

# ETHOS

KWARTALNIK  
INSTYTUTU  
JANA PAWŁA II KUL  
LUBLIN

Nr 144

październik-grudzień 2023

## SZTUCZNA INTELIGENCJA W ŚWIECIE WARTOŚCI LUDZKICH

W numerze m.in.

Zbigniew WRÓBLEWSKI  
Paweł FORTUNA

\* Moral Subjectivity  
and the Moral Status  
of Artificial Intelligence

Małgorzata ABASSY

\* Snucie opowieści  
jako narzędzie osvajania  
sztucznej inteligencji

Andrzej BORKOWSKI

\* Sztuczna inteligencja  
a problem języka i tekstu

Barbara KLONOWSKA

\* Ethical Machines

Izabela KACZMARZYK

\* Rzecz „godna widzenia”  
czy „apokaliptyczny potwór”?

Zbigniew NAWRAT

\* Medycyna i ethos robotów



KOLEGIUM REDAKCYJNE  
KWARTALNIKA „ETHOS”

Michel BASTIT, Sergio BELARDINELLI, Gerald J. BEYER, Rocco BUTTIGLIONE  
Christoph BÖHR, Barbara CHYROWICZ SSpS, John CROSBY, Aimable-André DUFATANYE  
Krzysztof DYBCIAK, Tomasz GARBOL, Dariusz GAWIN, Piotr GUTOWSKI, Michael JAMES  
Wojciech KACZMAREK, Karol KLAUZA, Stanislav KOŠČ, Ewa Agnieszka LEKKA-KOWALIK  
Hubert ŁASZKIEWICZ, Leszek MĄDZIK, Balázs M. MEZEI, Maciej NOWAK, ks. Sławomir NOWOSAD  
Antonio PARDO CABALLOS, Furio PESCI, Adam POTKAY, Giovanni SALMERI  
Stefan SAWICKI, ks. Jan SOCHOŃ, Manfred SPIEKER, Zbigniew STAWROWSKI  
Eleonore STUMP, ks. Alfred M. WIERZBICKI, Linda ZAGZEBSKI, Ireneusz ZIEMIŃSKI

ZESPÓŁ REDAKCYJNY  
KWARTALNIKA „ETHOS”

ks. Marek SŁOMKA – redaktor naczelny  
Jarosław MERECKI SDS – zastępca redaktora naczelnego  
Natalia GONDEK – sekretarz redakcji  
Tomasz GÓRKA – sekretarz redakcji  
Dorota CHABRAJSKA, Mirosława CHUDA, Bernadetta KUCZERA-CHACHULSKA, Patrycja MIKULSKA  
Ferdie McDERMOTT (redaktor językowy), Wojciech KRUSZEWSKI (redaktor językowy)

REDAKTORZY TOMU

Ewa Agnieszka LEKKA-KOWALIK, Krzysztof WIECZOREK

RECENZENCI TOMU

Krzysztof ABRISZEWSKI, Maurizio BALISTRERI, Ewa BIŃCZYK, Anna BUGAJSKA  
Riccardo CAMPA, Joanna CHŁOSTA-ZIELONKA, ks. Michał DROŹDŹ, Julius ERDMANN  
Jan FAZLAGIĆ, Beata GAJ, Anna GRANAT, ks. Grzegorz HOŁUB, Wojciech KALISZEWSKI  
Stephan KAMPOWSKI, Jan KREFT, Ewa KRASKOWSKA, Piotr MAZUR, Adam MAZURKIEWICZ  
Christian MILLER, Janusz PASTERSKI, Sanja SELJAN, Norbert SCHAFFELD, Katarzyna SMYK  
Natasza SZUTTA, Wiesław WÓJCİK, Artur ŻYWIOŁEK

Przygotowanie do druku

Dorota CHABRAJSKA, Mirosława CHUDA, Natalia GONDEK  
Tomasz GÓRKA, Patrycja MIKULSKA, Mirella NAWRACAŁA-URBAN

Przygotowanie komputerowe

Konrad NOWAKOWSKI

Projekt okładki

Leszek MĄDZIK

Grafika na okładce

Kamil KUZKO

ETHOS jest kwartalnikiem filozoficznym o charakterze interdyscyplinarnym  
publikowanym w wersji papierowej (jest to wersja pierwotna).  
Fragmenty artykułów dostępne są na stronie [www.ethos.lublin.pl](http://www.ethos.lublin.pl).



Ministerstwo Kultury  
i Dziedzictwa Narodowego

Dofinansowano ze środków  
Ministra Kultury i Dziedzictwa Narodowego  
pochodzących z Funduszu Promocji Kultury

ISSN 0860-8024  
e-ISSN 2720-5355  
[www.ethos.lublin.pl](http://www.ethos.lublin.pl)  
Nakład 400 egz.

