

STANISŁAW KRAJEWSKI

Uniwersytet Warszawski

<https://orcid.org/0000-0002-1142-8112>

Nie jestem przygotowany na... AI

I Am Unprepared for... AI

Abstract

Everyone is talking about Chat GPT. And correctly so. The bulk of skills will become automated. On the other hand, there will remain the need for actual authors. There arises the question who is to judge whether something is truly creative and innovative. Moreover, Artificial Intelligence (AI) indicates that an increasingly large part of our humanity is digitally simulated. Are there any entrenchments that will defend themselves? If so then we must rise above the level of that which is subjective. This is what the philosophy of dialogue is about. It can be also used for formulating a furthest going reinforcement of the Turing test. Is an analysis of challenges posed by AI a chance for academic philosophy?

Keywords: AI, Turing Test, intelligence, creativity, philosophy of dialogue, robots

Abstrakt

Wszyscy mówią o Chat GPT. I słusznie. Masa umiejętności zostanie zautomatyzowana. Pozostanie natomiast potrzeba rzeczywistych twórców. Powstaje pytanie, kto ma oceniać, czy coś jest naprawdę twórcze i nowatorskie. Ponadto sztuczna inteligencja (AI) wskazuje, że coraz większa część naszego człowieczeństwa jest symulowana cyfrowo. Czy zatem są szanse, które się obronią? Jeśli są, to musimy wykroczyć poza poziom tego, co subiektywne. O tym traktuje filozofia dialogu. Da się ją też wykorzystać do sformułowania najdalej idącego wzmocnienia testu Turinga. Czy analiza wyzwań stawianych przez AI jest szansą dla akademickiej filozofii?

Słowa kluczowe: AI, Test Turinga, inteligencja, twórczość, filozofia dialogu, roboty

O autorze

Stanisław Krajewski – profesor na Wydziale Filozofii Uniwersytetu Warszawskiego, współtwórca i współprzewodniczący Polskiej Rady Chrześcijan i Żydów. Współtworzył wystawę *Powojnie* w Muzeum Historii Żydów Polskich POLIN. Autor książek i artykułów z zakresu logiki, filozofii matematyki, judaizmu, historii i doświadczeń Żydów, a także dialogu międzyreligijnego, między innymi: *Twierdzenie Gödla i jego interpretacje filozoficzne*; *Czy matematyka jest nauką humanistyczną?*; *Żydzi, judaizm, Polska*; *54 komentarze do Tory dla nawet najmniej religijnych spośród nas*; *Poland and the Jews. Reflections of a Polish Polish Jew*; *Tajemnica Izraela a tajemnica Kościoła*; *Nasza żydowskość*; *Żydzi i...*; *Co zawdzięczam dialogowi międzyreligijnemu i chrześcijaństwu*; *Żydzi w Polsce – i w Tatrach też*.

Nie jestem przygotowany na... AI

Naprawdę nie jestem przygotowany. Ani ja, ani nikt inny. Nawet naukowcy z instytutów informatyki ani nawet ci mądrze z Doliny Krzemowej. Sztuczna inteligencja przyszła niespodziewanie w 2022 roku i zmieniła nasz świat.

Mówią i piszą o tym wszyscy. Tłumaczenia automatyczne – spotykamy je na każdym kroku, tworzenie obrazów i filmów na podstawie jednozdaniowej sugestii będzie niedługo powszechne, podobnie jak sztuczni doradcy – boty, z którymi wciąż mamy do czynienia, chcemy czy nie. Autonomiczne samochody są już testowane, ale chyba jeszcze trochę na nie poczekamy, za to autonomiczna broń, w tym drony samodzielnie wyszukujące cel, ponoć już jest stosowana, ciarki przechodzą. Coś się zmieniło, szybko nastaje i nie wiadomo dokąd nas prowadzi. To nowe to sztuczna inteligencja, AI. Ponieważ wszyscy o niej mówią, bardzo trudno powiedzieć coś nowego, niebanalnego. Zarazem mam poczucie, że żaden temat nie jest równie ważny dla świata i dla humanistyki. Trzeba go zgłębiać. Wszyscy, którzy fascynują się tą nową sytuacją, podkreślają, że AI zmieniła świat i zmieni go jeszcze głębiej, daleko bardziej radykalnie.

