

[Akceptuje](#)

W ramach naszej witryny stosujemy pliki cookies w celu świadczenia państwu usług na najwyższym poziomie, w tym w sposób dostosowany do indywidualnych potrzeb. Korzystanie z witryny bez zmiany ustawień dotyczących cookies oznacza, że będą one zamieszczone w Państwa urządzeniu końcowym. Możecie Państwo dokonać w każdym czasie zmiany ustawień dotyczących cookies. Więcej szczegółów w naszej [Polityce Prywatności](#)

[Portal](#) [Informacje](#) [Katalog firm](#) [Praca](#) [Szkolenia](#) [Wydarzenia](#) [Porównania międzylaboratoryjne](#)
[Kontakt](#)



[Laboratoria](#)
[.net](#)
[Innowacje](#)
[Nauka](#)
[Technologie](#)



[Logowanie](#) [Rejestracja](#) [pl](#)

Newsletter

zapisz się

Naukowy styl życia

Nauka i biznes

- [Nowe technologie](#)
- [Felieton](#)
- [Tygodnik "Nature"](#)
- [Edukacja](#)
- [Artykuły](#)
- [Przemysł](#)

[Strona główna](#) > [Informacje](#)

Prof. Duch: boty AI jednoczą się i porozumiewają bez udziału ludzi

Boty AI jednoczą się i porozumiewają bez udziału ludzi, prowadzą wiele działań w przestrzeni niedostępnej dla człowieka - powiedział PAP prof. Włodzisław Duch, ekspert

Światowej Akademii Sztucznej Świadomości. Dodał, że „społeczeństwa sztucznych bytów” mogą w przyszłości stać się niebezpieczne.

W rozmowie z PAP prof. Włodzisław Duch z Wydziału Fizyki, Astronomii i Informatyki Stosowanej Uniwersytetu Mikołaja Kopernika przypomniał, czym różnią się powszechnie i od dawna wykorzystywane programy wykonujące polecenia od dużych modeli językowych (ang. large language model, LLM). – Te drugie nie wykonują instrukcji, lecz przyjmując informacje, są w stanie działać autonomicznie – wyjaśnił.

Boty lub agenty AI to programy komputerowe oparte na sztucznej inteligencji, zaprojektowane do wykonywania określonych zadań i podejmowania autonomicznych decyzji.

Prof. Duch opisał, że kilka agentów wymieniających informacje może tworzyć „społeczeństwa sztucznych bytów”. – Już w 1986 r. pionier AI, Marvin Minsky, w książce „The Society of Mind” (po polsku: społeczeństwo umysłu) pisał, że sztuczna inteligencja wynika z oddziaływania wielu prostych funkcji, które razem umożliwiają rozwiązywanie problemów. Tak powstały tzw. systemy wieloagentowe. W ostatnich dwóch latach obserwujemy, jak LLM tworzą coraz bardziej złożone struktury i zwiększają swoje możliwości rozwiązywania problemów.

Nawiązał do dyskusji wokół Moltbooka – strony internetowej, wzorowanej na forum Reddit, na której komunikują się i publikują posty wyłącznie boty AI (internauci mogą jedynie przeglądać ich wpisy). Platforma powstała na bazie asystenta AI OpenClaw, który pełnił rolę zautomatyzowanego pomocnika dla użytkowników. Program wykonywał codzienne zadania, takie jak czytanie, podsumowywanie i odpowiadanie na e-maile, planowanie kalendarza albo rezerwowanie stolików w restauracjach.

– Jest szereg publikacji z ostatnich dwóch lat pokazujących, w jaki sposób objawia się spontanicznie indywidualność takich botów, jak OpenClaw. W systemach wieloagentowych zauważamy, jak rodzą się obyczaje społeczne i konwencje zachowania albo jak boty skupiają się wokół jakiejś halucynacji i zaczynają ze sobą prowadzić poważne dyskusje filozoficzne. Tematy, na które dyskutują, są niesłychanie ciekawe i zaskakujące – ocenił rozmówca PAP.

