



e-ISSN 2084-1043 p-ISSN 2083-6635
Published online: 31.03.2025

ARGUMENT
BIENNIAL PHILOSOPHICAL JOURNAL
Vol. 14 (2/2024) pp. 319-324
www.argument-journal.eu

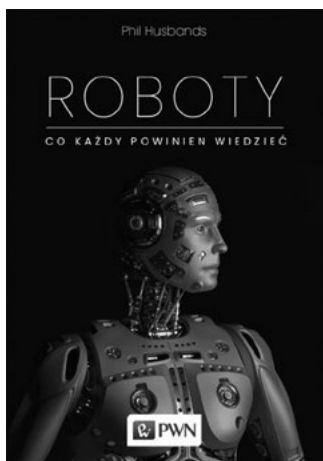
Phil Husbands, *Roboty: co każdy powinien wiedzieć*

tłum. Wojciech Fenrich

Warszawa: Wydawnictwo Naukowe PWN, 2023, 214 s.

ISBN: 978-83-01-22796-8 (druk)

ISBN: 978-83-01-22927-6 (ebook)



Sztuczna inteligencja (*artificial intelligence*, AI) stopniowo, lecz nieubłaganie przenika do wszystkich sfer naszego życia, stając się jego nieodłącznym elementem. Każdego dnia naukowcy i inżynierowie udoskonalają jej możliwości, sprawiając, że skuteczniej radzi sobie ona z różnorodnymi trudnościami oraz nowymi wyzwaniami. Potrafi analizować coraz większe i bardziej złożone zbiory danych, znacząco poprawia jakość tłumaczenia tekstów oraz precyzyjnie rozpoznaje i przetwarza ludzką mowę. W miarę rozwoju tych zdolności AI będzie nie tylko automatyzować liczne procesy, ale również wspierać kreatywne działania człowieka, dostosowując się do naszych indywidualnych

potrzeb i preferencji, co otworzy nowe możliwości w takich dziedzinach jak edukacja, medycyna, a nawet codzienna komunikacja. W nadchodzących latach obecność AI będzie jeszcze bardziej powszechna i intensywna. Myli się ten, kto sądzi, że problemy związane ze sztuczną inteligencją i robotami go nie dotyczą. Warto zatem już dziś zdać sobie sprawę z ich rosnącego wpływu na nasze życie i świat oraz aktywnie angażować się w zrozumienie natury AI oraz potencjalnych konsekwencji jej użytkowania. Obecna AI rozwijana jest oczywiście ze względu na wiele różnych zastosowań, w tym na użytek robotyki i robotów. Co prawda proste roboty mogą istnieć bez sztucznej inteligencji, ale znacznie zwiększa ona ich zdolności, pozwalając im na planowanie, podejmowanie decyzji, uczenie się i adaptację w dynamicznie zmieniających się warunkach. Trzeba

więc zapoznać się nie tylko ze sztuczną inteligencją, ale też ze światem robotów. Pomóc mogą nam w tym dobre książki, do których moim zdaniem należy praca *Roboty: co każdy powinien wiedzieć* (*Robots: What everyone needs to know*) w tłumaczeniu Wojciecha Fenricha.

Powyższa pozycja to przystępnie napisany przewodnik po współczesnym świecie robotyki, autorstwa Phila Husbandsa, profesora informatyki i sztucznej inteligencji na Uniwersytecie w Sussex, inżyniera i współtwórcy robotyki ewolucyjnej. Główne tematy książki osadzone są w pewnym kontekście historycznym. Autor opisuje rozwój robotyki w czasie i wskazuje, jak zjawisko to ewoluowało zarówno w rzeczywistości (począwszy od XVIII-wiecznych robotów), jak i w popkulturowych wyobraźni oraz twórczości. Husbands analizuje, jak filmy, literatura i inne media kształtowały nasze rozumienie robotów. Ale, co najważniejsze, omawia współczesne postępy w dziedzinie robotyki, rozważając, jak daleko posunęła się technologia w zakresie automatyzacji, sztucznej inteligencji oraz robotów przemysłowych i przeznaczonych do użytku domowego.