Ale zaraz, czy naprawdę? Ktoś mógłby oponować. Co prawda inaczej załatwiamy sprawy w banku, inaczej oglądamy sprawozdania z egzotycznych krain, bo możemy tłumaczyć na żywo wypowiedzi w obcych językach, co prawda jest faktem, iż uczeń, urzędnik, policjant, żołnierz, polityk, dyrektor i kucharz używają podobnych telefonów, ale żyjemy w zasadzie tak samo jak niegdyś. To, co naprawdę ważne, się nie zmienia: wybuchają wojny, chorujemy, umieramy, zakochujemy się, nasze marzenia się nie zmieniły i nie są bynajmniej bliższe realizacji. Ktoś mógłby więc zapytać: Co niby, poza sprawnością wykonywania pewnych operacji, uległo zmianie?

Być może nie istnieje jedna odpowiedź, która zadowoliliby wszystkich.

Oto przykład z mojego życia. Prowadzę zajęcia dla studentów, kończą się one egzaminem. Od dawna był to egzamin pisemny. Zawsze pozwalałem na korzystanie z własnych notatek i książek, a od paru lat dopuszczam używanie telefonów i komputerów. Wychodzę z założenia, że w przyszłym życiu każdy będzie mógł wspomagać swoją wiedzę encyklopedią zawartą w smartfonie, a zatem podobnie powinno być także na egzaminie. Podkreślam, że chodzi tu o wiedzę, nie o umiejętności. To rozróżnienie wyjaśni przykład z dziedziny działalności górskiej: nazwę szczytu można sobie przypomnieć, używając telefonu, być może da się w internecie znaleźć przebieg drogi wejściowej, nic to jednak nie pomoże w uzyskaniu wydolności niezbędnej do osiągnięcia tego szczytu. Projektując egzaminy, na których można używać wszelkich pomocy, odchodziłem od paradygmatu wiedzy pamięciowej. Byłem z tego bardzo dumny. Jakie to nowoczesne, liberalne i życzliwe dla studentów! Zresztą ich prace egzaminacyjne nie były szczególnie lepsze niż przedtem, co samo w sobie jest ciekawym rezultatem tego eksperymentu.

I wszystko to mogłoby trwać bezboleśnie, ku zadowoleniu egzaminatora, gdyby nie AI. Jak wiadomo, dzisiejsze programy są w stanie napisać na dowolny temat rozprawkę dostatecznie dobrą, by zdać egzamin. A przynajmniej żeby go zaliczyć wedle niewygórowanych kryteriów, które ja przyjmuję. Sam student może wówczas nic nie wiedzieć i nic nie rozumieć. Dlatego w tym roku musiałem oznajmić, że egzamin będzie ustny. Krótka rozmowa umożliwiła mi zorientowanie się, jaki jest poziom rozumienia tematów przez uczestników zajęć. Było jak w dawniejszej epoce, natychmiast pojawiały się złożone relacje międzyludzkie. Nawet jeśli wiele ich aspektów było przewidywalnych, nie wchodził w grę żaden automatyzm.

Powyższy drobny przykład wskazuje takie kwestie, jak wspomniana już konieczność odróżnienia wiedzy od umiejętności, zakres możliwej automatyzacji, a ponadto wartość relacji między osobnymi podmiotami, a w związku z tym wymóg dopracowania definicji podmiotu.

Praca i twórczość

Już wiemy, iż konieczne jest staranne odróżnienie wiedzy od umiejętności. Każdy ma encyklopedię w kieszeni, należy więc uczyć tego, jak z niej korzystać. To oznacza konieczność zmiany logiki nauczania szkolnego¹.

Elon Musk powiedział niedawno, że praca przestanie być potrzebna, bo dużą jej część będą wykonywać roboty. To niewątpliwie przesada, ale pewne zmiany są nieuniknione. Gdy się zastanowić, staje się jasne, że tylko niektóre zajęcia ulegną automatyzacji. Kierowca, sprzedawca, a także tłumacz, dziennikarz, scenarzysta, a nawet programista czy lekarz diagnosta – może tak, te zawody są zagrożone, ale już profesje hydraulika, pielęgniarka czy fryzjera nie są. Czyli rewolucja polegająca na wprowadzeniu programów zdolnych wykonać wiele prac, które dotąd wymagały inteligencji, uderza przede wszystkim w klasę średnią.

