

JAK OSWOIĆ SZTUCZNĄ INTELIGENCJĘ?

W nawiązaniu do ukazanych prognoz łączą się dwie, całkowicie odmienne wizje przyszłych wydarzeń – dobre AI kontra złe AI; połączenie Paruzji i Apokalipsy³⁶.

Z jednej strony, przyszłość jawi się jako czas pomyślności i przyjemności oraz świat bezpośredniej natychmiastowej łączności, nieograniczonej wiedzy i nieprawdopodobnych udogodnień. Moce obliczeniowe komputerów będą narastały, narodzą się innowacyjne, prężne gałęzie gospodarki i nowe rodzaje zajęć – cyberzawody. Osiągnięcie przez roboty świadomości ma, w znacznej mierze, przyczynić się do wzrostu ich użyteczności w społeczeństwie. Mogłyby na przykład podejmować samodzielne decyzje, pracować jako sekretarki, lokaje, asystenci i pomocnicy³⁷.

Z drugiej strony, obecne wektory rewolucji technologicznej mogą prowadzić do fazy przejęcia przez sztuczną inteligencję kontroli nad kluczowymi funkcjami poznawczymi człowieka. W świetle tej hipotezy nasuwa się myśl, że jeden tylko rodzaj globalizacji będzie naprawdę wszechobejmujący: cyfryzacja „wszystkiego”, jakaś forma totalitaryzmu. Wtedy jedna globalna świadomość zatroszczyłaby się o wszystko. Dziś nie sposób określić jeszcze, do czego to „wszystko” będzie się odnosić³⁸, ale taka wizja sztucznej inteligencji, jak już podkreślono w artykule, pojawia się często. Obawy budzi scenariusz, w którym AI przejmuje kontrolę oraz przewiduje zamiary człowieka i potrafi ustalić sposób realizacji określonego postępowania w ich obliczu.

Istnieje zatem mroczna prognoza przyszłości, w której sztuczna inteligencja jest postrzegana jako zagrożenie dla ludzkiej egzystencji i przyczynia się do powstania koszmarnego świata. Niewykluczone, że do połowy tego wieku komputery staną się tak sprawne, że będą mogły zarządzać dużymi miastami, a nawet państwami. Można im będzie powierzyć stałą kontrolę nad energetyką, przepływem surowców i produktów w obrocie gospodarczym, bankowością, handlem, transportem publicznym, zaopatrzeniem w wodę i usuwaniem nie-

³⁶ Por. B r o c k m a n, dz. cyt., s. 9.

³⁷ Por. M. K a k u, *Wizje czyli jak nauka zmieni świat w XXI wieku*, tłum. K. Pesz, Wydawnictwo Prószyński i S-ka, Warszawa 2000, s. 147-197.

³⁸ Por. A. Z y b e r t o w i c z i n., *Samobójstwo Oświecenia? Jak neuronauka i nowe technologie pustoszą ludzki świat*, Wydawnictwo Kasper, Kraków 2015, s. 430-440.

czystości, środowiskiem naturalnym. Każdy defekt w obwodach sterujących ich systemem, awaria, pętla sprzężenia zwrotnego będą grozić upadkiem lub paraliżem cywilizacji o tragicznych dla ludzkości skutkach („szalone roboty”, „roboty-zabójcy”). Niewykluczone również, że systemy sztucznej inteligencji rozrosną się, podobnie jak biurokracja, do ogromnych rozmiarów. W tym sensie nieumyślnym zagrożeniem mogą stać się posłuszne, „spolegliwe roboty”, które będą znakomicie spełniać swoją misję³⁹. Co więcej, jeśli któraś ze ścieżek powstania superinteligencji się ziści, to pojawić się może „stwór” wyposażony w nadludzką inteligencję, który mógłby zagrozić przyszłości gatunku ludzkiego, chociaż mielibyśmy jako ludzie jedną przewagę – to my bylibyśmy jego twórcami⁴⁰.