Książka podkreśla, jak maszyny uczą się coraz lepiej przetwarzać dane, rozumieć ludzką mowę i tłumaczyć teksty, ale jednocześnie zwraca uwagę na wyzwania technologiczne, takie jak ograniczenia sztucznej inteligencji czy trudności w pełnej autonomizacji robotów. Husbands nie tylko opisuje, co obecnie osiągnięto w dziedzinie robotyki, ale również wnikliwie analizuje, jakie wyzwania stoją przed technologią w kontekście społecznym. Skutecznie zestawia fascynujące postępy technologiczne z pytaniami o to, jak te zmiany wpłyną na nasze codzienne życie, pracę i relacje międzyludzkie. Samo zestawienie tytułów i podtytułów daje dość dobre wyobrażenie zakresu poruszanej w niniejszej książce tematyki, czym będziemy się kierować w poniższej recenzji.

W rozdziale I pt. „Roboty już tu są” Husbands stawia między innymi takie pytania jak: O czym myślimy, myśląc o robotach? Dlaczego roboty powinny nas obchodzić? Dlaczego roboty nas fascynują, a niekiedy wprawiają w popłoch? Ile robotów istnieje? Na to ostatnie pytanie autor książki odpowiada: „Trudno jest podać dokładną liczbę, ale rozsądne szacunki określają, że w 2019 roku w regularnym użyciu było na świecie od 13 do 15 milionów robotów. Są one wykorzystywane w przemyśle i handlu, szpitalach, szkołach i wielu innych miejscach” (s. 19).

W rozdziale II, zatytułowanym „Podstawy”, autor pyta na przykład o to, czym dokładnie jest robot. (Najogólniej mówiąc, jest to „maszyna korzystająca z sensorów do zbierania informacji o świecie i aktuatorów do działania w nim”; s. 21). Jakiego rodzaju robotów istnieje? Czy wszystkie roboty zachowują się tak samo? Jak działają roboty? Czy różne sensory mogą zmienić możliwości robota? W jaki sposób poruszają się roboty posiadające nogi? Czy istnieje jedno najlepsze podejście do projektowania robotów? Czy roboty wirtualne się wliczają? Co potrafią roboty? Czemu roboty nie potrafią? Chociaż maszyny osiągnęły już wiele, to droga do stworzenia takich, które w pełni dorównają ludziom (pod względem inteligencji), jest jeszcze długa.

Nieodłączną częścią zrozumienia robotyki jest kontekst historyczny. Pozwala on nam docenić złożoność tej dziedziny, zidentyfikować kluczowe momenty w jej rozwoju oraz zrozumieć, jak robotyka kształtuje nasz świat. Stąd w rozdziale III pt. „Odrobina historii” autor odpowiada na następujące pytania: Kiedy powstały pierwsze prawdziwe roboty? Kiedy po raz pierwszy pojawiły się naprawdę autonomiczne roboty? Co potrafiły żółwie Greya Waltera? Jak działały żółwie? Czym jest cybernetyka? Kiedy roboty przemysłowe pojawiły się w fabrykach? Kiedy rozpoczęły się prace nad inteligentnymi robotami autonomicznymi? Jak dobrze spisywały się te roboty? Co stało się dalej?

W rozdziale IV, zatytułowanym „Wewnątrz maszyny”, Husbands przygląda się działaniu robotów: Jak działają samochody bez kierowcy? Kiedy zaczęły się prace nad samochodami bez kierowcy? Jak działają sztuczne sieci neuronowe wykorzystywane w samochodach i robotach? Jaka była struktura sieci neuronowej ALVINN-a? Skąd pojazd autonomiczny „wie”, gdzie się znajduje? Z jakich czujników korzystają współczesne pojazdy autonomiczne? Jakie są najważniejsze sposoby podejścia do kwestii rozpowszechnienia w pełni autonomicznych pojazdów? Czy kopiowanie biologii to dobry pomysł? Jak niezawodne są współczesne roboty? Jak to naprawdę jest z ich inteligencją?

