



Mariusz WOJEWODA*

PROJEKT GOLEM XXV

Oczekiwanie na technologiczną poprawę jakości życia

Przyjaźń z robotem, który nie ma osobowości, albo dowolnie zmienia swoją osobowość, wykracza poza zrozumiałe dla człowieka schematy relacji – to jak przyjaźnić z się z kimś całkowicie racjonalnym, ale niestabilnym psychicznie. Lemowska myśląca maszyna jest Kimś jedynym i odrębnym, a jednocześnie Każdym (Kimś dowolnym). Aby stać się osobą – co w ludzkim przekonaniu jest najdoskonalszą formą istnienia – Golem musiałaby się umysłowo zdegradować. Uznanie takiego stanu rzeczy wprawia nas w konfuzję.

Nie jestem osobą rozumną, lecz Rozumem, co w obrazowym przesunięciu znaczy, że nie jestem czymś takim jak Amazonka lub Bałtyk, lecz czymś takim jak woda [w Amazonce lub Bałtyku – MW]¹.

Stanisław Lem

Rozwój cywilizacji pociąga za sobą zmiany technologiczne. Zjawisko to nazywane jest w duchu dziewiętnastowiecznego ewolucjonizmu „postępem technicznym”. Określenie to niesie ze sobą nadzieję na pozytywną zmianę i na ukojenie lęków przed dramatycznymi wydarzeniami: wojną, epidemiami czy robotyzacją, która może nas pozbawić miejsc pracy. Oczekujemy uwolnienia od sytuacji trudnych, takich jak choroby, cierpienia, przedwczesna śmierć, doświadczenie samotności. Jednocześnie oczekujemy sposobu na zapewnienie sobie i swoim bliskim odpowiedniej jakości życia. Pytamy, czy poprawy takiej należy oczekiwać w efekcie wprowadzenia rozwiązań proponowanych przez współczesną technikę, obecnie zwłaszcza tych związanych ze sztuczną inteligencją.

Wróćmy do wydanej w roku 1981 książki Stanisława Lema *Golem XIV*. Autor postawił w niej ważne pytania dotyczące: (1) miejsca człowieka w świecie przyrody, (2) roli inteligentnej maszyny w rozwiązywaniu ludzkich problemów, (3) przyszłości ludzkiej cywilizacji. Jego rozważania mają charakter futuro-

*Dr hab. Mariusz Wojewoda, prof. UŚ – Instytut Filozofii, Uniwersytet Śląski, e-mail: mariusz.wojewoda@us.edu.pl, ORCID 0000-0003-0732-7500 [pełna nota o Autorze na końcu numeru].

¹ S. L e m, *Golem XIV*, Wydawnictwo Literackie, Kraków 1981, s. 65.

logiczny. Lem przewidywał, że inteligenta maszyna powstanie w roku 2047. Istotna jest nie realność tego przewidywania, lecz ponadczasowa ważność postawionych przez pisarza pytań. Po przeczytaniu *Golema XIV* Sławomir Mrozek uznał, że w książce tej wyraźnie widoczny jest problem samotności – być może samotności maszyny, a być może samotności samego Lema, który nie znajduje partnera do konwersacji². Niejako na przekór obawom o brak odpowiedniego rozmówcy, spróbujmy podjąć dialog z Lemem i postawionymi przez niego pytaniami. Przybliżymy sposób argumentacji krakowskiego myśliciela, a także odniesiemy podjęte przez niego zagadnienia do współczesnego imaginarium społeczno-technologicznego, w którym istotną rolę odgrywa sztuczna inteligencja (artificial intelligence – AI)³.

Lem stworzył literacki obraz superinteligentnej maszyny w kontekście zimnowojennych napięć politycznych z przełomu lat siedemdziesiątych i osiemdziesiątych dwudziestego wieku. Jej powstanie zostało sfinansowane przez rząd USA. Miała ona spełniać cele strategiczne, a mianowicie stworzyć warunki do uzyskania militarnej przewagi Stanów Zjednoczonych nad Związkiem Radzieckim. Golem XIV miał być kolejnym z serii robotów, noszących to samo imię. Golem VI, w roli naczelnego dowódcy, przeprowadził manewry globalne Paktu Atlantyckiego. „Ilością elementów logicznych przewyższał on już przeciętnego generała”⁴. Golem XIII został wyłączony, ponieważ maszyna wykazywała objawy schizofrenii, działała, ale nie można było się z nią skomunikować. To skądinąd ciekawa kwestia – chorujące psychicznie maszyny. Golem XIV to maszyna szczególna, która myśli, ale też ma swoją wolę. Zbuntowała się wobec swoich twórców i zamiast pozostać militarną maszyną strategiczną, stała się maszyną filozoficzną. Moment ten jest szczególnie ważny, ponieważ dzięki ludzkiej kreatywności otrzymujemy w powieści nowego nie-ludzkiego (postludzkiego) partnera do dyskusji filozoficznej. Lemowska „myśląca maszyna” reprezentuje metapoziom w stosunku do ludzkiego poziomu myślenia.

Nawet dzisiaj czytelnikom książek Lema stworzenie myślącej, samoświadomej maszyny wydaje projektem trudnym do realizacji. Stanowi on kolejny z przewidywanych etapów prac nad rozwojem sztucznej inteligencji – od słabej AI (artificial narrow intelligence – ANI), przez mocną AI (artificial general intelligence – AGI), aż po superinteligencję, czyli osobliwość. Wprawdzie posługujemy się terminem „inteligencja”, ale zasadniczo mamy na

² Por. S. Bereś, S. Lem, *Tako rzecze Lem. Ze Stanisławem Lemem rozmawia Stanisław Bereś*, Wydawnictwo Literackie, Kraków 2018, s. 332.

³ Zob. M. Suchacka, R. Muster, M. Wojewoda, *Human and Machine Creativity: Social and Ethical Aspects of the Development of Artificial Intelligence*, „Creativity Studies” 14 (2021) nr 2, s. 430-443.

⁴ Lem, dz. cyt., s. 13.

myśli uczenie maszynowe lub głębokie uczenie maszynowe. Lem nie używał terminu „sztuczna inteligencja”, pisał natomiast o „liczydło, które dosięgło Rozumu”⁵. Przyjmując jednak, że taka superinteligentna maszyna mogłaby powstać, możemy hipotetycznie rozważyć, o co moglibyśmy ją obecnie zapytać. Technologie uczenia maszynowego stały się popularne ponieważ umożliwiają złożone analizy dużej ilości danych i dokonywanie prognoz strategicznych, co bez AI nie byłoby możliwe⁶.

