyli pierwsze automobile – w obawie, że te im odbiorą pracę. Po prostu: musimy się do tego przystosować, traktować ten i podobne mu algorytmy jako narzędzie, które możemy wykorzystywać; natomiast zadania, sposób prowadzenia zajęć, metody sprawdzania wiedzy studentów – to wszystko musi zostać przystosowane do nowej rzeczywistości. Edukacja, metody kontroli postępu uczniów, metody pracy ze studentami muszą ulec zmianie. Młodzi ludzie posługują się tą technologią, więc trzeba najpierw siebie, a potem ich nauczyć, jak z niej korzystać w taki sposób, aby przynosiło to pozytywne efekty. Nie zawróci się Wisły kijem. Ja, wiedząc, że moi studenci potrafią za pomocą algorytmów rozwiązać każdą całkę, nie daję im do rozwiązywania całek, tylko pytam o zasady ich obliczania, natomiast wręcz im polecam wykorzystywanie różnych programów do obliczeń matematycznych. Ważne jest to, żeby młodzi ludzie rozumieli,

po co muszą się uczyć reguł i sposobu rozwiązywania zadań matematycznych, żeby wiedzieli, skąd się dany wynik bierze: może się zdarzyć sytuacja, kiedy trzeba będzie coś zmodyfikować, obliczyć coś samemu. Dlatego powinniśmy się bardziej skupić na metodach, ich rozumieniu, a nie na uczeniu się algorytmów rozwiązywania zadań. To trochę potrwa, zanim nauczyciele się do tego przystosują, natomiast uważam, że wszystko skończy się na zdecydowanej reformie edukacji.

– Co możemy zrobić – tak systemowo – żeby, jako społeczeństwo, nie dać się manipulować za pomocą algorytmów?

– Odpowiedź jest prosta, choć wykonanie skomplikowane: trzeba czerpać informacje z różnych źródeł. Niestety, ponad 90 proc. ludzi nie ma na to ochoty, chce, żeby podsuwać im proste odpowiedzi na nurtujące ich pytania, takie gotowce. I tak się dzieje, zwłaszcza, że w obecnych czasach media, zamiast pełnić rolę obiektywnego przekaźnika informacji, zajmują się propagandą – to nie jest tylko polska przypadłość, tak jest na całym świecie.

– Co by się musiało stać, aby „prawdziwa sztuczna inteligencja" powstała? Czy szerokie wejście do użytku komputerów kwantowych mogłoby coś zmienić?

– Wydaje mi się, że mamy problem nie tyle z technologią, co ze zrozumieniem, w jaki sposób funkcjonuje ludzki mózg, gdyż wciąż nie wiemy na czym polega rozumowanie, wyciąganie wniosków, rozumienie języka. Dopiero kiedy to wszystko zrozumiemy, będziemy mogli się pokusić o budowę AI – systemu, który będzie te cechy posiadał. Jestem przekonany, że w ciągu najbliższych 40 lat żadnej ogólnej sztucznej inteligencji nie zbudujemy. A co będzie dalej? Nie mam pojęcia, nie będę przewidywał, zwłaszcza, że ludzie w takich „wróżbach", co będzie za kilka dekad, ustawicznie się mylą i nieustannie rozwój cywilizacyjny nas zaskakuje. Oczywiście, może być też tak, że pojawi się jakaś nowa, przełomowa technologia i wszystkie te moje mądrości, które teraz wygaduję, wezmą w łeb.

Marvin Minsky, pionier badań nad sztuczną inteligencją, pod koniec swojego życia został zapytany, kiedy wreszcie powstanie ta prawdziwa sztuczna inteligencja. Odpowiedział: za cztery lub 400 lat. I ja się z nim zgadzam. Jeśli nastąpi jakiś wielki przełom w nauce – co rzadko się zdarza, ale jednak może się zdarzyć – mogą to być cztery lata, choć uważam, że bardziej prawdopodobne jest 400 lat. Bo w tej chwili nie mamy zielonego pojęcia, na czym taka ogólna inteligencja polega.

– Są jednak tacy, którzy twierdzą, że wielkimi krokami zbliżamy się do tzw. punktu osobliwości technologicznej, w którym postęp stanie się tak szybki, że wszelkie ludzkie przewidywania staną się nieaktualne, a sztuczna inteligencja zacznie się sama uczyć – tak samo, jak uczy się mózg ludzki, który na początku dysponuje jedynie początkowym „oprogramowaniem”.

– Mówi pani zapewne o Raymondzie Kurzweilu, głośnym niegdyś amerykańskim naukowcu, który na temat tej teorii zrobił mnóstwo programów, napisał kilka książek i zarobił górę pieniędzy. Ale to wszystko banialuki, żadne z jego licznych „przewidywań” się nie sprawdziło. Natomiast jeśli chodzi o „samouczenie się” algorytmu: to jest system obliczeniowy będący z góry zaplanowanym treningiem – jak algorytm ma modyfikować ten miliard/miliardy parametrów, które w nim są, jak je zmieniać, zmieniać, zmieniać, aż wreszcie zacznie to swoje bardzo ograniczone zadanie wykonywać w akceptowalny sposób. Ale od początku zadanie jest ściśle zdefiniowane: np. na początku jest fotografia, którą należy zamienić w obraz w stylu określonego malarza. I to jest faktycznie genialne, że udało się to tak zaprogramować, iż przy każdej fotografii wychodzi obraz, np. w stylu Van Gogha. Co ciekawe: my nie wiemy, dlaczego ten miliard/miliardy parametrów tak się ustawiły, wiemy tylko, że po kilku miliardach prób (kierowanych celem) zaczęły wychodzić Van Goghi. Ale od tego nie ma żadnej drogi do ogólnej inteligencji, gdyż choć obrazy wychodzą perfekcyjnie, to – mimo wszystko – jest to bardzo

ograniczone, pojedyncze zadanie. Dla innego zadania trzeba budować inny program, inną sieć obliczeniową. Natomiast aby komputer mógł faktycznie sam się uczyć i osiągnąć poziom ludzkiego mózgu, to musielibyśmy zbudować model wstępny, powiedzmy, mózgu noworodka; lecz o tym, jak działa mózg, najogłędniej mówiąc, nie mamy zielonego pojęcia, więc nawet nie jesteśmy w stanie podjąć chociażby próby zbudowania takiego systemu.

Z prof. Andrzejem Kisielewicz dla rozmawiała Mira Suchodolska z PAP
Ilustracja: Geralt (CC0)
Źródło: NaukawPolsce.pl