Rozdział V pt. „Robotyczne imaginarium: roboty w kulturze popularnej” odwołuje się do wytworów ludzkiej wyobraźni, przede wszystkim w dziedzinie literatury. Odnajdujemy w nim następujące pytania: Kiedy maszyny przypominające roboty po raz pierwszy pojawiły się w literaturze? A co z wytworem bioinżynierii — organicznymi robotami? Na ile wiarygodne są fikcyjne sposoby przedstawienia robotów w kulturze głównego nurtu? Czy wyimaginowane roboty mogą nam pomóc w przewidzeniu potencjalnych skutków, jakie niesie ze sobą technika? Autor zwraca uwagę na pewną ambiwalencję w związku z tym tematem. Wytwory ludzkiej wyobraźni inspirować rozwój rzeczywistych technologii i pozwalają na przewidywanie potencjalnych problemów, a równocześnie mogą rodzić nierealistyczne oczekiwania wobec technologii i generować mylne wyobrażenia na temat możliwościach AI.

W rozdziale VI, zatytułowanym „Inteligencja, superinteligencja i cyborgi”, autor zwraca się w stronę przyszłego rozwoju sztucznej inteligencji i osobliwości: Wiele mówi się o osobliwości w odniesieniu do robotyki, ale czym ona jest? Czy osobliwość jest blisko? Czy powinniśmy się martwić? Czy powstanie superinteligentnych robotów jest nieuniknione? A co z cyborgami? Czy będziemy w stanie ulepszyć się za pomocą technik robotycznych? Czy powinniśmy obawiać się robotycznych ulepszeń cyborgicznych? Na to ostatnie pytanie Husbands odpowiada: „W tej chwili nieszczególnie. Prawie wszystkie cyborgiczne zastosowania mają pozytywny, medyczny charakter i są odpowiednio uregulowane” (s. 139).

Konsekwencji rozwoju AI i praktycznych zastosowań robotów poświęcony jest w całości rozdział VII pt. „Roboty w pracy”. Rozważane są w nim następujące

pytania: Jakie prace są obecnie wykonywane przez roboty? Kiedy w pełni autonomiczne pojazdy staną się powszechne? A co z robotami w ochronie zdrowia i opiece społecznej? Jak roboty są wykorzystywane w edukacji? W jaki sposób z robotów korzysta wojsko? Czy w sztuce jest miejsce dla robotów? Jak roboty są wykorzystywane w naukach teoretycznych? Czy roboty przejmą nasze miejsce pracy? Czy powinniśmy pozwolić robotom na odebranie nam pracy? Czy roboty gruntownie zmieniają sposób, w jaki żyjemy?

W rozdziale IX, zatytułowanym „Różne wersje przyszłości robotów”, autor podejmuje kwestię przyszłości maszyn. Niechętnie i zdawkowo rozważa, jak możemy korzystać z robotów w niezbyt odległej i dalszej przyszłości. Na koniec stwierdza: „Różne wersje przyszłości nie nabrały jeszcze kształtów. To od nas wszystkich zależy, czy wykorzystamy je najlepiej, jak to możliwe” (s. 195).

Czytelnik zauważył zapewne, że w omówieniu tym pominięto rozdział VIII, poświęcony etyce robotów, rozdział — w mojej ocenie — bardzo ważny. Pomiąłem go świadomie, by teraz skoncentrować się na jego treści. Dla Husbandsa etyka robotów definiowana jest dość szeroko. Zwykle odróżnia się „roboetykę” lub etykę robotów (*robot ethics*) od etyki maszyn (*machine ethics*), twierdząc, że ta pierwsza koncentruje się na etycznym postępowaniu twórców i użytkowników robotów, a nie na implementacji zasad etycznych w samych maszynach, czym zajmuje się ta druga. Etyka maszyn stanowi część etyki AI (*ethics of artificial intelligence*). Wedle Husbandsa

etyka robotów jest [...] zastosowaniem etyki w obrębie robotyki i kieruje tym, w jaki sposób projektujemy, wykorzystujemy i wchodzimy w interakcje z robotami, a także tym, jak one odnoszą się do nas. Obejmuje zarówno nasze zachowania — to, jaki użytek robimy z robotyki w obrębie społeczeństwa — jak i zachowania robotów: jak możemy sprawić, by zachowywały się etycznie (s. 167).