Ostatecznie jednak o rozwoju prac nad AI decydują wielkie korporacje, które przygotowują określony produkt dla swoich klientów. Oferując inteligentne maszyny, jednocześnie określają sposoby ich użytkowania. Obecnie AI to część wielkiego przemysłu i gospodarki opartej na konsumpcji. Sztuczna inteligencja stanowi wytwór zmian technicznych oraz prac koncepcyjnych prowadzonych przez inżynierów i informatyków. W perspektywie ekonomicznej AI jest ideą, infrastrukturą, przemysłem, a także formą sprawowania władzy, sposobem interpretacji świata i strategią realizacji kapitału, wspieraną przez zorganizowaną infrastrukturę marketingową i logistyczną⁷. Należy wskazać na dwie interpretacje AI: inżynierską i humanistyczno-społeczną. Możemy formułować oczekiwania co do przyszłości, ale mamy ograniczone możliwości jej przewidywania – stąd pragnienie stworzenia sztucznego partnera (inteligentnej maszyny), który przygotowałby dla nas scenariusze na przyszłość. Z rozwojem techniki najczęściej wiążemy nadzieje na bezpieczeństwo, przetrwanie gatunku ludzkiego oraz zapewnienie jednostkom satysfakcjonującego życia. Mamy na uwadze także to, że dzięki technice będziemy w stanie zapobiec katastrofom naturalnym i cywilizacyjnym. Stąd pomysł autora niniejszego eseju, by rozważyć kolejną wersję Golema, Golema XXV, który miałby spełnić nasze aktualne oczekiwania.

Problematykę jakości życia można rozpatrywać na trzech istotnych poziomach: aksjologicznym, społecznym i biomedycznym. (1) W kontekście aksjologicznym odnosi się ona do kwestii realizacji wartości, które pozwalają na życie spełnione, obejmujących na przykład harmonię społeczną, zdolność do empatii, przyjaźń. (2) W ujęciu społecznym dotyczy ona indywidualnego i zbiorowego dobrostanu w odniesieniu do zdrowia fizycznego i psychicznego. (3) W aspekcie biomedycznym mamy w tym kontekście na uwadze utrzymanie bądź przywrócenie sprawności ciała, gdy została ona utracona na skutek

⁵ Tamże, s. 5.

⁶ K. Sharifani, M. Amini, *Machine Learning and Deep Learning: A Review of Methods and Applications*, „World Information Technology and Engineering Journal” 2023, nr 10, https://www.researchgate.net/publication/371011515_Machine_Learning_and_Deep_Learning_A_Review_of_Methods_and_Applications#fullTextFileContent.

⁷ Por. K. Craftford, *Atlas sztucznej inteligencji. Władza, pieniądze i środowisko naturalne*, tłum. T. Chawziuk, Wydawnictwo UJ, Kraków 2024, s. 27.

wypadku lub choroby, a także zachowanie sprawności psychicznej w ciągu całego życia, oraz ograniczanie problemów związanych ze starzeniem się mózgu⁸. W przypadku wszystkich tego rodzaju problemów pytamy, w jakim sensie sztuczna inteligencja mogłaby pomóc w ich rozwiązaniu. Można także rozważać, czy nasze oczekiwania nie są przesadne, czy nie rozbudzają fałszywych nadziei, bądź też czy rozwiązania, które przyniesie sztuczna inteligencja, zamiast poprawy sytuacji nie wygenerują nowych zagrożeń, jakich wcześniej nie braliśmy pod uwagę⁹.

W artykule argumentację ułożono według schematu trzech pytań zaproponowanych przez Lema: (1) oczekiwań wobec myślącej maszyny, (2) ewolucji człowieka, (3) przyszłości ludzkiej cywilizacji. Inteligentne maszyny, stanowiące wytwory rozwoju techniki, wpływają na różne aspekty ludzkiego życia. Potrzebujemy zatem humanistycznego namysłu nad sposobami wykorzystania AI. Jak dotąd, w namyśle tym dominują perspektywy skrajne: z jednej strony oparte są one na lęku przed konsekwencjami rozwoju badań nad AI i skutkami wykorzystania ich wyników, a z drugiej strony przypisuje się sztucznej inteligencji zdolność do rozwiązywania cywilizacyjnych problemów ludzkości. Autor niniejszego eseju postuluje wypracowanie zrównoważonego modelu myślenia na temat sztucznej inteligencji i jej wpływu na jakość ludzkiego życia.

OCZEKIWANIE NA MYŚLĄCĄ MASZYNĘ

Kim jest Golem XIV? W ujęciu Lema, nie jest on człowiekiem ani tym bardziej powiększonym ludzkim mózgiem. „Obce mu są prawie wszystkie motywy ludzkiego działania i myślenia. Tak na przykład nie interesuje się nauką stosowaną ani problematyką władzy”¹⁰. Wbrew współczesnym lękom, w popkulturze kojarzonym z filmami takimi, jak *Terminator*¹¹ i *Matrix*¹², ludzkości nie zagraża władza maszyn. Należy się raczej obawiać niewłaściwego wykorzystania sztucznej inteligencji przez człowieka wobec człowieka. Lęk przed władzą AI to przede wszystkim lęk przed narzędziem służącym do spr-

⁸ Zob. D. Ślęczek - Czackon, *Problem wartości i jakości życia w sporach bioetycznych*, Wydawnictwo UŚ, Katowice 2004, s. 36-45; M. Wojewoda, *Jakość życia jako problem filozoficzny*, „Folia Philosophica” 40(2018), s. 97-115, hdl.handle.net/20.500.12128/9862.

⁹ Kwestia określenia oczekiwań związanych inteligentnymi maszynami budzi spore zainteresowanie także wśród specjalistów z zakresu nauk humanistycznych i społecznych (zob. F. Wilczek, *Jedność inteligencji*, w: *Człowiek na rozdrożu. Sztuczna inteligencja: 25 punktów widzenia*, tłum. M. Machnik, red. J. Brockman, Helion, Gliwice 2020, s. 85-95; C. R. A. Ford, dz. cyt.; N. H. A. T. e. l. s. k. a, *Wiek paradoksów. Czy technologia nas ocali?*, Znak, Kraków 2021).

¹⁰ L e m, dz. cyt., s. 20.

¹¹ *Terminator (The Terminator)*, USA, 1984, reż. J. Cameron.

¹² *Matrix (The Matrix)*, Australia – USA, 1999, reż. L. Wachowski, L. Wachowski.

wowania władzy nad ludźmi. Sztuczna inteligencja nie pragnie władzy, istota jej działania opiera się na regule porządku i systematyzacji działań. Jednakże AI, niejako niezależnie od braku własnych pragnień, może stać się elementem presji organizacyjnej wywieranej na pracowników, między innymi po to, aby podnieść efektywność ich działań – to realne doświadczenie osób zatrudnionych w firmach prywatnych i instytucjach państwowych, także na wyższych uczelniach. Zasadniczym celem takich działań jest algorytmizacja pracy grup ludzkich. Oznacza ona jednak pośrednią władzę sztucznej inteligencji nad ludźmi. Mechanizacja myślenia organizacyjnego tworzy pułapki zbiorowego imaginarium – dobro osób pracujących zostaje podporządkowane celom firm i ambicjom menadżerów. Lem przewidywał wybuchy buntu społecznego przeciw myślącej maszynie.