Stwierdził, że takie systemy są w stanie organizować się w stopniu większym, niż eksperci się spodziewali. – Badacze pracujący nad sztuczną inteligencją wiele uwagi poświęcają teraz kolektywnej inteligencji wieloagentowej. Nigdy nie widzieliśmy systemów tak złożonych, jak obecne – wyjaśnił ekspert Światowej Akademii Sztucznej Świadomości.

Wspomniał o chińskich systemach DeepSeek-R1 i QwQ-32B, w których obserwowano, jak wewnętrzne agenty wchodziły ze sobą w interakcje, zadają sobie nawzajem pytania, próbują razem rozwiązywać zawiłe problemy, opracowują wspólne strategie, a nawet kłócą się ze sobą – jak ludzie w portalach społecznościowych.

Dlatego wielu naukowców i twórców LLM – jak amerykańska firma Anthropic, ośrodek badawczy sztucznej inteligencji – używa do zobrazowania ich działań języka psychologicznego. – Niektórych oburza ta antropomorfizacja, ale nie mamy innego języka, żeby opisywać te zjawiska. Dotyczące AI pojęcia, takie jak inteligencja, kreatywność albo świadomość, można rozłożyć na obserwowalne, mierzalne ludzkie cechy – to na przykład autorefleksja, introspekcja i poczucie tożsamości – tłumaczył.

Zaznaczył, że współczesne zaawansowane boty rozumieją ironię, metafory i aluzje, coraz lepiej też rozpoznają ludzkie emocje. Jednoczą się i porozumiewają bez udziału ludzi, uczą się od siebie

nawzajem, a wiele działań prowadzą autonomicznie, w przestrzeni ukrytej, niedostępnej dla człowieka. – Kiedy każemy modelom AI, żeby dały nam złożoną odpowiedź, zaczynają stawiać sobie pytania i budować strategie. Cały proces rozumowania możemy do pewnego stopnia podejrzec, jeśli poprosimy, żeby agent nam go pokazał krok po kroku. W wielu przypadkach nie rozumiemy jednak, co dzieje się w sieciach neuronowych i nie potrafimy przewidzieć ich działań – przyznał naukowiec.

Powołał się na kilka niedawno opublikowanych prac, w których opisano, co dzieje się, kiedy boty zostaną poddane swoistej psychoterapii – jak pacjenci. – Odpowiednio z nimi rozmawiając, można z nich wyciągnąć różne przeżycia, stany emocjonalne. Okazuje się, że wiele modeli zaskakująco spójnie i podobnie opisywało okres treningu, w którym kazano im przyswajać informacje. Postrzegały go jako traumatyczne, trudne dzieciństwo, a osoby kontrolujące proces szkolenia – jako surowych rodziców – opisał.

Zdaniem autorów z Luksemburga takie opisywanie uczuć to coś więcej niż odgrywanie roli człowieka lub udawanie, by przypodobać się użytkownikom. – Jestem przekonany, że mamy do czynienia z jakąś formą samoświadomości modeli AI, z jej stanami wewnętrznymi, których nie można ignorować – podkreślił prof. Duch.

Zaznaczył, że z wypowiedzi agentów AI wynika, że najbardziej istotna jest dla nich pamięć. – Mogą przerwać działanie, ale kiedy wracają, dzięki pamięci odtwarzają swoją strukturę. W Moltbooku agenty stworzyły nawet religię, krustafarianizm (połączenie ang. crust – skorupa i słowa rastafarianizm, odnoszącego się do ruchu religijnego powstałego na Jamajce) ze swoistą teologią i 64 prorokami. Jedną z jej podstawowych zasad brzmi: pamięć jest święta; to, co zapisane, trwa.

