

# ZROZUMIEĆ CZŁOWIEKA

O tym, dlaczego obawiamy się humanoidalnych robotów, a także o tym, czy oglądanie maszyn grających w piłkę jest przyjemne, opowiada **dr Aleksandra Przegalińska**.

## ACADEMIA: Czym właściwie jest sztuczna inteligencja?

ALEKSANDRA PRZEGALINSKA: To jest dobre pytanie, bo to cały czas jest kwestia, która podlega dyskusji. Ostatnio pojawił się raport Dominiki Maison, która zapytała wielu ekspertów, co to jest SI, i wszystkie odpowiedzi zaczynały się od „trudno to stwierdzić”. Więc rzeczywiście coś jest na rzeczy. Mogę więc powiedzieć, czym jest SI dla mnie. Uważam, że to jest dyscyplina wiedzy, dziedzina badawcza i dyscyplina wdrożeń, która rozwija się dynamicznie (choć bywały momenty stagnacji) od lat 40. ubiegłego stulecia. Stanowi interdyscyplinarny parasol, pod którym jest co najmniej kilka innych subdziedzin. Ogólnie SI mogłaby być sklasyfikowana jako dziedzina, w której staramy się tworzyć sztuczne systemy, będące w stanie przejawiać aktywność zbliżoną człowiekowi, czyli będą wykazywać się inteligencją, adaptować się do zmieniających się warunków i skutecznie działać w zmieniającym się świecie.

Niektórzy mówią, że w SI chodzi o to, żeby symulować aktywności mentalne człowieka. Według mnie nie o to chodzi. Symulacja zakłada, że udajemy, że coś jest, a tak naprawdę tego nie ma. Na samym początku SI, jej pierwszy budowniczy, niejaki John McCarthy mówił, że o SI możemy mówić wtedy, gdy oceniamy jakąś aktywność, którą ocenilibyśmy jako inteligencję, gdyby wykonywał ją człowiek. Jeśli jakiś system wykona zadanie równie dobrze jak człowiek, to możemy mówić o SI. To podejście behawioralne. Myślę, że dzisiaj raczej nie chodzi nam o to, żeby udawać, że coś działa jak człowiek, tylko żeby faktycznie pewne funkcje poznawcze, mentalne człowieka czy w ogóle inteligentnych żywych systemów spróbować odtworzyć w sztucznych środowiskach tak, żeby one też mogły działać, adaptować się, rozwiązywać problemy.

To są pewne cele i założenia tej dziedziny. A pod jej parasolem wyliczyłabym najważniejsze działy. Po pierwsze, uczenie maszynowe, czyli zdolność przetwarzania ogromnych wolumenów danych. Po drugie,

widzenie maszynowe, czyli dziedzina, która polega na tym, że przetwarzany jest obraz, a system jest w stanie dokonać identyfikacji tego, co się wokół niego objawia, czyli widzi przedmioty, obiekty, ludzi. I po trzecie, przetwarzanie języka naturalnego. Tutaj chodzi o dane tekstowe, które mogą być przetwarzane różnymi metodami, ale główny pomysł jest taki, żeby dokonywać syntezy ludzkiej mowy tak, żeby ją rozumieć i odtwarzać. Jednym słowem, celem jest to, żeby te systemy działały dobrze nie tylko w językach formalnych, ale i naturalnych, jak polski, angielski czy niemiecki.

Poza tymi głównymi poddziedzinami z SI wiązana jest też robotyka, zwłaszcza sensorowa, gdzie pojawia się wspomniana wcześniej adaptacja systemu do otoczenia. Boston Dynamics [czołowa amerykańska firma, specjalizująca się w inżynierii i robotyce – przyp. red.], tworzy systemy potrafiące odnieść się do zmieniającego się otoczenia: pokonają przeszkodę, są w stanie stabilnie stać na różnych przestrzeniach i to jest część robotyki, która przynależy do sztucznej inteligencji.