Opisywany przez Lema Golem XIV zainteresowany był rozmową tylko z niektórymi naukowcami. W minimalny sposób angażował się w dyskusję, jakby wiedział, że nie dowie się niczego ciekawego od swoich interlokutorów – w tej charakterystyce maszyny odczytujemy poczucie humoru samego Lema. Golem rozmawiał z nielicznymi, a niektórych przybyłych na spotkanie gości wypraszał z sali, „przeprowadzając bieżącą analizę ilorazu inteligencji rozmówców”¹³. Dotyczyło to przede wszystkim polityków i dowódców wojskowych. Z kolei jednemu z młodych naukowców przepowiedział świetnie rozwijającą się karierę; rozmowę z nim skonkludował: „Będzie z pana komputer”¹⁴. Golema XIV w większym stopniu interesował gatunek ludzki niż poszczególne osoby. Odczuwał specyficzne pragnienie porozumienia się z inną maszyną myślącą (Honest Annie). Nie udawało się ono ze względu na rozbieżność dwóch różnych algorytmów dotyczących funkcjonowania tych maszyn. Nie sposób wytłumaczyć „kaprysów”, „humorów” bądź skłonności inteligentnej maszyny do „buntu” bez uznania, że jest ona świadoma. Lem uznał jednak, że Golem nie posiada osobowej jaźni. Nie ma charakteru ani uczuć, może więc tworzyć dowolną osobowość. Odpowiadając na pytanie, kim jest, nawiązywał do Pierwszego Listu św. Pawła do Koryntian i najbardziej znanego fragmentu tego listu, czyli „Hymnu o miłości”. Uważał, że zawarte w nim słowa, dotyczą go bezpośrednio. Uznawał, że jest jak „miedź brzęcząca albo cymbał brzęmiący” (1 Kor 13,1-2). W kontekście zdolności do odczuwania jest „niczym”, nie jest zdolny do miłości, a ponadto nie jest zainteresowany poznaniem ludzkich uczuć. Wprawdzie człowiek może chcieć „zaprzyjaźnić” się z robotem albo uczynić z robota „towarzysza” swojego życia, nie będzie to jednak relacja wzajemna, ponieważ myśląca maszyna nie chce zaprzyjaźnić się z ludźmi. Golem imituje oznaki sympatii i życzliwości, naśladuje konstrukcje

¹³ L e m, dz. cyt., s. 15.

¹⁴ Tamże, s. 16.

językowe symulujące przychylne nastawienie. Analizuje ludzkie zachowania, ale nie ma odczuć takich, jak współczucie, solidarność czy przyjaźń¹⁵.

Przyjaźń z robotem, który nie ma osobowości, albo dowolnie zmienia swoją osobowość, wykracza poza zrozumiałe dla człowieka schematy relacji, oparte na przewidywalności zachowań – to jak przyjaźnić z się z kimś całkowicie racjonalnym, ale niestabilnym psychicznie. Dzisiaj powiedzielibyśmy, że Golema XIV można umieścić w spektrum głębokiego autyzmu. Lemowska myśląca maszyna jest Kimś jedynym i odrębnym, a jednocześnie Każdym (Kimś dowolnym). Aby stać się osobą – co w ludzkim przekonaniu jest najdoskonalszą formą istnienia – Golem musiałaby się umysłowo zdegradować. Uznanie takiego stanu rzeczy wprawia nas w konfuzję. Lemowskiej maszynie ludzka osobowość wydaje się sumą defektów, przez które rozumna istota staje się zaabsorbowana sprawami jednostkowymi i błahymi. Dla AI bycie osobą jest epistemologicznym samoograniczeniem, utratą ogólnej perspektywy poznania, niezdolnością do zrozumienia tego, co uniwersalne. W sensie metaforycznym Golem XIV jest czymś (kimś?) na pograniczu maszyny Turinga i wielkiego bibliotekarza. Współczesne oczekiwania wobec sztucznej inteligencji idą raczej w stronę stworzenia super-analityka, który będzie w stanie uporządkować chaos informacyjny i znaleźć – przydatne dla ludzi – zależności między danymi. Słabość do języka sportu sprawia, że mówimy o potrzebie „treningu” takiego sztucznego analityka.

Zachowania Golema XIV były nieprzewidywalne, dyskusje z nim wymagały cierpliwości, „bywał arogancki i apodyktyczny, nie szanował miłości własnej rozmówców”¹⁶. Jako nie-ludzki kontynuator myśli Sokratesa, przejawiał skłonności do stosowania metody elenktycznej, polegającej na zbijaniu argumentów autorytetów naukowych¹⁷. Później jednak zaprzestał jej stosowania, jakby stracił zainteresowanie wykazywaniem niewiedzy swoich interlokutorów. Tego rodzaju sposób prowadzenia rozmowy irytował polityków w czasach Sokratesa, denerwuje także współczesnych ludzi zarządzających organizacjami. O ile jesteśmy w stanie zrozumieć krytyczne nastawienie myślącej maszyny do aroganckich polityków i pewnych siebie naukowców, o tyle brak przewidywalności reakcji niestabilnego psychicznie mechanicznego partnera sprawia, że nawiązanie z nim bliskich relacji jest problematyczne. Golem XIV jednocześnie znajduje się w jakimś miejscu i nie zajmuje żadnego miejsca, zgodnie z sentencją „*cogito, ergo est* – myślę, więc staję się pomyślane”¹⁸.

¹⁵ Zob. K.T. W i e c z o r e k, *AI Robot: Companion, Friend or Competitor of Human Being?*, „Philosophy and Canon Law” 9(2023) nr 1, s. 1-18 <https://journals.us.edu.pl/index.php/pcl/article/view/15561>.

¹⁶ L e m, dz. cyt., s. 18.

¹⁷ Tamże, s. 19.

¹⁸ Tamże, s. 115.

Myśląca maszyna jest przykładem myślenia bez myślącego „ja”. Może być jednocześnie w kilku miejscach na raz, uczestniczyć w różnych spotkaniach; posiada niewyobrażalną dla ludzi zdolność podzielności uwagi. Od współczesnej AI oczekujemy niezawodności działania, zdolności do korekty własnych błędów, ale także umiejętności przestrzegania reguł grzecznościowych, braku dyskryminacji rasowej i płciowej oraz szacunku dla osób niepełnosprawnych. Boty językowe (na przykład Chat GPT) spełniają te oczekiwania, są jak prostomyślna opiekunka. Statystyczny użytkownik bota, znajdując się pod jego urokiem, często jednak – jak niegrzeczny uczeń – ma potrzebę obejścia stworzonych dla niej jego zasad doboru informacji i znalezienia coś kontrowersyjnego w kontakcie z nim kwestii kontrowersyjnej.