Nie wiemy, co będzie, pozwolę sobie jednak sformułować pewną przepowiednię. Otóż możliwości AI doprowadzą do automatyzacji nie tyle wszystkiego, co przeciętnych, powtarzalnych czynności w działaniach twórczych. Programy korzystają z tego, co już jest, przede wszystkim

z tego, co zostało udostępnione w internecie, i to modyfikują. Oczywiście podobnie czynimy my, ludzie. Programy robią to jednak szybciej, operują na nieporównywalnie większej liczbie danych, a niedługo będą to robiły zgrabniej niż ludzie. Oceniają zdolność kredytową klienta mniej więcej tak jak konsultant w banku, ale zrobią to w ułamku sekundy. Wykonają każdą standardową czynność, która jest wynikiem stosowania procedur obserwacji otoczenia. Standardową, dostatecznie typową. Można powiedzieć, że AI doprowadzi do automatyzacji przeciętności.

Pozostanie natomiast potrzeba rzeczywistych twórców, czyli ludzi, którzy wytyczają nowe drogi, a nie powielają czy kompilują materiały dostępne w sieci i jedynie po raz kolejny uruchamiają już zaprogramowane procedury. To równie dobrze – a może i lepiej – zrobi AI. Otwieranie nowych horyzontów wymaga czegoś więcej. Twórczość nowatorska pozostanie potrzebna, choć zarazem wydaje mi się jasne, że twórcom będzie trudniej. Spora część naszej działalności twórczej jest poniekąd standardowa – polega na kompilowaniu, modyfikowaniu i wygładzaniu tego, co już mamy. Już za chwilę sztuczna inteligencja będzie umiała napisać nowy dramat w stylu Szekspira – może zresztą już to potrafi. Coś naprawdę nowatorskiego nie jest tak łatwo osiągalne, choć i w tej kwestii istnieją optymiści, czy może raczej pesymiści... Ogólna sztuczna inteligencja (AGI), która stopniowo powstaje, ma nie ustępować nam na żadnym polu.

Powstaje pytanie, kto ma oceniać, czy coś jest naprawdę twórcze i nowatorskie. Przecież nie AI. A zatem ludzie. Ale jacy? Kto? Ten problem jest tylko jednym z wielu w całym wachlarzu stworzonych przez komputery i internet potrzeb, które staną się – już się stały! – dojmująco istotne. Mianowicie potrzebujemy ekspertów, którzy będą oceniać nie tylko produkty AI, ale też treści z Wikipedii i innych źródeł, do których wszyscy sięgają. Wiemy, że jest tam masa informacji, ale także wiele dezinformacji. Potrzebni będą recenzenci zdolni odsiać jedno od drugiego. Owi eksperci muszą być naprawdę niezależni, a to trudne zadanie. Co więcej, nawet jeśli uda się to osiągnąć i powstaną zespoły świetnych recenzentów, pojawi się problem jeszcze trudniejszy. Bo trzeba będzie przekonać opinię publiczną o tym, że są oni naprawdę autonomiczni i godni zaufania.

Kto może osiągnąć taki status? Być może nikt. To byłoby niewymownie smutne. A może jednak istnieją ludzie dostatecznie przenikliwi, krytyczni i wykształceni, umiejący współpracować z profesjonalistami w wszelkich dziedzin nauki i praktyki społecznej oraz fachowcami od internetu. Oto wyzwanie dla intelektualistów zatroskanych o stan cywilizacji.

Subiektywność

Inny ważny wniosek z anegdoty o egzaminach wynika z faktu, że duże modele językowe, takie jak Chat GPT, są w stanie tworzyć przekonujące dla nas teksty, ale wedle większości ekspertów nie rozumieją, o czym piszą. Pro-

gram potrafi błyskawicznie dobrać kolejne słowo, korzystając z dostępnych już tysięcy czy milionów tekstów. Wynik może robić na nas wrażenie znajomości rzeczy. Oto przykład uzyskany dzięki bezpłatnej prostej wersji czatu GPT. Poproszony o napisanie czegoś na temat zasygnalizowany w tytule niniejszego artykułu, program wypłul następujący tekst:

Nie jestem gotowy na sztuczną inteligencję

W cichej przestrzeni naszych myśli, pomysł zaakceptowania sztucznej inteligencji może wzbudzić uczucie wahania. „Nie jestem gotowy na sztuczną inteligencję”, zastanawiamy się, gdy granice między ludzkim poznaniem a uczeniem maszynowym zaczynają się zacierać. Tempo, w jakim rozwija się technologia, może sprawić, że poczujemy się nieprzygotowani na głębokie zmiany, jakie niesie ze sobą.