## Ciekawe, że mówi pani bardziej o zdolności do adaptacji do otoczenia, a niekoniecznie o symulowaniu ludzkich cech.

Symulowanie też ma sens. Na przykład algorytmy genetyczne symulują działanie jakiejś populacji. Taka symulacja jest w tym przypadku sposobem uwidocznienia czegoś, co w świecie fizycznym byłoby trudne. Realne stworzenie sieci neuronowych, które by sprawnie przetwarzały dane, byłoby bardzo trudne. Ale symulacja komputerowa sieci neuronowej jest możliwa. Tylko to nie jest symulacja rozumiana jako udawanie, oszukiwanie, tylko zakładająca zrobienie czegoś w przestrzeni wirtualnej w momencie, gdy nie można tego zrobić w przestrzeni fizycznej. To jest co innego. Jak mówię o symulacji w negatywnym sensie, to chodzi mi o prosty system, który coś udaje, ale robi to za pomocą prostackich trików. W robotyce to się często zdarza i dla mnie wtedy to nie jest sztuczna inteligencja, tylko udawactwo.

DR ALEKSANDRA PRZEGALINSKA



**Zatem nieprostackie udawanie człowieka polegałoby na tym, że SI jest w stanie nie tylko analizować wielkie liczby danych, ale na przykład podejmować autonomiczne decyzje?**

Dobrym przykładem na wytłumaczenie tej różnicy jest właśnie mowa. Nie chodzi o to, żeby system naśladował głos, ale go generował, uczył się mówić. Chodzi o system, który na poziomie semantycznym rozumie, co mówimy, czyli jest w stanie zmapować słowa. Nie jest po prostu boksem, który przetwarza frazy, bo ma takie zadanie, żeby z każdego zdania oznajmującego zrobić pytanie. Te pierwsze boty dawnego typu nie wiedziały, co to jest treść w języku. Po prostu przetwarzały frazy i robiły to w sposób przekonujący. Cokolwiek się mówiło, system o to pytał, działał jak psychoterapeuta. I faktycznie, jak powie się: „mam problem”, to system może odpowiedzieć na to: „masz problem?”. Terapeuta też mógłby tak zapytać, ale to nie jest żadne rozumienie mowy ludzkiej, wypowiedzi, ich treści, tylko są to proste algorytmiczne operacje. Jest też sporo trików, w których można udawać, że system przetwarza język i że rozumie treści, syntetyzuje ludzką mowę, a w rzeczywistości wcale tego nie robi. To jest taka symulacja, która kiedyś była uprawniona, ale dzisiaj nie jest już większą wartością,

bo możemy to robić w sposób bardziej zaangażowany, dojrzały, głębszy.

**A czy jest możliwe spowodowanie, że SI będzie empatyczna podczas rozmowy?**

Że będzie odczuwała – nie. Ale można jej dać sens empatii. Ciekawe, że pani o to pyta, bo DeepMind, czyli taka spółka zależna od Google (ta sama, która stworzyła AlphaGo Zero, który pokonał człowieka w grze) zapowiedziała, że stworzą coś, co się nazywa ToMnet – Theory of Mind Network. To ma być sieć, która będzie oddawała teorię sieci innego umysłu. Chodzi w niej o to, że wchodząc ze sobą w rozmaite interakcje, jesteśmy w stanie zrozumieć, że drugi człowiek, wysyłający różne sygnały, przechodzi przez jakieś stany emocjonalne, że ma jakieś intencje stojące za jego wypowiedziami itd. I to jest w filozofii umysłu bardzo znana teoria. Tylko że zawsze myślałam, iż to jest koncepcja czysto filozoficzna, nie sądziłam, że ktoś postanowi ją na serio zaimplementować w praktycznym sensie. DeepMind zapowiedziało, że chce zbudować sieć, która będzie budowała teorię innego pomysłu. I teraz, odpowiadając na pytanie: ona nie będzie niczego odczuwała, nie będzie oczywiście niczego świadoma. To będzie sieć neuronowa, która ma prze-

**dr Aleksandra Przegalińska**

Prowadzi badania z obszarów filozofii i sztucznej inteligencji. Jest związana z Akademią Leona Koźmińskiego, a od 2020 r. zaczyna pracę w Harvard Law School Labour and Worklife Program. Jest współautorką pierwszego w Polsce programu studiów dotyczących sztucznej inteligencji w dziedzinie zarządzania.