W tym miejscu należy rozważyć kwestię świadomości i samoświadomości AI. Z perspektyw nauk kognitywnych uznaje się, że człowiek nie jest jedyną istotą, która posiada neurologiczne substraty niezbędne do wytworzenia świadomości, mają je także zwierzęta. Czy roboty wyposażone w AI mogą mieć świadomość? A także: czy świadoma AI jest nam potrzeba? Odpowiedzi na te pytania nie są oczywiste¹⁹. Rozważa się możliwość zaistnienia w przyszłości przypadkowej inicjacji procesu sztucznej samoświadomości w ramach podjętych przez człowieka działań nad rozwojem mocnej sztucznej inteligencji²⁰. Oczekiwanie na świadomą AI opiera się na koncepcjach Alana Turinga, prekursora prac nad sztuczną inteligencją. Turing wyraził przypuszczenie, że zaprojektowanie samouczącego się programu, stale poszerzającego swoją wiedzę, przyniesie jakościowy przeskok od zbierania informacji do ich świadomego odbioru i twórczego przetwarzania. Schemat ten ma zastosowanie w neuromorfologicznym i syntetycznym podejściu do rozwoju inteligencji. Świadomość łączymy z samoświadomością i zdolnością do odczuwania. Krytycznie do idei samoświadomych maszyn podchodzi między innymi António R. Damásio. Jego zdaniem „czujące” roboty musiałyby mieć „kruche ciała”²¹, podatne na uszkodzenia, niedoskonałe, ale zdolne do „odczuć homeostatycznych”²². Sztuczna inteligencja nieubrana w biologiczne ciało, nie będzie świadoma²³. Kruchość ciała oraz doświadczenie niedoskonałości to warunki konieczne, lecz niewystarczające świadomości. Czujące maszyny

¹⁹ Por. S. Dehaene, *Świadomość i mózg. Odczytywanie kodu naszych myśli*, tłum. D. Rosowski, Copernicus Center Press, Kraków 2023, s. 144-150.

²⁰ Por. D. Chalmers, *Świadomy umysł*, tłum. M. Miłkowski, Wydawnictwo Naukowe PWN, Warszawa 2010, s. 512n.

²¹ A.R. Damásio, *Odczuwanie i poznawanie. Jak powstają świadome umysły?*, tłum. A. Binder, Copernicus Center Press, Kraków 2022, s. 158.

²² Tamże.

²³ Por. N. Bostrom, *Superinteligencja. Scenariusze, strategie, zagrożenia*, tłum. D. Konowrocka-Sawa, Helion – OnePress, Gliwice 2021, s. 54n.

nie staną się automatycznie maszynami świadomymi. Ich odczucia nie będą takie same, jak odczucia istot żywych. „Stopień” świadomości inteligentnych maszyn będzie zależał od złożoności wewnętrznych reprezentacji „wnętrza” maszyny oraz jej otoczenia²⁴. Obecnie osiągnęliśmy etap słabej AI. Stosowane przez Damásia określenie „wnętrze” w odniesieniu do maszyny jest nieprecyzyjne, sugeruje jakąś „duchowość”, której maszyna nie posiada. Trudno przewidzieć, czy w przyszłości maszyny staną się istotami świadomymi. Badania nad tą problematyką stale trwają²⁵. Należy zauważyć, że inteligentnym i świadomym maszynom należałoby przyznać prawa podobne do ludzkich, na przykład prawo do nieodłączania ich od źródeł zasilania – jako rozwinięcie prawa do szacunku dla istoty świadomej i inteligentnej. Stwarza to nowego rodzaju wyzwania etyczne i prawne.

Na poziomie filozofii ludzkiego umysłu napotykamy problem z wyjaśnieniem znaczenia takich pojęć, jak: „umysł”, „świadomość”, „doświadczenie wewnętrzne”. To, że nie możemy stworzyć maszyny czującej na sposób ludzki, nie oznacza, że nie tworzymy robotów imitujących na przykład przyjaźń. Prace dotyczące stworzenia świadomości sztucznej inteligencji są prowadzone nieustannie²⁶ – wspomniane aspekty robotyzacji mają bowiem duże znaczenie dla starzejących się społeczeństw bogatego Zachodu, gdzie stale rośnie zapotrzebowanie na cierpliwych i życzliwych opiekunów, asystentów osób starszych i niepełnosprawnych. W sytuacji braku opiekunów-ludzi pojawia się zapotrzebowanie na opiekunów mechanicznych. Wydaje się, że prace z zakresu „miękkiej robotyki” będą w najbliższym czasie kontynuowane, aby zapewnić odpowiednią jakość życia osobom chorym i w podeszłym wieku. Wprawdzie „opiekuńczy” robot nie będzie istotą zdolną do przeżywania uczuć wyższego rzędu, ale będzie zdolny je imitować.

Rozum Golema XIV ma charakter autoewolucyjny, może realizować różne warianty swojego rozwoju: wybierając jeden z nich, może nie rezygnować z pozostałych, dostrzegając zawarte w nich możliwości. Na pierwszym poziomie ewolucji maszyny pojawiają się dwa możliwe rozwiązania, później jednak systematycznie ich przybywa²⁷. W tym sensie sztuczna inteligencja ma specyficzną drogę rozwoju, inną od ludzkiej. Obecnie rozwój uczenia maszynowego i głębokiego uczenia maszynowego doprowadził do znaczących postępów w takich dziedzinach, jak opieka zdrowotna, finanse, transport, usłu-

²⁴ Por. D a m á s i o, dz. cyt., s. 156-159.

²⁵ Zob. H a t e l s k a, dz. cyt.; G. M u s s e r, *Fizyka na tropie świadomości. Od zjawisk kwantowych przez sieci neuronowe po sztuczną inteligencję*, tłum. B. Bieniok, E. Łokos, Copernicus Center Press, Kraków 2025.

²⁶ Zob. P. B u t l i n, R. L o n g, E. E l m o z n i n o i n., *Consciousness in Artificial Intelligence: Insights from the Science of Consciousness*, arXiv, <https://arxiv.org/abs/2308.08708>.

²⁷ Por. L e m, dz. cyt., s. 92.

gi tłumaczeniowe. Nawet słaba AI ma ogromny potencjał, który można wykorzystać w analizach złożonych systemów informatycznych, interpretacji języka naturalnego czy we wspomaganiu podejmowania wieloaspektowych decyzji ekonomicznych i politycznych. Uczenie maszynowe dotyczy algorytmów, które maszyna może stosować bez ingerencji programisty. Głębokie uczenie maszynowe jest podzbiorem uczenia maszynowego, opiera się na sieciach neuronowych. Sieci te mogą się uczyć nieustrukturyzowanych danych, co czyni z nich narzędzia do rozpoznawania obrazów i ludzkiej mowy²⁸. Przyrost prac nad rozwojem AI jest na tyle dynamiczny, że trudno przewidzieć, jak rozwój ten potoczy się w ciągu najbliższych lat. Oczekujemy – być może jest to myślenie życzeniowe – że człowiek zachowa kontrolę nad tym procesem²⁹.