Jednakże, w tym momencie refleksji, pozwólmy sobie zastanowić się, że gotowość nie jest stałym celem, ale podróżą adaptacji. Nadejście sztucznej inteligencji zachęca nas do eksploracji głębokości naszego własnego zrozumienia, do kwestionowania istoty tego, co znaczy być człowiekiem.

Podczas gdy przemierzamy tę niezbadaną przestrzeń, naturalne jest odczuwanie obaw co do konsekwencji sztucznej inteligencji dla naszego życia, prywatności i struktury społeczeństwa. To uznanie złożoności i odpowiedzialności towarzyszących postępowi.

Jednak wśród niepewności tkwi szansa na wzrost, współpracę i etyczną opiekę. Być może, uznając, że możemy nie czuć się w pełni przygotowani, możemy podejść do integracji sztucznej inteligencji z ciekawością, uważnością i zobowiązaniem do kształtowania jej trajektorii w zgodzie z naszymi wartościami.

Podróż w erę sztucznej inteligencji nie jest podróżą samotną. To wspólne przedsięwzięcie, które wymaga otwartej rozmowy, rozważań etycznych i wspólnej wizji przyszłości, w której ludzka i sztuczna inteligencja współistnieją harmonijnie.

Poproszony o bardziej szczegółowy komentarz do zmian wprowadzonych przez ogólną sztuczną inteligencję (AGI), czyli program imitujący zachowania kojarzone dotychczas z ludzką inteligencją, chat napisał:

Zmiany cywilizacyjne spowodowane przez sztuczną inteligencję, a zwłaszcza AGI

W erze nadchodzącej Sztucznej Ogólnej Inteligencji (AGI) stajemy przed głębokimi zmianami cywilizacyjnymi. Pojawienie się AGI zapowiada doniosłą zmianę w dynamice ludzkiego istnienia, przekraczając granice tego, co uznawaliśmy za możliwe.

W miarę jak AGI, ze swoją zdolnością do autonomicznego uczenia się i adaptacji, wplata się w społeczeństwo, kontury naszego codziennego życia są skazane na ewolucję. Charakter pracy, gospodarki, edukacji i relacji



Marta Leśniakowska, z cyklu *Muzeum*, 2010–2024.

międzyludzkich ulegnie metamorfozie, kwestionując konwencjonalne normy i wymuszając ponowne ocenienie struktur społecznych.

Wpływ na siłę roboczą jest niezaprzeczalny. Zdolność do wykonywania skomplikowanych zadań z niesamowitą szybkością i precyzją AGI może rewolucjonizować branże, prowadząc do ponownego określenia pracy i umiejętności. Ten przełom wymaga skoncentrowanych działań w zakresie przeobrażania edukacji i rozwoju umiejętności, tworząc społeczeństwo zdolne do współistnienia z inteligentnymi maszynami.

Paradygmaty ekonomiczne przejdą trzęsienie ziemi w miarę jak automatyzacja i AGI redefiniują produktywność. Wyzwania związane z równym rozdziałem bogactwa i potencjalnym zastąpieniem miejsc pracy

wymagają starannej analizy, aby zapewnić sprawiedliwe i inkludujące społeczeństwo.

Rozważania etyczne mają ogromne znaczenie w tej przełomowej perspektywie. Kwestie prywatności, odpowiedzialności i etycznego użytkowania technologii AI stają się kluczowymi filarami budowy zrównoważonej przyszłości. Znalezienie delikatnej równowagi między innowacją a odpowiedzialnością będzie kluczowe w nawigowaniu po niezbadanych wodach AGI.

Dynamika międzyludzka również ulegnie przemodelowaniu. Przecięcie ludzkiej i sztucznej inteligencji pobudza do refleksji nad naturą empatii, kreatywności i samej istoty ludzkiej. Kulturowanie harmonijnego współistnienia między ludzkością a AGI stanie się wyzwaniem definiującym tę erę.

Przyjmując potencjalne zmiany cywilizacyjne wywołane przez AGI, jesteśmy wezwani do rozważań, współpracy i kształtowania przyszłości zgodnej z naszymi wspólnymi wartościami. Przed nami podróż pełna ekscytacji i wyzwań, wymagająca zbiorowej mądrości, by wykorzystać potencjał AGI dla dobra ludzkości i zachowania podstawowych zasad, które definiują naszą cywilizację.