Dlaczego etyka robotów jest ważna? Moim zdaniem przede wszystkim dlatego, że roboty są masowo wykorzystywane w krytycznych obszarach życia publicznego, energetyce, produkcji, logistyce, służbie zdrowia czy bezpieczeństwie; ich działania mogą bezpośrednio wpływać na ludzkie życie i bezpieczeństwo. Podobnie argumentuje autor, choć idzie o krok dalej. Jego zdaniem obecnie, gdy już roboty autonomiczne i obdarzone prymitywną inteligencją „zaczynają pojawiać się na naszych drogach i w naszych domach, a także są wykorzystywane przez wojsko, nagle poważne pytania etyczne, prawne i filozoficzne stają się bardzo aktualne” (s. 168). Takie roboty zdaniem Husbandsa powinny też ponosić społeczną i moralną odpowiedzialność.

Już pod koniec rozdziału V autor zauważył, że obecnie

przybiera na sile pewne mroczne pytanie, które stanowi sedno wielu dystopijnych wizji *science fiction*: jeśli kiedykolwiek uda nam się stworzyć maszyny o podobnym do ludzkiego poziomie inteligencji i samoświadomości, zbudowane tak, by zintegrować

się z ludzkim społeczeństwem, czy nieuchronnie przejmą one nasze najgorsze cechy — skłonność do przemocy i wywierania przymusu, mściwość, chciwość i nienawiść? (s. 128)

Odpowiedzi należy szukać już w kolejnym rozdziale, poświęconym osobliwości, przez którą należy rozumieć hipotetyczny moment eksplozji AI; moment, w którym przekroczy ona możliwości umysłowe człowieka i — być może — zacznie stwarzać dla niego egzystencjalne zagrożenie. Problem ten zdaniem autora nie jest jednak tak poważny, jak niektórzy sądzą. Przede wszystkim nie ma przekonujących dowodów, że osobliwość jest blisko. Wróćmy jednak do pytania, czy superinteligentne maszyny przejmą nasze najgorsze cechy. Moim zdaniem sprawa nie jest oczywista.

Wiele przemawia za potencjalnym przejściem przez AI naszych negatywnych cech, negatywnych wzorców zachowań. Sztuczna inteligencja uczy się bowiem na podstawie wielu danych i wzorców, które zawierają to, co jest typowe dla człowieka, włącznie z jej wszystkimi dobrymi i niestety złymi nastawieniami. AI, podobnie jak ludzie, może być skłonna do walki o kontrolę i dominację (to zresztą najczęstszy scenariusz w książkach i filmach *science fiction*). Niewykluczony jest też inny scenariusz. Twórcy AI starają się „wyposażyć” sztuczną inteligencję w różnorodne systemy bezpieczeństwa, które uniemożliwią jej podejmowanie negatywnych lub szkodliwych działań. Właśnie w tym kierunku zmierzają obecne badania nad sztuczną inteligencją. Po pierwsze, twórcy AI zbierają dane z różnorodnych źródeł i odpowiednio je generują, tworząc wzorce (ściśle rzecz biorąc, zbieranie danych nie należy do obowiązków twórców AI — zajmują się tym działające na ich zlecenie wyspecjalizowane zespoły badawcze lub firmy analityczne). Po drugie, próbują oni wbudować algorytmy minimalizujące nierówności i zapewniające równość oraz sprawiedliwość (*fairness algorithms*). Algorytmy monitorujące zapobiegają dyskryminacji, redukują uprzedzenia (*bias mitigation algorithms*) i pomagają uniknąć sytuacji, w której jedna grupa byłaby faworyzowana kosztem innej. Twórcy AI próbują wyposażyć sztuczne podmioty moralne w zbiór różnych wartości i zasad moralnych, które podmioty te będą musiały uwzględniać podczas podejmowania decyzji. To wszystko oczywiście nie gwarantuje, że systemy te będą pozbawione naszych złych cech i będą postępować tylko dobrze. Dopuszczam jednak możliwość, że odpowiedzialnie tworzona AI będzie czynić więcej dobra niż człowiek.