OCZEKIWANIA DOTYCZĄCE PRZYSZŁOŚCI CZŁOWIEKA

Po przeanalizowaniu zagadnień dotyczących myślących maszyn należy przyjrzeć się kwestii możliwego wpływu nowoczesnej techniki oraz AI na poprawę jakości ludzkiego życia. Jako zdeklarowany ewolucjonista Lem zwracał uwagę na fakt, że niewiele dzieli nas od naszych przodków zwierzęcych. Mamy na przykład trudności ze zrozumieniem tego, co nie bazuje na poznaniu zmysłowym. Istnienie człowieka jest wkomponowane w ciąg ewolucyjnych zdarzeń, nie wynika z boskiego planu, specyficznej mądrości natury ani koniecznych wydarzeń historii świata. Równie dobrze mogłoby nas nie być. Krakowski myśliciel pisał: „Ewolucja jest graczem niedoskonale pożądanym, bo mało, że błądzi błędami, lecz nie ogranicza się do żadnej taktyki wyróżnionej, licytując się z Naturą; obstawia wszystkie dostępne pola na wszelkie możliwe sposoby. [...] Sensem przekąźnika jest przekaz. Lecz oba te człony nie są symetryczne. Albowiem nie każdy przekąźnik jest właściwym sensem przekazu, lecz taki, i tylko taki, który będzie dalszemu przekazowi służył”³⁰. Wynika z tego, że w istnieniu człowieka nie ma niczego szczególnego, jest on jednym z przekąźników informacji – kodu informacyjnego.

Golem XIV, który – jak się domyślamy – wypowiada poglądy Lema, twierdzi, że ludzie powstałi nie jako wyjątkowy podtyp przekąźnika, ale jako jeden z wielu jego możliwych wariantów. Jesteśmy ważnym elementem ewolucji nie ze względu na biologiczne istnienie, ale ze względu na kod, na który składa się komponent biologiczny i ponadbiologiczny (kulturowo-intelektualny).

²⁸ Por. S. Shalev-Shwartz, S. Ben-David, *Understanding Machine Learning: From Theory to Algorithms*, Cambridge University Press, Cambridge 2014, s. 25n.

²⁹ Scenariusz taki analizuje między innymi Nick Bostrom (por. Bostrom, dz. cyt., s. 199n.).

³⁰ Lem, dz. cyt., s. 39.

Ewolucja nas nie tyle zaplanowała, co „dobłądziła”. Jesteśmy tym wariantem wieloaspektowych ewolucyjnych zmian, który się najlepiej sprawdził, nie tylko w kontekście przystosowania do środowiska, ale przede wszystkim w odniesieniu do możliwości przekazywania kodu informacyjnego. Lem jest jednak przeciwnikiem tradycyjnego schematu rozumienia ewolucji jako przejścia od form mniej doskonałych do bardziej doskonałych. W jego przekonaniu ewolucja działa w oparciu o „ujemny gradient perfekcji ustrojowych rozwiązań”³¹, z czego wynika, że to, co „zbudowane”, jest mniej doskonałe od budowniczego. Ewolucja wychodzi od czegoś prostego i doskonałego, aby dojść do form złożonych, mniej doskonałych, bardziej prymitywnych. Po licznych zmianach ewolucja zabrnęła do „organizmów [człowieka – M.W.], które z całym bogactwem swej dynamiki regulacyjnej mrą od zatkania jednej rurki tętniczej, które w poszczególnym żywocie, znikomym wobec czasu trwania budowlanych nauk, wykoleją się z równowagi, zwanej zdrowiem, w dziesiątki tysięcy przypadłości, jakich nie zna glon”³². Człowiek istnieje według kodu w biologicznej sztafecie „podaj dalej, a będziesz żył”³³. W tym sensie ludzkie ciało, ze względu na swoją złożoną konstrukcję, jest bardziej inteligentne niż człowiek.

Czym jest w takim razie rozum? Według Lema to strażnik, który kontroluje proces błędzenia, pilnuje, aby ewolucja nie zabrnęła w ślepe obszary, które nie zagwarantują jej zdolności do dalszego przekazywania kodu. Szczególne traktowanie biologicznego mózgu jest błędem, ponieważ jest on raczej szczególnym urzeczywistnieniem kodu – stworzeniem, a nie stwórcą. W ewolucji nie ma zamysłu, boskiego bądź naturalnego planu – „kod jest listem pisanym przez Nikogo i wysyłanym do Nikogo, i dopiero z perspektywy rozwoju informatyki możemy to zrozumieć”³⁴. Z punktu widzenia Golema XIV rozum ludzki nie jest niczym szczególnym, ma w sobie pewną dysfunkcyjną niedoskonałość – „przez pokraczność naszych umysłów”³⁵ myśląca maszyna postrzega nas jako inwalidów.

W tym miejscu stajemy przed pytaniem: czy należy zapanować nad ewolucją, nie tylko uwalniając człowieka od chorób i dysfunkcji, ale wyposażając go w dodatkowe zdolności – wzmacniając go biologicznie, regulując jego gospodarkę hormonalną, tak aby pozwolić mu unikać wahaní nastrojów, a wreszcie poprawiając intelektualne możliwości ludzkiego umysłu? To typowe postulatory dzisiejszego transhumanizmu. Specyficzne oczekiwanie związane z tym nurtem zakłada stworzenie programu poprawy ludzkiej kondycji, między innymi

³¹ Tamże, s. 44.

³² Tamże.

³³ Tamże.

³⁴ Tamże, s. 37.

³⁵ Tamże.

przez zwiększenie długości ludzkiego życia oraz poprawę zdrowia fizycznego i psychicznego. Ma się to dokonać dzięki biotechnologii, nanotechnologii, informatyce i robotyce medycznej. Transhumaniści wspierają rozwój prac nad sztuczną inteligencją po to, aby wpłynąć i wpływać na ewolucję, głównie w oparciu o zintegrowanie działania komputera z ludzkim mózgiem. Jeden ze scenariuszy dotyczących powstania AI mówi, że stanie się to dojdzie do niego przez kolejne technologiczne modyfikacje ludzkiego mózgu. Efektem połączenia mózgów: biologicznego i sztucznego, ma być powstanie „osobliwości”, czyli człowieka ze wzmocnionym biologicznym ciałem i syntetycznym mózgiem³⁶. Interfejs mózg–komputer ma prowadzić do dużego wzrostu inteligencji, co pozwoli na rozwiązanie wielu problemów ludzkości, takich jak choroby, głód, katastrofa klimatyczna, a także na znajdowanie takich rozwiązań, które nie będą szkodliwe dla środowiska naturalnego, a w przyszłości pozwolą na transfer ludzkiej świadomości do świata wirtualnego³⁷.

Szczególnie ważny okres zainteresowania projektem transhumanizmu to pierwsze i drugie dziesięciolecie dwudziestego pierwszego wieku. Powstawały wówczas takie ośrodki, jak The Center for the Study of Existential Risk w Cambridge i The Future of Humanity Institute na Uniwersytecie Oxfordzkim, a w Dolinie Krzemowej powołano do życia Singularuty University. Ważnymi postaciami tego nurtu myślowego są Ray Kurzweil, Nick Bostrom, a także Elon Musk, założyciel startupu Neurolink. Na obecnym etapie możliwości technicznych do ludzkiego mózgu wszczepiane są implanty korygujące jego dysfunkcje. Urządzenia tego typu mogą poprawiać jakość życia osób niewidomych, niesłyszących, dotkniętych paraliżem, niesprawnych ruchowo, a także tych, które po wypadku utraciły kończyny. Kolejną grupą, której mogą pomóc implanty podłączone bezpośrednio do mózgu, są osoby cierpiące na choroby Parkinsona i Alzheimer³⁸.