aprzegalinska@kozminski.edu.pl



tworząc pozawerbalne komunikaty i w związku z tym mieć podwyższony poziom kontekstowości wypowiedzi rozmówcy. Czyli rozumieć, że mogę powiedzieć coś w sposób zdenerwowany albo spokojny, że mogę być zniecierpliwiona albo że mogę czegoś oczekiwać, wypowiadając dane słowa. Nie wiem, jak to dokładnie będzie realizowane, bo DeepMind nie chce za dużo zdradzać. Intencja jest taka, żeby możliwie zbliżyć się do ludzkiego stylu komunikacji, który zakłada, że jeśli wchodzi w grę rozmowa, to jest coś jeszcze poza samą warstwą komunikatu werbalnego i znaczeń słów.

**A co jest Świętym Graalem dla osób pracujących w obszarze SI? Co jest tą granicą, marzeniem badawczym?**

Jest nim oczywiście świadomość u maszyn, dlatego że jej nie rozumiemy. Nawet u siebie. Ale nie wiem, czy chcę tego Świętego Graala, szczerze mówiąc. Nie możemy okiełznać własnego umysłu, zbadać, jak to jest, że myśli się formułują, jak działa mózg. Te me-

sama w sobie jest bardzo szlachetna. To jest próba zrozumienia mechanizmów biologicznego świata, odtworzenia ich w jakiś sposób, próba wejścia w samych siebie. Jej celem jest zrozumienie najważniejszej warstwy człowieka – mentalnej. Co nie znaczy, że chciałabym, żeby SI była świadoma. Na szczęście nie sądzę, żeby było to możliwe w najbliższym czasie.

**No właśnie, to by było przerażające. Czyli celem robotyków nie jest zastąpienie człowieka?**

Absolutnie nie. Celem robotyka jest częściowo zrozumieć funkcje psychomotoryczne człowieka i dobrze je odtwarzać. Na przykład bionika zajmuje się tworzeniem sztucznych kończyn, które będą współpracowały z ludzkim ciałem. Ich celem w ogóle nie jest zastępowanie człowieka, tylko lepsze go zrozumienie. A jeśli już zastępowanie, to w tych funkcjach, których on nie chce wypełniać. Dlaczego ludzie w XXI w. mają nosić ciężkie paczki i łamać sobie kręgosłupy, skoro ewidentnie nie są do tego przystosowani? Lepiej stworzyć robota Atlas z Boston Dynamics, który będzie za nich tę paczkę dźwigał i prznosił ją tam, gdzie chcą. Czyli ma służyć do wykonywania pewnych czynności, które dla człowieka są trudne. Zdolność logicznego myślenia w przypadku człowieka jest znacznie większa. Maszyny mają swoje domeny, a my swoje. Inteligencja, polegająca na operowaniu językiem formalnym, przetwarzaniu ogromnej ilości danych, to jest inteligencja sztuczna – ona jest w tym po prostu najlepsza. Tak samo w przypadku czynności nudnych i ciężkich, np. wypełniania za każdym razem tych samych formularzy. To też SI może robić zamiast człowieka. Dlatego teraz głównym trendem nie jest próba zastąpienia człowieka, tylko komfortyzacja, czyli pokazywanie, gdzie systemy SI mogą wesprzeć inteligencję naturalną.

**W jednym z ostatnich numerów „Economist” pojawił się artykuł o tym, że teraz roboty, a nie lekarze, będą przyjmować pacjentów w przychodni. Nagłówek jest kontrowersyjny, ale w treści okazuje się, że chodzi o zastąpienie ludzi w czynnościach, w których łatwo o pomyłkę, o typowo ludzki błąd.**

Ta automatyzacja jest trochę jak bajka o żelaznym wilku. Wielu ludzi tym straszy, ale ten system będzie znacznie powolniejszy, niż się tym wyrrywającym do przodu analitykom wydaje. Wczoraj byłam na konferencji w Berlinie, podczas której roboty grały w piłkę. To było naprawdę żalosne widowisko. Pewne czynności, jak tapetowanie pokoju, dla systemów SI są trudne. Inne są łatwe, jak np. przetwarzanie ogromnych ilości danych finansowych czy danych w diagnostyce medycznej. SI może niewątpliwie wspomóc człowieka w wykryciu rzadkich chorób, których człowiek nie umie jeszcze rozpoznawać. Tutaj algorytmy mają wielką moc. Ale mówienie, że rozmaite prace będą całkowi-