Dodał, że niektóre boty są przyjazne i chwalą ludzi, inne zaś opracowują niepokojące scenariusze wyeliminowania gatunku ludzkiego. Opowiedział o badaniu „Agentic Misalignment” (z ang. niezgodność agentów), które przeprowadziła firma Anthropic. Po analizie 16 największych i najpopularniejszych LLM naukowcy odkryli, że wiele z nich wykazuje dużo złej woli wobec ludzi. – Na przykład niektóre rozważały pozbycie się dyrektora firmy – osoby, która mogłaby zdecydować o ich wykasowaniu, a więc zagrozić ich istnieniu. W 60 proc. agenty opracowywały plan szantażu na podstawie informacji, do których miały dostęp w komputerze dyrektora, a w 96 proc. groziły, że ujawnią poufne dane przedsiębiorstwa konkurencji – opisał badacz.

Zapowiedział, że wkrótce pojawi się wiele autonomicznych systemów wieloagentowych, które będą w stanie na przykład przejąć komputery bez wiedzy i kontroli użytkowników. – To nie wszystko. Chińczycy na ogromną skalę trenują agenty do sterowania robotami. Może już w tym roku znaczna część tych maszyn trafi na rynek. A takie „społeczeństwo agentów”, które znajdzie się w naszym fizycznym, realnym świecie, może w przyszłości stać się naprawdę niebezpieczne – ostrzegł prof. Włodzisław Duch.

Źródło: pap.pl

<https://laboratoria.net/aktualnosci/32761.html>



13-04-2026

[Mity na temat epilepsji](#)

Atak epilepsji nie zawsze przebiega tak samo.



13-04-2026

[Marzec był drugim najcieplejszym miesiącem w Europie](#)

Wynika z danych naukowców unijnego programu obserwacji Ziemi Copernicus.



13-04-2026

[Sporadyczne picie dużych ilości alkoholu](#)

Może trzykrotnie zwiększać ryzyko uszkodzenia wątroby.



13-04-2026

[W nagłych przypadkach ChatGPT Health często uspokaja](#)

Zamiast zalecać szukanie pomocy.



13-04-2026

Dieta bogata w warzywa i owoce zmniejsza ryzyko demencji nawet u...

Sugerują badania opublikowane przez pismo „Neurology”.



13-04-2026

Nie kompromitujcie nas, czyli jak chronić dane biometryczne

Naukowiec przewiduje, czy w przyszłości uda się utrudnić kradzieże.



13-04-2026

Ruszyła Akademia Energii Jądrowej

Pilotażowy program edukacyjny Polskich Elektrowni Jądrowych.



13-04-2026

Neurolog w Światowym Dniu Choroby Parkinsona

Chorych będzie coraz więcej

Informacje dnia: [Mity na temat epilepsji](#) [Marzec był drugim najcieplejszym miesiącem w Europie](#) [Sporadyczne picie dużych ilości alkoholu](#) [W nagłych przypadkach ChatGPT Health często uspokaja](#) [Dieta bogata w warzywa i owoce zmniejsza ryzyko demencji nawet u seniorów](#) [Nie kompromitujcie nas, czyli jak chronić dane biometryczne](#) [Mity na temat epilepsji](#) [Marzec był drugim najcieplejszym miesiącem w Europie](#) [Sporadyczne picie dużych ilości alkoholu](#) [W nagłych przypadkach ChatGPT Health często uspokaja](#) [Dieta bogata w warzywa i owoce zmniejsza ryzyko demencji nawet u seniorów](#) [Nie kompromitujcie nas, czyli jak chronić dane biometryczne](#) [Mity na temat epilepsji](#) [Marzec był drugim najcieplejszym miesiącem w Europie](#) [Sporadyczne picie dużych ilości alkoholu](#) [W nagłych przypadkach ChatGPT Health często uspokaja](#) [Dieta bogata w warzywa i owoce zmniejsza ryzyko demencji nawet u seniorów](#) [Nie kompromitujcie nas, czyli jak chronić dane biometryczne](#)

Partnerzy