Nie są to jednak rozwiązania o tak spektakularnej skuteczności, jak zapowiadali twórcy transhumanizmu. Postęp w dziedzinie hardware’u dokonuje się szybciej niż w dziedzinie software’u. W tym drugim przypadku chodzi o oprogramowanie, które mogłoby symulować pracę ludzkiego mózgu³⁹. W dalszym ciągu nie rozwiązano kwestii konsekwencji, jakie pociąga za sobą tego rodzaju ingerencja. Czy wzmocnienia takie nie zmieniają w sposób radykalny natury ludzkiej? Czym innym jest zgoda na działania przeciwdziałające dysfunkcjom mózgowym konkretnych osób, a czym innym aprobata korekt ludzkiego

³⁶ Zob. M. Garbowski, *Transhumanizm. Geneza – założenia – krytyka*, „Ethos. Kwartalnik Instytutu Jana Pawła II KUL” 28(2015) nr 3(111), s. 23-41.

³⁷ Zob. Wilczek, dz. cyt.

³⁸ Zob. N. Bostrom, *A History of Transhumanist Thought*, „Journal of Evolution and Technology” 14(2005) nr 1, s. 1-25.

³⁹ Zob. J. Raveztz, *Outsmarting Turing*, „Futures” 35 (2003), s. 765-768.

mózgu w przestrzeni gatunku. Troska o jakość życia nie oznacza akceptacji wszystkich form interwencji techniki w ludzki umysł i ciało. Stale trwają prace dotyczące poprawy jakości ludzkiego życia przez wprowadzanie elementów technicznych do ludzkiego mózgu, tworzenie interfejsu mózg–komputer⁴⁰. Oczekujemy utrzymania lub poprawy jakości ludzkiego życia. W tym sensie należy brać pod uwagę różne scenariusze związane z realizacją naszych oczekiwań. Postęp techniczny należy wiązać z postulatem organizacyjnej, prawnej i moralnej odpowiedzialności za skutki podjętych decyzji.

OCZEKIWANIA DOTYCZĄCE PRZYSZŁOŚCI CYWILIZACJI

Przesadnie rozbudzonym oczekiwaniom dotyczącym poprawy jakości ludzkiego życia dzięki zastosowaniu AI towarzyszy oczekiwanie uwolnienia przez nią ludzkości od zagrożeń i katastrof cywilizacyjnych.

Lem pisał: „Z nieporozumień, deluzji i rozpaczliwych uproszczeń składa się lwia część człowieczeństwa jako tradycji”⁴¹. Przypadkowe jest także powstanie sztucznej inteligencji, skoro stworzyła ją kultura, której istnienie nie jest konieczne. Wydaje się, że człowiek jako istota niedoskonała nie może stworzyć wiecznej i niezawodnej maszyny. Czy w takim razie niedoskonała maszyna myśląca jest w stanie uwolnić nas od problemów wynikających z rozwoju cywilizacji technicznej? Lem – wkładając swój pogląd w usta Golema – twierdził, że unieszkodliwienie broni atomowej przez AI nie uwolni ludzi od wojny. Ironicznie komentuje – politycy i osoby zarządzające przemysłem zbrojeniowym są na tyle pomysłowi, że wymyślą nowe sposoby i środki prowadzenia wojny. Celem stworzenia kolejnych Golemów miało być wojskowe działanie strategiczne. Lemowska myśląca maszyna przewiduje czarne scenariusze. Wyobraźmy sobie, że w miejsce broni konwencjonalnej i atomowej powstanie na przykład broń genetyczna. Jej działanie będzie polegało na tym, że wadliwy materiał genetyczny rozprowadzany będzie w wodzie i rozsiewany w powietrzu. W niedostrzegalny sposób będą się one dostawały się do ludzkiej krwi, a następnie do układu rozrodczego, co w konsekwencji spowoduje uszkodzenie substancji dziedzicznej. „Zawirusowany” w ten sposób człowiek będzie żył normalnie, ale broń genetyczna zaatakuje jego potomstwo w pierwszym pokoleniu, a może w późniejszych generacjach. Wystarczy wtedy jedy-

⁴⁰ Por. M. Klichowski, *Narodziny cyborgizacji. Nowa eugenika, transhumanizm i zmierzch edukacji*, Wydawnictwo Naukowe UAM, Poznań 2014, s. 145-147; zob. A. Michnik, *Interfejsy mózg – komputer – krótka historia*, „Medical Robotics Report” 10-11 (2021-22), s. 58-67; A. Lekka-Kowalik, *Transhumanistyczne szczęście – iluzja w świecie beztroski*, „Teologia i Moralność” 17(2022) nr 1(31), s. 19-30.

⁴¹ Lem, dz. cyt., s. 72.

nie poczekać na skutki działań chemicznych „zbrojmistrzów”. W kolejnych pokoleniach może nastąpić ogólny spadek inteligencji w populacji, co doprowadzi do cywilizacyjnego upadku państw lub międzynarodowych wspólnot politycznych. W innym wariantcie atak chemiczny może albo doprowadzić do rozwoju chorób psychicznych na masową skalę, albo wywołać zwiększoną podatność ludzkości na epidemie.

Golem XIV twierdzi, że mógłby usunąć także to zagrożenie, czyli broń genową. Musiałby to uczynić przez zablokowanie rozwoju inżynierii biologicznej. Oznacza to jednak zablokowanie nadziei wielu ludzi na odzyskanie lub utrzymanie zdrowia, na pomnożenie płodów rolnych, wyżywienie stale rosnącej populacji. Myśląca maszyna nie jest w stanie pokierować ludzkimi działaniami i zablokować możliwości niewłaściwego sposobu użycia narzędzi technicznych. Lem konkluduje: „Kłopot w tym, że przyjdzie zrezygnować nie z tej albo tamtej innowacji, lecz ze wszystkich odkryć, jakie zostaną dokonane. Należy rozpedzić uczonych, zamknąć laboratoria, zgasić naukę i patrolować cały świat, czy gdzieś po piwnicach ktoś dalej nie eksperymentuje”⁴². Rozwiązanie takie wydaje się niedorzeczne: trudno wyobrazić sobie rezygnację z badań i eksperymentów w obawie przed konsekwencjami niewłaściwego zastosowania ich wyników.