Zasygnalizowane potencjalne zagrożenia wskazują, że sztuczna inteligencja będzie musiała podlegać coraz ściślejszej kontroli, aby nie pojawiły się niepożądane konsekwencje. Roboty należy wyposażać w urządzenia zabezpieczające, które nie pozwoliłyby im przejąć władzy na Ziemi. Być może powinna powstać nowa gałąź sztucznej inteligencji zajmująca się stricte problemem utrzymywania układów sztucznej inteligencji pod kontrolą⁴¹. W praktyce zagadnienie kontroli nie jest jednak proste. Wręcz przeciwnie – jawi się jako mocno skomplikowane, a sama technologia wydaje się być współczesnym paradoksem, gdyż z jednej strony tworzy ją człowiek, ale z drugiej tak naprawdę nic o niej nie wie⁴².

Należy podkreślić, że ostrzeżenia dotyczące niekontrolowanego postępu technologicznego są znane od lat. Nie są one natomiast szerzej publikowane, a dotychczasowe oficjalne dyskusje wydają się niedostateczne. Dzieje się tak między innymi dlatego, że – jak pisze w swoim esej *Dlaczego przyszłość nas niepotrzebuje?* Bill Joy⁴³ – opisywanie zagrożeń nie sprzyja osiąganiu zysków, a najnowocześniejsze technologie mają wyraźnie komercyjne zastosowanie i rozwija się je niemal wyłącznie w korporacjach. W wieku triumfującego komercjalizmu wraz z nauką dostarczają serii innowacji fenomenalnie lukratywnych.

Rozważając o przyszłości inteligentnych maszyn i o tym, czy przejmą one władzę i zaczną same decydować o sobie, warto dodać, że najpierw ktoś lub

³⁹ Por. K a k u, dz. cyt., s. 174-197.

⁴⁰ Por. B o s t r o m, dz. cyt., s. 11.

⁴¹ Por. K a k u, dz. cyt., s. 174-197.

⁴² Interesujące scenariusze w kontekście rozwoju technologii prezentuje Natalia Hatałska (por. N. H a t a ł s k a, *Wiek paradoksów. Czy technologia nas ocali?*, Wydawnictwo Znak, Kraków 2021). Każdy ze scenariuszy można zawsze zmienić, ostatecznie przyszłość zależy od naszych wyborów tu i teraz.

⁴³ Por. B. J o y, *Dlaczego przyszłość nas nie potrzebuje?*, w: *Wybierz czerwoną pigułkę: Nauka, filozofia i religia w Matrix*, red. G. Yeffeth, tłum. W. Derechowski, Wydawnictwo Helion, Gliwice 2003, s. 229-245.

coś musi zdefiniować zadanie maszyny, czyli to, co maszyna ma wykonać. Dla człowieka określanie celów jest zazwyczaj powiązane z warunkami biologicznymi i psychologicznymi, historią osobistą, środowiskiem kulturowym, historią cywilizacji. Zamiary są unikalną cechą ludzi. Jeśli chodzi o maszyny, to ludzie nadają im cel po ich skonstruowaniu. Stąd, rozważając przyszłość, nie należy pomijać problematyki związanej z zadaniami AI w oparciu o prawa matematyki i fizyki. Inteligentnym systemom można wgrać szkodliwe cele, dlatego też kluczowe będzie zaprojektowanie technologicznej infrastruktury zdolnej do wykrywania i kontrolowania zachowań systemów szkodliwych. Konieczne jest tworzenie prawnych i ekonomicznych ram stymulujących zachowania AI na bazie intelektualnych i technologicznych zasobów, jakimi dysponujemy jako ludzie.

Jednak nie po raz pierwszy jesteśmy w sytuacji, kiedy nowo powstała technologia wydaje się stwarzać zagrożenie dla naszej egzystencji. Wynalezienie bomby atomowej oraz tworzenie jądrowych arsenałów groziło i stale grozi zagładą świata. Z kolei kiedy po raz pierwszy pojawiła się technologia rekombinowania DNA, towarzyszył temu lęk, że zmodyfikowane genetycznie organizmy wydostaną się na wolność i doprowadzą do śmierci wielu ludzi na całym świecie. Jak się wydaje, osiągnięcia w dziedzinie uczenia maszynowego stwarzają stosunkowo niewielkie zagrożenia w porównaniu z bronią nuklearną czy śmiertelnościami organizmami. Stąd można dywagować, że do sztucznej inteligencji również się przyzwyczaimy, co zresztą już się dzieje.