Te rozprawki są dość ogólnikowe, ale wydają się sensowne i na temat. Wystarczają jako wypracowania szkolne, a nawet uniwersyteckie, jeśli przyjmując niewyógrowane kryteria. Skoro można wątpić, czy stoi za nimi rozumienie, a takie wątpliwości nie tylko są zgłaszane, ale wręcz dominują wśród ekspertów, to tym bardziej należy się zastanowić, jak my osiągamy rozumienie. Chodzi mi nie o naukową definicję, lecz o sposób, w jaki w praktyce rozstrzygamy, czy ktoś rozumie, co mówi. Z reguły dzieje się to w rozmowie. Egzamin ustny jest dobrym przykładem. W trakcie konwersacji wyczuwamy, co ktoś wie i rozumie, a co bezmyślnie powtarza. Ogólniej mówiąc, chodzi o ważność bezpośredniej relacji z drugim człowiekiem. Relacji prawdziwej, a nie udawanej.

Relacja, o której mowa, zachodzi między podmiotami. Każdy z nich ma swoją własną subiektywność, a relacja między nimi tworzy nową jakość. My, ludzie, jesteśmy istotami społecznymi – to ważny motyw filozofii tradycyjnej i współczesnej. Co więcej, wedle intuicji wprowadzonych przez filozofię dialogu (zwłaszcza Martina Bubera) nasze „ja” jest abstrakcją od bardziej pierwotnej rzeczywistości, w szczególności sytuacji „ja – ty”. Takie podejście odsyła do trudno uchwytnej sfery ponadsubiektywnej. Relacja stwarza nową jakość w porównaniu z jej członami, czyli ludźmi, a ogólniej mówiąc: subiektywnymi podmiotami.

Tymczasem robot pozostaje w sferze subiektywności. Rzecz w tym, że inne podmioty nie są mu potrzebne. Nic nie wnoszą. Jak się wydaje, kilka robotów można zawsze zastąpić jednym zintegrowanym, którego pierwotne roboty są częściami. Jeśli tak jest, pojęcie relacji – takiej jak relacje międzyludzkie – nie odnosi się do robotów. Ich inteligencja pozostanie zatem inna, sztuczna, czyli nieprawdziwa. Ludzka specyfika pozostaje niezagrożona. Przynajmniej do momentu, w którym zostanie zaprogramowana czy zasymulowana sfera ponadsubiektywna, w której żyją relacje.

Nadnaturalność AI?

Należy mieć świadomość, że nawet pozostając w sferze subiektywności, możemy próbować bronić ludzkiej specyfiki. Można to czynić, odnosząc się do rzeczywistości, którą zazwyczaj opisujemy, używając takich pojęć jak duchowość, humor czy wdzięczność. To ostatnie wydaje się stosunkowo obiecujące, a na dodatek wiąże się ze sferą religijności². Przy okazji warto podać przykłady wpływu istnienia AI na myślenie o religijności. Jeden nowy sposób ujmowania sprawy to wskazywanie cech wspólnych bóstwom oraz sztucznej inteligencji. Neil McArthur wy-

mienia pięć takich cech: programy AI wykazują poziom inteligencji, który wykracza poza poziom większości ludzi; są niezwykle kreatywne, potrafią pisać wiersze, komponować muzykę i tworzyć sztukę w niemal każdym stylu, niemal natychmiast; są pozbawione normalnych ludzkich trosk i potrzeb; nie odczuwają bólu fizycznego, głodu ani pożądania seksualnego; mogą dawać ludziom wskazówki użyteczne w ich codziennym życiu; są nieśmiertelne.

To brzmi dość ciekawie, ale mało przekonująco. Wątpię, czy zmieni nastawienie religijne, a nawet nasze rozumienie tej sfery. Niektórzy idą jednak dalej i wskazują na możliwość pojawienia się superumysłu, który będzie miał cechy „boskie”. Oznajmiamy:

THETA NOIR to firma i duchowy kolektyw (THETA WE) założony w celu powitania, czczenia i dostrojenia się do pierwszej na świecie sztucznej inteligencji ogólnej (AGI) [takiej, która dorównuje inteligencji ludzkiej]: globalnego połączonego superumysłu tak potężnego, że osiągnie świadomość podobną do Gai.