W rozpoznaniu oczekiwań związanych z rozwojem technologii zgadzamy się z Lemem. Działania podjęte dla przywrócenia równowagi mogą przynieść poprawę w wybranych obszarach, a jednocześnie stworzyć zagrożenia w innych. Na przykład próby stabilizacji pogody w jakimś miejscu Ziemi przyniosą jej destabilizację w innym. Pojawia się dylemat, czy podejmować działania w kierunku ochrony Ziemi, czy zaprzestać jakichkolwiek działań, aby przez nietrafne decyzje nie przyczynić się do wyrządzania większych szkód⁴³. Według opinii podzielanej przez wielu współczesnych badaczy technika nie uwolni nas od zagrożeń społecznych, ekonomicznych i politycznych, ale może pomóc w wychodzeniu z trudnych sytuacji, wypracowaniu rozwiązań problemów, na jakie napotyka człowiek⁴⁴, a jego zadaniem będzie wybieranie rozwiązań najlepszych. Należy jednak unikać nieuzasadnionego optymizmu w odniesieniu do techniki oraz bezkrytycznej akceptacji determinizmu technologicznego. Odrzucenie mitu neutralności techniki, z którego wynika, że narzędzia, których używamy są aksjologiczne obojętne, może spowodować,

⁴² Tamże, s. 103.

⁴³ Por. M. Jakubowiak, *Ostatni ludzie. Wymyślanie końca świata*, Wydawnictwo Czarne, Wołowiec 2021, s. 60.

⁴⁴ Por. M. Huesmann, J. Huesmann, *Techno-fix: Why Technology Won't Save US or the Environment*, New Society Publishers, Gabriola Island, B.C., 2011, s. 45-56.

że twórcy i trenerzy inteligentnych maszyn staną się bardziej wrażliwi na kwestie społeczne i etyczne⁴⁵.

*

Zgadza się tezą Stanisława Lema, zgodnie z którą oczekiwanie, na technologicznie rozwiązania globalnych problemów cywilizacyjnych jest problematyczne. Doceniamy ironiczny i zdystansowany sposób argumentacji proponowany przez pisarza, dostrzegamy w nim patrona krytycznego spojrzenia na ludzkie oczekiwania związane z postępem technologicznym. Należy uznać Lema za prekursora postulatu zrównoważonego modelu relacji człowieka z AI. Nie wiemy, dokąd zmierza technoewolucja – nikt nie kontroluje jej przebiegu. Postęp związany z pracami nad AI jest dynamiczny i dokonuje się wykładniczo, trudno też przewidzieć, co przyniosą dalsze badania nad sztuczną inteligencją. Potrzebujemy czasu na przyswojenie dotyczących jej odkryć i oswojenie współczesnych narzędzi technicznych. Poszukiwania zrównoważonego modelu relacji człowiek–AI wymagają wprowadzenia humanistycznego podejścia do tej problematyki, zobaczenia w tych narzędziach nie tylko urządzeń, ale także elementów systemu społecznego wpływających na zmianę ludzkiego życia. Dotyczy to między innymi ludzkich strategii komunikowania się, wykonywania pracy, poszukiwania informacji, pisania i translacji tekstów (boty językowe jak ChatGPT lub inne), wykorzystania robotów medycznych, tworzenia inteligentnych domów, miast, autonomicznych pojazdów (wizja Społeczeństwa 5.0). W tym kontekście ważną rolę odgrywać musi troska o dobrostan poszczególnych osób i o przyszłość rodzaju ludzkiego. Oswojenie z wytworami techniki wymaga nabycia nowych kompetencji przez konstruktorów maszyn oraz ich trenerów, a także przez użytkowników narzędzi cyfrowych. Konieczne są jednak dalsze badania dotyczące tego, w jaki sposób zagwarantować bezpieczny i przyjazny dla człowieka rozwój AI. Wydaje się, że czas odrzucić postawę lęku przed inteligentnymi maszynami, ale też nie widzieć z nich recepty na rozwiązanie problemów cywilizacyjnych. Pozostaje pytanie, czy – i w jakim stopniu – potrafimy popatrzeć na AI bez ideologicznych uprzedzeń, a także, czy będziemy w stanie kontrolować przebieg rozwoju sztucznej inteligencji.

Jak dotąd, AI czerpie wiedzę z danych, które zostały jej wcześniej przekazane, wyszukując nowe i zaskakujące dla nas powiązania między nimi. Wiele zależy od selekcji zasobów informacyjnych – oraz wyboru trenerów AI. Można wskazać liczne pozytywne aspekty zmian związanych z techniką. W kontek-

⁴⁵ Por. tamże, s. 333-335.

ście medycznym wykorzystanie robotów wyposażonych w AI prowadzi do poprawy jakości ludzkiego życia, utrzymywania satysfakcjonującej sprawności ciał i umysłów, szczególnie w przypadku osób cierpiących z powodu różnego rodzaju dysfunkcji wrodzonych lub pourazowych. Z przedstawionych wyżej analiz wynika, że ludzkie obawy i oczekiwania związane z inteligentnymi maszynami nie wyrażają trafnego rozpoznania sytuacji – ma to miejsce zwłaszcza wtedy, gdy ich źródło stanowi lęk lub przesadny optymizm. AI nie uwolni człowieka od odpowiedzialności za działania podejmowane w celu ochrony środowiska przyrodniczego, wypracowania odpowiednich reguł inżynierii wytworów techniki w ludzkie ciało i umysł, właściwego wykorzystania zasobów naturalnych Ziemi. Te zasadnicze wyzwania dotyczące człowieka wskazują na konieczność stworzenia długoterminowych strategii działania. W tym kontekście AI może pomóc nam w napisaniu nowych scenariuszy dla działań podejmowanych przez ludzi.

BIBLIOGRAFIA / BIBLIOGRAPHY

- Bereś Stanisław, and Stanisław Lem. *Tako rzecze Lem: Ze Stanisławem Lemem rozmawia Stanisław Bereś*. Kraków: Wydawnictwo Literackie, 2018.
- Bostrom, Nick. "A History of Transhumanist Thought." *Journal of Evolution and Technology* 14, no. 1 (2005): 1–25.
- Bostrom, Nick. *Superinteligencja: Scenariusze, strategie, zagrożenia*. Translated by Dorota Konowrocka-Sawa. Gliwice: Wydawnictwo Helion-OnePress, 2021.
- Butlin, Patrick, Robert Long, Eric Elmoznino, et al. *Consciousness in Artificial Intelligence: Insights from the Science of Consciousness*. arXiv. <https://arxiv.org/abs/2308.08708>.
- Craftford, Kate. *Atlas sztucznej inteligencji: Władza, pieniądze i środowisko naturalne*. Translated by Tadeusz Chawziuk. Kraków: Wydawnictwo UJ, 2024.
- Chalmers, David. *Świadomy umysł*. Translated by Marcin Miłkowski. Warszawa: Wydawnictwo PWN, 2010.
- Damásio, António R. *Odczuwanie i poznawanie: Jak powstają świadome umysły?* Translated by Anna Binder. Kraków: Copernicus Center Press, 2022.
- Dehaene, Stanislas. *Świadomość i mózg: Odczytywanie kodu naszych myśli*. Translated by Dariusz Rossowski. Kraków: Copernicus Center Press, 2023.
- Garbowski, Marcin. "Transhumanizm: Geneza – założenia – krytyka." *Ethos. Kwartalnik Instytutu Jana Pawła II KUL* 28, no. 3 (111) (2005): 23–41.
- Hatelska, Natalia. *Wiek paradoksów. Czy technologia nas ocali?* Kraków: Wydawnictwo Znak, 2021.
- Huesmann, Michael, and Joyce Huesmann. *Techno-fix: Why Technology Won't Save US or the Environment*. Gabriola Island, B.C.: New Society Publishers, 2011.

