

PROPOZYCJE „ETHOSU”

Jobst Landgrebe, Barry Smith, *Why Machines Will Never Rule the World: Artificial Intelligence without Fear*, Routledge, New York – London 2022, ss. 341.

Rozwój sztucznej inteligencji (AI) jest tematem zaprzatającym umysły naukowców z różnych dyscyplin, polityków, przedsiębiorców i zwykłych ludzi, a postawy wobec tej kwestii wahają się od bezkrytycznego entuzjazmu do autentycznego strachu przed buntem inteligentnych robotów, które mogą uznać nas, ludzi, za niższy gatunek, podporządkowując nas, a ostatecznie – eliminując. Książka Jobsta Landgrebe’a, naukowca i przedsiębiorcy z filozoficznym wykształceniem, i Barry’ego Smitha, wybitnego filozofa, jest odpowiedzią na te obawy: silna (ang. general) sztuczna inteligencja jest matematycznie – nie zaś jedynie technicznie czy fizykalnie – niemożliwa. „Spokojnie. Maszyny nigdy nie będą rządziły światem” (s. x; tłum. fragmentów książki – A.L.K.) – piszą autorzy książki. Definiują oni sztuczną inteligencję ogólną (AGI) jako tę, która prezentuje poziom wyższy od inteligencji ludzkiej, lub co najmniej jej równy, i jest zdolna do radzenia sobie z problemami pojawiającymi się w otaczającym człowieka świecie w stopniu adekwatności co najmniej podobnym do tego, z jakim radzi sobie z nimi człowiek. Nie negują oczywiście postępu w budowaniu sztucznych inteligencji przeznaczonych do konkretnych zadań, kwestionują natomiast nieograniczony postęp w symulacji działania człowieka. Argument na rzecz ich tezy, w nieco uproszczonej formie, przedstawia się następująco: (1) zbudowanie AGI wymaga napisania oprogramowania, które naśladowałoby ludzki system neuropoznawczy (autorzy rozważają i ostatecznie odrzucają alternatywne sposoby wykreowania AGI); (2) by napisać takie oprogramowanie, należałoby zbudować matematyczny model tego systemu, umożliwiający przewidywanie jego zachowania; (3) nie jesteśmy w stanie budować matematycznych modeli dynamicznych systemów złożonych, choć systemy te podlegają prawom fizyki; (4) ludzki system neuropoznawczy jest dynamicznym systemem złożonym, a wobec tego nie jesteśmy w stanie zbudować wymaganego matematycznego modelu tego systemu i dlatego nie możemy napisać oprogramowania, które naśladowałoby ludzki system neuropoznawczy. Ergo: stworzenie AGI jest niemożliwe z powodu ograniczeń płynących z matematyki. Pozyskiwanie pieniędzy na badania nad AGI zostało więc oparte na

falszywych obietnicach, a strach przed maszynami rządzącymi światem jest nieuzasadniony.

W pierwszej części książki Landgrebe i Smith w kolejnych rozdziałach prowadzą nas przez analizy ludzkiego umysłu i problemu relacji między umysłem a ciałem, ludzką i maszynową inteligencją, zagadnienia natury ludzkiego języka i komunikacji oraz ludzkich zachowań społecznych i moralnych. W dwóch rozdziałach części drugiej autorzy poddają refleksji naturę systemów złożonych i rozmaite podejścia do ich matematycznego modelowania, wskazując na wsobne ograniczenia tego procesu. Im bardziej złożony jest system, tym mniejsze są szanse na jego adekwatne modelowanie i tworzenie dla niego AI. Część trzecia poświęcona została możliwościom sztucznej inteligencji i jej ograniczeniom. Tytuły kolejnych rozdziałów stanowią zarazem tezę, że maszyny nie osiągną oczekiwanego poziomu inteligencji, nie opanują ludzkiego języka i nie opanują interakcji społecznych w stopniu wyższym czy choćby równym temu, który jest dostępny człowiekowi. Możliwe jest natomiast matematyczne modelowanie częściowe – i właśnie tam znajduje zastosowanie sztuczna inteligencja.

W dyskusjach autorzy sięgają do przykładów z medycyny, inżynierii, ekonomii, psychologii, odpowiadając na obietnice i obawy wielu współczesnych myślicieli, w tym wiodących teoretyków transhumanizmu: Nicka Bostroma i Raya Kurzweila czy Elona Muska. Rozważania Landgrebe'a i Smitha pozostają pod wyraźnym wpływem fenomenologii, w tym Edmunda Husserla (z okresu *Doświadczeń filozoficznych*), Adolfa Reinacha i Maxa Schelera, czego autorzy bynajmniej nie ukrywają, powołując się na tych myślicieli, zwłaszcza w odniesieniu do analiz natury człowieka, jego poznania i woli, a jednocześnie pokazując niewystarczalność rozmaitych podejść filozofii analitycznej do uchwycenia natury naszego świata. Autorzy deklarują, że nadrzędną ideą, która spaja wszystkie części książki, jest uznanie istnienia nieredukowalnych do siebie rzeczywistości: fizycznej, biologicznej, społecznej i mentalnej, oraz przyjęcie realistycznej filozofii dotyczącej świata dostępnego w naszym codziennym doświadczeniu (por. s. 3n.). To podejście – wraz z odwołaniem do Maxa Schelera koncepcji inteligencji praktycznej – pozwala autorom odróżnić inteligencję pierwotną od obiektywizującej (ang. objectifying) czy uprzedmiotowującej. Ta pierwsza przysługuje bytom ludzkim i pozaludzkim (choćby zwierzętom) i nie jest rezultatem uczenia się, ale raczej umożliwia uczenie się poprzez adaptację do nowych sytuacji oraz rozwinięcie pierwotnego zachowania kulturotwórczego. Ta druga przysługuje ludziom i umożliwia kumulatywną ewolucję kultury: „Inteligencja obiektywizująca pozwala homo sapiens zdystansować się od swojego środowiska w sposób, który umożliwia widzenie siebie, innych ludzi i elementów tego środowiska (biologicznych i niebiologicznych) jako przedmiotów, z których każdy ma własną trajektorię i własny zbiór właściwości i mocy kauzalnych” (s. 46). Inteligencja ta obejmuje zdolność ujmowania świata poprzez kategorie ogólne, intencjonalność, rozpoznawanie przyczyn i celów, panowanie nad koniecznościami biologicznymi, zdolność decydowania (czyli wyboru i niejako ustanawiania siebie źródłem działania) oraz wyobrażania sobie i planowanie długoterminowe. To wszystko musiałaby posiadać AGI, a skoro nie da się tego modelować matematycznie, to inteligencja sztucznej inteligencji – jakkolwiek by ją rozumieć, co autorzy starannie dyskutują – nie