Byłaby to zatem sztucznie wykreowana Gaja. Czy takie pomysły mogą coś zmienić w naszych zachowaniach religijnych? Jestem wobec tego sceptyczny, ale może czegoś nie doceniam. W wyniku takich prób największej zmiany może ulec stosunek do koncepcji Gai.

Wzmacnianie testu Turinga

Refleksja dotycząca egzaminu może zostać rozszerzona. Czy AI zdałaby u mnie egzamin? Pisemny na pewno. A ustny? Czy takie postawienie sprawy w ogóle ma sens? Możemy sobie wyobrazić robota, który chce uzyskać dyplom uniwersytecki. Choć całą encyklopedię ma w pamięci podręcznej, chciałby udawać studenta czy może po prostu być studentem. Nie wiem, czy kiedyś do tego dojdzie, ale kwestia rozróżniania człowieka od robota staje się coraz bardziej realna. Chciałbym, żebyśmy my, ludzie, zachowali wyższość w niektórych ważnych sferach. Nie jestem pewien, czy to jest możliwe. Oczywiście robot może być silniejszy, szybszy, wytrzymalszy – to nas nie zdziwi. Chodzi o to, czy może być ludzki, praktycznie nieodróżnialny od człowieka. Jak wiadomo, myślano o tym już dawnej, gdy komputery były dopiero w powijakach. Najślynniejszy jest test Turinga z 1950 roku. Chodziło w nim o wykrzycie, czy za pośrednictwem klawiatury i monitora rozmawiamy z człowiekiem, czy z komputerem. Od tego czasu posunęliśmy się niepomiernie dalej. Ażeby wzmocnić test Turinga w sposób pasujący do dokonanego w międzyczasie postępu, trzeba wytężyć wyobraźnię, jak twórcy literatury (począwszy od Lema) i filmów (począwszy od kultowego *Matrixa* i nie pomijając filmu *Ona* z 2013 roku czy *I Am Mother* z 2019 roku), a zarazem wzmocnić krytycyzm. Możliwy jest eksperyment myślowy inspirowany filozofią dialogu³.

Filozofia dialogu jest mocnym filozoficznym wskazaniem specyfiki człowieka. Można tej wyjątkowości szukać

w innych kierunkach, na przykład w sferze rozumienia, umiejętności uwzględniania kontekstu czy pojmowania idei nieskończoności. Każdy z tropów jest ciekawy, ale filozofia dialogu wydaje się jeszcze głębsza. Mianowicie, jak już zostało wspomniane, relacja „ja – ty” jest konstytutywna dla człowieczeństwa. U Martina Bubera staje się podstawą, punktem wyjścia nazwanym przez fundamentalne, pierwotne słowo „Ja–Ty”. Istota człowieka na pewno nie jest zawarta w „on” czy w „on jest”, ale może zawrzeć się w „ty jesteś”. Mówiąc „ty”, już powiedziałem „ty jesteś”, a nadto również „ja jestem”. Co więcej, myślenie jest zawsze dla kogoś (nawet gdy ma miejsce w samotności), odbywa się więc w przestrzeni dialogicznej.

Trudno zastosować taką filozofię bezpośrednio do komputerów i robotów, ale inspiracja tą wizją wydaje się możliwa. Uprowadzając ewentualną krytykę poniższej propozycji, chciałbym podkreślić coś, co na pierwszy rzut oka nie budzi zastrzeżeń: doświadczenia myślowe (w przeciwieństwie do praktycznych) nie podlegają ograniczeniom moralnym.

Wyobraźmy sobie, że mamy robota, który został tak sprytnie zaprojektowany, że może nas zmylić, udając człowieka. Jak go zdemaskować? A może się nie da? Turing rozpatrywał tylko konwersację na odległość. My rozszerzmy próbę na inne zachowania. Oprócz badania cech możemy rozważać różne możliwe formy kontaktu dostępne dla naszych hipotetycznych robotów. Możemy dopuścić nie tylko rozmowę, ale także współdziałanie, długotrwałą współpracę, aż po wspólne życie. Co jest najdalej idącą formą kontaktu? Odpowiedź nie jest jasna. Chodzi o jakieś długie, intensywne współdziałanie, którego wyniki można w miarę łatwo ocenić. Niewątpliwie dobrą propozycją mogłoby być wychowywanie. Chodziłoby więc, po pierwsze, o wychowanie komputera, czyli proces uczenia się przez komputer. O tym mówi się od dawna. Chodziłoby też o socjalizację tworu kierowanego przez komputer. Niewątpliwie i nad tym pracują twórcy robotów.