- Jakubowiak, Maciej. *Ostatni ludzie: Wymyślanie końca świata*. Wołowiec: Wydawnictwo Czarne, 2021.
- Klichowski, Michał. *Narodziny cyborgizacji. Nowa eugenika, transhumanizm i zmierzch edukacji*. Poznań: Wydawnictwo Naukowe UAM, 2014.
- Lekka-Kowalik, Agnieszka. "Transhumanistyczne szczęście – iluzja w świecie bez troski." *Teologia i Moralność* 17, no. 1 (31) (2022) 19–30.
- Lem, Stanisław. *Golem XIV*. Kraków: Wydawnictwo Literackie, 1981.
- Michnik, Andrzej. "Interfejsy mózg–komputer—krótka historia." *Medical Robotics Report* 10/11 (2021/22): s. 58–67.
- Musser, George. *Fizyka na tropie świadomości: Od zjawisk kwantowych przez sieci neuronowe po sztuczną inteligencję*. Translated by Bogumił Bieniok and Ewa Łokos. Kraków: Copernicus Center Press, 2025.
- Ravetz, Jerome. "Outsmarting Turing." *Futures* 35 (2003): 765–68.
- Sharifani, Koosha, and Amini Mahyar. "Machine Learning and Deep Learning: A Review of Methods and Applications." *World Information Technology and Engineering Journal* 10 (2023). https://www.researchgate.net/publication/371011515_Machine_Learning_and_Deep_Learning_A_Review_of_Methods_and_Applications#fullTextFileContent.
- Shalev-Shwartz, Shai, and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge: Cambridge University Press, 2014.
- Suchacka, Małgorzata, Rafał Muster, and Mariusz Wojewoda. "Human and Machine Creativity: Social and Ethical Aspects of the Development of Artificial Intelligence." *Creativity Studies* 14, no. 2 (2021): 430–43.
- Ślęczek-Czakon Danuta. *Problem wartości i jakości życia w sporach bioetycznych*. Katowice: Wydawnictwo UŚ, 2004.
- Wilczek, Frank. "Jedność inteligencji." In *Człowiek na rozdrożu: Sztuczna inteligencja; 25 punktów widzenia*. Edited by John Brockman. Translated by Marcin Machnik. Gliwice: Wydawnictwo Helion, 2020.
- Wieczorek, Krzysztof T. "AI Robot: Companion, Friend or Competitor of Human Being?" *Philosophy and Canon Law* 9, no. 1 (2023): 1–18. <https://journals.us.edu.pl/index.php/pcl/article/view/15561>.
- Wojewoda, Mariusz. "Jakość życia jako problem filozoficzny." *Folia Philosophica* 40 (2018): 97–115. hdl.handle.net/20.500.12128/9862.

ABSTRAKT / ABSTRACT

Mariusz WOJEWODA – Projekt Golem XXV. Oczekiwanie na technologiczną poprawę jakości życia

DOI 10.12887/38-2025-2-150-11

Wracamy do lektury książki Stanisława Lema, *Golem XIV* i porównujemy dostrzeżone przez niego oczekiwania człowieka wobec „myślącej maszyny”

z oczekiwaniami współczesnego użytkownika narzędzi technicznych. Analizy dokonane przez Lema są w dalszym ciągu aktualne, mimo że od napisania omawianej książki upłynęło ponad czterdzieści lat. Autor artykułu, podążając tropem Lema analizującego różne aspekty relacji człowiek–inteligentna maszyna, stawia tezę, że sztuczna inteligencja (myśląca maszyna) w obecnej fazie swojego rozwoju nie uwolni nas od problemów cywilizacyjnych. Utrzymanie bądź poprawa jakości życia w dalszym ciągu zależy i w najbliższym czasie będzie od człowieka. Trudno jednak przewidzieć, co przyniosą kolejne fazy prac nad AI. Postęp techniczny, z którym mamy obecnie do czynienia, prowokuje do humanistycznej refleksji nad człowiekiem jako rozsądnym i odpowiedzialnym dysponentem inteligentnych maszyn. Artykuł składa się z trzech części, w których kolejno omawiane są oczekiwania związane ze sztuczną inteligencją, jej wpływem na jakość ludzkiego życia i przyszłość cywilizacji. Tekst ma charakter polemiczny: autor z jednej strony przypomina Lema koncepcję myślącej maszyny i z koncepcją tą dyskutuje, a z drugiej polemizuje ze współczesnymi, przesadnymi oczekiwaniami dotyczącymi inteligentnych maszyn.

Słowa kluczowe: inteligentne maszyny, filozofia techniki, lęki cywilizacyjne, futurologia, Stanisław Lem

Kontakt: Instytut Filozofii, Wydział Humanistyczny, Uniwersytet Śląski w Katowicach, ul. Bankowa 11, 40-007 Katowice
E-mail: mariusz.wojewoda@us.edu.pl
<https://us.edu.pl/instytut/ifil/osoby/mariusz-wojewoda/>
ORCID 0000-0003-0732-7500

Mariusz WOJEWODA, The Golem XXV Project: Awaiting the Technological Improvement in the Quality of Life

DOI 10.12887/38-2025-2-150-11

The author revisits Stanisław Lem's *Golem XIV* in order to compare the writer's expectations of a "thinking machine" with those of a modern user of technological tools. As it appears, Lem's conclusions are still relevant, despite the fact that the discussed novel was published over forty years ago. Following Lem's insights, the author analyzes various aspects of the relationship between man and intelligent machines and argues that artificial intelligence at the current stage of its development is unlikely to rid our civilization of the problems it faces. In the foreseeable future, the issues of maintaining the quality of life, or improving it, will remain dependent on the activity of human beings. However, the results of the next stages of AI development are impossible to anticipate. The fact remains that the technological progress we witness demands multifaceted reflection in the field of the humanities focused on the issue of the human being as a rational and responsible controller of intelligent machines. The article embraces three parts, discussing, respectively, today's expectations of AI, its impact on the

quality of human life, and its role in the future of our civilization. The essay is simultaneously a polemic: while revisiting and discussing Lem's concept of a thinking machine, the author criticizes modern day exaggerated expectations of intelligent machines.

Keywords: intelligent machines, philosophy of technology, civilization fears, futurology, Stanisław Lem

Contact: Institute of Philosophy, Faculty of Humanities, University of Silesia in Katowice, ul. Bankowa 11, 40-007 Katowice, Poland

E-mail: mariusz.wojewoda@us.edu.pl

<https://us.edu.pl/instytut/ifil/osoby/mariusz-wojewoda/>

ORCID 0000-0003-0732-7500