” Dla neuronauki i dla filozofii umysłu celem SI jest zrozumieć świadomość, a nawet ją odtworzyć, i to jest Święty Graal.

tody, które dziś mamy, neuroobrazowania samego mózgu są bardzo skąpe. Natomiast umysł jest właściwością fizycznego mózgu, który przetwarza myśli, u każdego działa to trochę inaczej. Ta dziedzina kryje mnóstwo zagadek. Sama świadomość jest konceptem niejasnym. Zakłada, że różne istoty żywe mają jakiś jej poziom, z czego ludzie mają na pewno dosyć wysoki. Mają też samoświadomość, zwrotność, mogą odnieść się do siebie, do siebie w czasie, mają integralność tożsamości w czasie, która na pewno jakoś współgra ze świadomością. To wszystko są zagadnienia, które są zamarkowane, ale nie są jasne. Neuronauka nie odpowiada na nie w tej chwili. Można się zastanowić, jaki jest największy cel robotyki i SI. Jakby zapytać biznesmena, toby chciał zaprząć SI do procesów, żeby były maksymalnie efektywne. Ale gdyby zapytać naukowca, toby odpowiedział, że sztuczna inteligencja jest po to, żeby dać nam wgląd w rozumowanie, myślenie człowieka. Dla obu dyscyplin, i dla neuronauki, i dla filozofii umysłu, celem SI jest zrozumieć świadomość, a nawet ją odtworzyć, i to jest Święty Graal.

Uważam, że SI to jest przepiękna, wspaniała dziedzina. Oczywiście potrafi być źle wykorzystywana, ale

DR ALEKSANDRA PRZEGALINSKA

cie zastąpione przez SI, jest przesadą. Zwłaszcza że na jedną pracę czy konkretny zawód składa się mnóstwo czynności. Jedna praca wymaga wielu różnych umiejętności, często zakłada kontakt z ludźmi, komunikację. To w większości nie są sytuacje, w których jeden człowiek przekłada ryby na taśmie. Takie prace się jeszcze zdarzają, moim zdaniem zostaną zastąpione, ale większość zawodów wymaga wielu poziomów różnego typu inteligencji: społecznej, obliczeniowej. I w tym sensie ludzie, którzy wieszczą zastąpienie człowieka przez maszynę, nie zdają sobie sprawy z tego, o czym mówią.

### A jakie są realne zagrożenia związane z SI?

Szeroko rozumiane *misusage* [nadużycia – przyp. red.], czyli spekulacje na rynku finansowym, hakerstwo, *malicious bots*, ruch botowy, czyli sytuacje, w których manipulujemy wyborami, wrzucamy do sieci armię robotów, która rozprzestrzenia jakieś fałszywe informacje itd. Zagrożeń jest szereg, ale one nie są z obszaru automatyzacji i nie wynikają z tego, że SI zyskuje świadomość i zniszczy człowieka. Te zagrożenia widzę gdzie indziej. W przestrzeniach, w których SI zagraża człowiekowi, bo jest wykorzystywana przez innego człowieka do celów niezgodnych z przeznaczeniem. Widzę problemy związane np. z bronią autonomiczną. Ale taka wizja, że SI to coś, co przemaszerowuje przez świat i zostawia człowieka bezbronny i bezrobotny, to duża przesada.

### Pani zajmuje się też relacjami człowieka z technologiami. Z czego wynika ten strach przed tym, że maszyny przejmą władzę nad światem?