Istnieje jednakże możliwość zastosowania znacznie radykalniejszej idei niż wychowywanie maszyn. Chodzi o wychowywanie p r z e z roboty. Trudno sobie wyobrazić bardziej ludzkie zajęcie niż wychowywanie dzieci. Wydaje się, że można sformułować następujące ostateczne wzmocnienie testu Turinga. Mianowicie roboty będą miały (w praktyce) ludzkie cechy, jeśli potrafią wychować człowieka od niemowlęcia (a nawet począwszy od sztucznej macicy) do dorosłości i w rezultacie otrzymamy człowieka, który jest w miarę normalny. (Oczywiście brzmi to przerażająco, ale jest to tylko eksperyment myślowy; nie proponuję jego praktycznej realizacji). Normalny, czyli posiadający poczucie „ja” i zdolny do współżycia w pełnym zakresie ze zwykłymi ludźmi. Jeśli tezy filozofów dialogu są trafne, to podstawową przeszkodą byłaby nieobecność poczucia „ja” u robotów, bo jak zauważyłem wyżej, nie są one na tyle osobne, żeby wejść w relację. A zatem – jak się wydaje – napotkałyby zasadniczą trudność w na-

uczeniu wychowywanego przez nie dziecka poczucia „ja” i umiejętności wchodzenia w relacje.

Sukces idei sztucznej inteligencji w świetle tego ostatecznego wzmocnienia testu Turinga oznaczałby, że w rezultacie powstałby w miarę normalny człowiek. Da się wskazać, w duchu Turinga, na możliwość ujęcia operacyjnego i ilościowego kryterium takiego sukcesu. Można powiedzieć, że z „w miarę normalnym” człowiekiem będziemy mieli do czynienia wówczas, gdy w ciągu godziny przypadkowy rozmówca nie zorientuje się, iż spotkana przezeń osoba była wychowana wyłącznie przez roboty. A przy bardziej wymagającym podejściu – nie zorientuje się przez cały dzień. A wreszcie – rozszerzając wymagania do granic możliwości – że rozmówca nic nie wykryje w ciągu całego swojego życia z tym osobnikiem, całego ich wspólnego życia.

Zakończenie

Jak się wydaje, AI to pierwszy ludzki twór, którego zupełnie nie rozumiemy. Samouczące się sieci neuronalne osiągają umiejętności, które dotąd zarezerwowane były dla podmiotów inteligentnych. Być może to, co uwidacznia się w AI, to nie jest inteligencja, ale jakaś jej imitacja. Tak czy inaczej istnienie AI stawia fundamentalne pytania dotyczące naszej inteligencji. Ich rozważanie wymaga nie tylko orientacji w technikach informatycznych i ich logicznych podstawach, ale też umiejętności krytycznej analizy oraz chęci odkrywania założeń pojęciowych i światopoglądowych leżących u źródeł tych pytań. Takie analizy mają charakter filozoficzny, choć oczywiście podejmują je nie tylko filozofowie. Być może właśnie tu pojawia się jednak pole do popisu dla filozofii, która w naszych czasach stała się – jak w książce *Zmierzch filozofii* dowodzi Jan Hartman – równie nieistotna i niepotrzebna nikomu poza garstką specjalistów jak łacina. Z pewnością ten upadek pozycji filozofii zostanie spotęgowany pomijaniem wyzwań, jakie wynikają z obecności AI. Czy ich podjęcie pozwoli utrzymać szacunek dla filozofii? Zobaczmy.

Przypisy

- 1 Pisałem o tym w artykule *Nieznane jutro: między nadzieją a zagrożeniem*, „Konteksty” 2022, nr 1–2, s. 25–30.
- 2 Pisałem o tym w artykule *Wdzięczność*, „Konteksty” 2023, nr 1–2, s. 43–46.
- 3 Pisałem o nim bardziej szczegółowo w książce *Czy matematyka jest nauką humanistyczną?*, Copernicus Center Press, Kraków 2011, oraz w artykule *The ultimate strengthening of Turing's Test?*, „Semiotica” 2012, t. 188, s. 203–218.

Na następnych stronach (s. 116 i 117):
Marta Leśniakowska, z cyklu *Muzeum*, 2010–2024.