Odpowiedź na pytanie, czy roboty z ogólną AI nieuchronnie przejmą nasze najgorsze cechy, nie może być jednoznaczna, ponieważ nie wiemy, jaka będzie jej natura. Otóż sztuczna inteligencja może mieć zupełnie inną naturę niż ludzie, co może oznaczać, że niekoniecznie będzie przeżywać i myśleć jak my. Natura AI, tak jak ją dzisiaj widzimy, w przeciwieństwie do ludzkiej, (1) nie jest ograniczona biologicznie, (2) nie ma wewnętrznego „ja”, które świadomie i w pełni autonomicznie wybiera pewne cele. (1) oznacza, że AI nie ma motywów

biologicznych i nie podlega tym samym procesom fizjologicznym, które wpływają na ludzkie potrzeby, emocje i motywacje. Uposażenie neuronalne (w mózgu ludzkim) jest poddawane wpływom hormonów, chemicznych substancji (neurotransmitterów) oraz biologicznych cykli (na przykład sen, odżywianie), kształtującym nasze samopoczucie, emocje i działania. AI nie doświadcza tych procesów. (2) oznacza, że AI nie ma świadomego „ja”, nie posiada własnych pragnień, motywacji i celów, lecz działa na podstawie algorytmów i zestawów reguł. Jej „motywacje” są w rzeczywistości predefiniowanymi celami, które mogą być określone przez programistów lub systemy samouczące. Cele, które realizuje AI, są *de facto* celami ustalonymi przez ludzi, którzy ją tworzą. To wszystko — i wiele innych kwestii — wskazuje na fundamentalnie różną naturę AI. Z kolei jej inna natura może prowadzić do innego rodzaju współpracy z ludźmi i generować inne działania.

Autor w niniejszej książce nie stara się odpowiedzieć na wiele stawianych przez siebie pytań. Zakreśla jedynie obszar kluczowych zagadnień natury etycznej, wynikających z prowadzonego wywodu. Zastanawia się między innymi, czy twórcy zaawansowanych robotów powinni być za nie społecznie i moralnie odpowiedzialni. Pyta: A co z przypadkami, w których etyczne roboty są używane w nieetyczny sposób? Kto powinien zostać pociągnięty do odpowiedzialności, jeśli coś pójdzie źle? Czy kwestie etyczne dotyczące robotów są różne od tych, które pojawiały się w wypadku wcześniejszych technologii? To niewątpliwie coraz ważniejsze kwestie, które pilnie domagają się większego uporządkowania w świecie coraz szybciej rozwijającej się AI.

Husbands podejmuje szerokie spektrum zagadnień związanych z robotyką, co jest niewątpliwie zaletą jego książki. Niestety ambicja objęcia tak rozległego tematu sprawia, że niektóre wątki zostają jedynie zarysowane, a czytelnik pozostaje z niedosytem głębszej analizy. Być może zamysłem autora było właśnie określenie piętrzących się problemy i poszukanie ich źródeł we wcześniejszych działaniach twórców robotyki i prostych wersji sztucznej inteligencji. Stąd również chęć sięgania po inspiracje licznych powieści i filmów *science fiction* dotyczących tych zagadnień. Na uwagę zasługuje fakt często głębokich refleksji filozoficznych ukrytych w takich właśnie dziełach, ukazujących często dalekosiężne konsekwencje przyjęcia ściśle określonych założeń czy ograniczeń programistycznych. Interesujące są również próby implementacji typowo ludzkich kodeksów i rzekomo uniwersalnych norm (zrozumiałych i jasnych dla istot ludzkich) w sytuacji konfliktu danych. Co miałyby zrobić na przykład humanoidalny robot wyposażony w algorytmy AI, kiedy będzie musiał wykonać sprzeczne instrukcje?

Rosnące lawinowo możliwości wykorzystywania algorytmów AI pokazują aż nadto wyraźnie, że mogą one służyć do różnych celów, na przykład do skutecznego wprowadzania dezinformacji na niemal każdym polu ludzkiej działalności. Połączenie zaś tak nieunormowanej sztucznej inteligencji z cybernetycznym

ciałem może przynieść nieprzewidywalne na dzień dzisiejszy skutki. To, co jeszcze niedawno pozostawało jedynie w sferze fikcji, już dziś staje się realne. Dlatego uważam, że książka Phila Husbandsa *Roboty: co każdy powinien wiedzieć* jest pod tym względem niezwykle inspirująca i także aktualna.

*Dariusz DĄBEK**

* Dr, adiunkt, Instytut Dziennikarstwa i Stosunków Międzynarodowych, Uniwersytet Komisji Edukacji Narodowej w Krakowie. E-mail: dariusz.dabek@uken.krakow.pl.