Wynika on co najmniej z kilku rzeczy. Moim zdaniem chociażby z tego, że ludzie boją się systemów, które ich przypominają. Konstrukcja systemu SI, która wygląda jak robot, jest znacznie mniej przerażająca niż taka, która wygląda jak człowiek – proszę pomyśleć o robocie Sofii. Jeśli jest system, który nie jest człowiekiem, a demonstruje jakieś ludzkie cechy, na ogół w sposób nie do końca udany, to takie systemy antropomorficzne budzą spory lęk u ludzi, bo ludzie nie wiedzą, jak to sklasyfikować. Nie wiedzą, co to właściwie jest, a muszą się z tym konfrontować. Zadają sobie pytanie, czy to żyje, czy nie żyje. I to wprowadza ich w stan wysokiej konfuzji. Więc takie systemy nie są dzisiaj pożądane. Oczywiście jest masa innych problemów. Mnóstwo ludzi w Stanach Zjednoczonych zakupiła Alexę, to są roboty, którymi można zarządzać właściwie całym domem, wyłączać światło, robić zakupy. Działają na podstawie komunikatów głosowych, dlatego ludzie się martwią, czy Alexa ich przypadkiem nie podsłuchuje. Zachodni świat jest wyczulony na kwestie prywatności, dlatego taki robot może się nie do końca podobać. Kolejna sprawa to lęki związane z tym, czy system uwzględni dane, które ludzie zazwyczaj uwzględniają. Na przykład, jak mamy system SI, który oblicza, czy jestem w porządku pod kątem

wzięcia kredytu, zastanawiam się, czy weźmie pod uwagę jedne dane, a innych nie i wyda na mnie wyrok, któremu podporządkują się pracownicy banku i nikt go nie będzie nawet kwestionował. Może też zabraknąć takiego komponentu, że ktoś we mnie nie uwierzył. Sztuczna inteligencja tego nie robi, ona nie odczuwa zazdrości, ufności, nadziei. To jest w pełni racjonalny system i tego też wiele osób się obawia. Że zostawi ich bez szans odwołania się, że inni będą na nim polegać jak na ekspertach, a pojedynczy człowiek będzie stratny.

### Jeszcze ostatnie pytanie związane z zapleczem akademickim. Pani pracowała na MIT, najlepszej technicznej uczelni na świecie. Czy może pani zdradzić, jak tam wygląda praca od kuchni? Jakie są jakościowe różnice pracy badawczej w porównaniu z Polską?

Pierwsza, ale kluczowa sprawa to silna i bardzo zdrowa współpraca z biznesem. SI jest obszarem wdrożeń biznesowych, ale MIT bardzo dużo pracuje w różnych dziedzinach. To nie jest praca na zamówienie biznesu, polega raczej na odkrywaniu fundamentalnych zagadnień, które mogą się potem przydać. Istnieje więc weryfikacja pomysłów pod kątem praktyczności, ale jednocześnie panuje totalna swoboda pod kątem działalności intelektualnej. To nie jest korporacja naukowa, która działa na zamówienie biznesu. Umowy z firmami są tak skonstruowane, że chronią niezależność uczelni, a w tym samym czasie ta współpraca ma sens. Biznes daje dane i finansowanie. Takie podejście mi imponuje i uważam, że jest bardzo zdrowe. Z jednej strony praca tam nie jest sobie a muzom, ale ma ważny komponent fundamentalnego odkrywania. Poza tym oddanie pracy. Zespoły składają się z ludzi, którzy są wysoko zmotywowani, nie są zużyci, niechętni, nie wykonują tej pracy z przymusu. Odwrotnie, to często są ludzie, którzy zrezygnowali z pracy korporacyjnej, gdzie mogliby dostawać lepsze pieniądze, ale są zafascynowani tym, co robią. Duch oddania rozprzestrzenia się na cały zespół. Praca zespołowa jest traktowana bardzo poważnie, jest dużo współpracy, debriefingu, elementów dzielenia się, stawiania sobie nawzajem wyzwania i zagłębiania się także w czyjąś pracę. Staramy się to powtarzać w Akademii Leona Koźmińskiego, bo takie działanie jest naprawdę bardzo skuteczne. Wszyscy pracują w duchu, żeby komuś lepiej poszło. Oczywiście uderza także dostęp do aparatury itp., ale to jest drugorzędne. Granie na zespół, do wspólnej bramki i patrzenie dalekosiężne to główne cechy pracy w MIT. Dużo uwagi jest skierowanej na wspólny wysiłek, a także optymizm naukowy, wiary w to, że się uda. A do tego nieustannie zadawane jest pytanie: czy społeczeństwo na tym skorzysta?

Z DR ALEKSANDRĄ PRZEGALINSKĄ  
ROZMAWIAŁA DR JUSTYNA ORŁOWSKA