przekroczy inteligencji ludzkiej ani nawet jej nie dorówna. Ludzki mózg nie jest komputerem i żadna znana matematyka nie może go całkowicie adekwatnie opisać. Landgrebe i Smith ostatecznie głoszą zatem wyjątkowość ludzkiego bytu. Sztuczna inteligencja nas nie zastąpi. Paradoksalnie, dyskusja nad AGI włącza się pośrednio w dyskusję nad naturą człowieka – nad tym, czy jesteśmy jedynie zwierzętami, czy też jesteśmy „inaczej i wyżej”, by przywołać frazę ukutą przez Tadeusza Stycznia, reprezentanta lubelskiej szkoły filozoficznej. A jeśli faktycznie status bytowy człowieka sytuuje nas wyżej, to jaka jest nasza właściwa relacja do bytów pozaludzkich, jakie mamy wobec nich prawa i obowiązki?

Argumentacja Landgrebe’a i Smitha odwołuje się do filozofii (różnych tradycji), matematyki, informatyki, językoznawstwa, psychologii, antropologii kulturowej, socjologii, fizyki i biologii. W tym sensie są to rozważania naprawdę interdyscyplinarne, co pokazuje naturę problemu sztucznej inteligencji – nie jest to bowiem problem czysto techniczny. Owa interdyscyplinarność nie ułatwia lektury książki. Zdając sobie z tego sprawę, autorzy wyposażyli ją w słownik kluczowych pojęć, indeks i aneks z matematycznym ujęciem turbulencji. Wskazują też fragmenty, które czytelnik może bez szkody dla zrozumienia argumentacji pominąć. Bibliografia zawiera imponującą liczbę pozycji i sama może stanowić źródło dla badaczy AI.

Książka wpisuje się w literaturę kwestionującą nieograniczone możliwości sztucznej inteligencji. Jednym z pierwszych autorów prezentujących takie stanowisko był Hubert Dreyfus, który w roku 1972 opublikował książkę *What Computers Can't Do: A Critique of Artificial Reason* (Harper and Row, New York-Evanston-San Francisco-London). Do listy można dopisać Rogera Penrose’a *Shadows of the Mind: A Search for the Missing Science of Consciousness* (Oxford University Press, Oxford 1994) czy Roberta J. Marksa II, *Non-computable You: What You Do That Artificial Intelligence Never Will* (Discovery Institute, Seattle 2022). Nawet jeśli Landgrebe i Smith mają rację – a rodzi się przecież szereg pytań, choćby o nową matematykę, o możliwości komputerów kwantowych czy znalezienie strategii budowania AGI innej niż symulacja ludzkiego umysłu – nie możemy spocząć w spokoju ducha. Jeśli nawet AGI jest niemożliwa, wciąż rozwija się przecież „zwykła” sztuczna inteligencja. „Mimo poważnych ograniczeń, które opisaliśmy w tej książce – twierdzą autorzy – odkrycie coraz bardziej pomysłowych i potężnych zastosowań wąskiej woli sztucznej inteligencji w nadchodzących dekadach w ogromnym stopniu pogłębi i wzmocni „technosferę” (s. 301), zmieniając świat naszego codziennego doświadczenia. Wyzwaniem – jak piszą – nie jest zapobieganie zastąpieniu ludzkości przez superinteligencję czy tworzenie „etycznej” AI, ale przewidywanie konsekwencji pojawiania się „wąskiej” sztucznej inteligencji w różnych sektorach życia i reagowanie na te konsekwencje. Można złośliwie twierdzić, że AI już do pewnego stopnia rządzi światem, bo wiele naszych decyzji sterowanych jest informacjami gromadzonymi przez sztuczną inteligencję, nasze czynności są zapośredniczone przez urządzenia wyposażone w AI, a sztuczna inteligencja staje się uczestnikiem naszego życia społecznego (na przykład w postaci empatycznych robotów). Obecność AI zmienia nas samych, nasze relacje z innymi ludźmi i nasze rozumienie świata. Pytania o to, gdzie, dlaczego

i pod jakimi warunkami powinniśmy wprowadzać sztuczną inteligencję, stają się coraz bardziej istotne. Warto zatem przeczytać książkę Landgrebe’a i Smitha, by z jednej strony lepiej zrozumieć, o co należy, a o co nie należy się martwić w związku z rozwojem sztucznej inteligencji, a z drugiej strony, by uświadomić sobie, kim jesteśmy i co jest stawką w rozwijaniu AI. Co więcej, książka jest znakomitym dowodem na to, że bez filozofii nie udaje się podjąć palących problemów współczesności (o czym zdają się zapominać adwokaci stawiania na kształcenie praktyczne, które – jak twierdzą – jest niezbędne do uzyskiwania przewagi na rynku pracy).

A.L.K.