# ETHOS

Rok 36  
2023 nr 4(144)

## SZTUCZNA INTELIGENCJA W ŚWIECIE WARTOŚCI LUDZKICH

- \* Od Redakcji – Rozwijać czy nie rozwijać? Oto jest pytanie (A.L.K., K.W.) **5**
- \* From the Editors – To Develop or not to Develop, that Is the Question (A.L.K., K.W.) **11**
- \* JAN PAWEŁ II – Misja Kościoła w erze komputerów (Oroędzie papieskie na XXIV Światowy Dzień Środków Społecznego Przekazu 27 maja 1990 roku, Watykan, 24 I 1990) **19**
- \* JOHN PAUL II – The Christian Message in a Computer Culture (Message for the 24th World Communications Day 27 May 1990, Vatican, 24 January 1990) **23**

## SZTUCZNA INTELIGENCJA PODMIOTEM?

- \* Zbigniew WRÓBLEWSKI, Paweł FORTUNA – Moral Subjectivity and the Moral Status of Artificial Intelligence: A Philosophical and Psychological Perspective **29**  
DOI 10.12887/36-2023-4-144-05
- \* Małgorzata ABASSY – Snucie opowieści jako narzędzie osvajania sztucznej inteligencji. Między modelującą funkcją języka a podmiotowością maszyn **49**  
DOI 10.12887/36-2023-4-144-06

## SZTUCZNA INTELIGENCJA W ŚWIECIE PRZEDSTAWIONYM

- \* Andrzej BORKOWSKI – Sztuczna inteligencja a problem języka i tekstu. W kręgu *Golema XIV* Stanisława Lema **67**  
DOI 10.12887/36-2023-4-144-07

- \* Barbara KLONOWSKA – Ethical Machines: Representations of Artificial Intelligence in Ian McEwan’s *Machines Like Me* and Kazuo Ishiguro’s *Klara and the Sun* **87**  
DOI 10.12887/36-2023-4-144-08
- \* Grzegorz TRĘBICKI – Mariaż dwóch światów. Quasi-chrześcijańska wizja ewolucji człowieka i kosmosu w *Pieśniach Hyperiona* Dana Simmonsa **107**  
DOI 10.12887/36-2023-4-144-09

#### OBAWY I OCZEKIWANIA

- \* Bartłomiej KNOSALA – Zarządzanie środowiskiem naturalnym przez sztuczną inteligencję. Ograniczenia i wyzwania narracji postnatury **127**  
DOI 10.12887/36-2023-4-144-10
- \* Izabela KACZMARZYK – Rzecz „godna widzenia” czy „apokaliptyczny potwór”? Konsekwencje maszynizacji górnośląskiej rzeczywistości kulturowej **146**  
DOI 10.12887/36-2023-4-144-11
- \* Zbigniew NAWRAT – Medycyna i ethos robotów **169**  
DOI 10.12887/36-2023-4-144-12

#### KAROL WOJTYŁA – JAN PAWEŁ II INSPIRACJE

- \* Fr. Władysław ZUZIAK, Barbara ŻMUDA-FRYDRYCHOWSKA – Man’s Participation in the World: Wojtyła – Rorty **193**  
DOI 10.12887/36-2023-4-144-13

#### MYŚLAĆ OJCZYŻNA...

- \* Barbara TRYGAR – „Wielka jest potęga pamięci...” *Mój Ojciec* Kazimierza Brauna **211**

## OMÓWIENIA I RECENZJE

- \* Łukasz MŚCISŁAWSKI – When Wisdom and BigTech-AI Collide (rec. E. Spence, *Stoic Philosophy and the Control Problem of AI Technology: Caught in the Web*, Lanham, Rowman & Littlefield, Lanham 2021) **223**
- \* Damian SZCZĘCH – Plan na życie w świecie zdominowanym przez sztuczną inteligencję (rec. C. Campbell, *AI by Design: A Plan for Living With Artificial Intelligence*, CRC Press, New York 2022) **231**
- \* Propozycje „Ethosu” (J. Puckett Jr., *Apologetyka Radości. C.S. Lewisa argument z Pragnienia na rzecz istnienia Boga*, tłum. P. Bylica, Fundacja Prodoteo, Warszawa 2023) **236**

## PRZEZ PRYZMAT ETHOSU

- \* Dorota CHABRAJSKA – Gwarancja dla odrobiny umysłu **245**
  
- \* Noty o autorach **253**
- \* Contents **257**



OD REDAKCJI

## ROZWIJAĆ CZY NIE ROZWIJAĆ? OTO JEST PYTANIE

Niewiele jest dziś tematów, które tak bardzo rozbudzają i ożywiają zbiorową wyobraźnię, jak zagadnienia związane z rozwojem sztucznej inteligencji (SI, ang. artificial intelligence – AI), często mające posmak sensacji. Ogromne rzesze ludzi śledzą na bieżąco medialne doniesienia o nowych technologiach i nowych obszarach ich zastosowań, a także o kontrowersjach wywoływanych przez nie zawsze w pełni przemyślany i kontrolowany proces wdrażania technologii SI. Powszechnie panuje przekonanie, że w niedalekiej przyszłości dojdzie do radykalnych i zasadniczych zmian, zarówno w najbliższym otoczeniu człowieka, jak też w ludzkim sposobie bycia i działania. Niektóre scenariusze, szkicowane przez takich myślicieli jak na przykład Nick Bostrom<sup>1</sup>, Max Tegmark<sup>2</sup>, Ray Kurzweil<sup>3</sup>, czy Kevin Warwick<sup>4</sup> zakładają, że w niedługim czasie nastąpi eksplozja superinteligencji, która osiągnie potencjał umożliwiający jej kontrolę nad światem. 22 marca 2023 roku na stronach Future of Life Institute opublikowano list otwarty wzywający wszystkie laboratoria pracujące obecnie nad sztuczną inteligencją do natychmiastowego zatrzymania na co najmniej sześć miesięcy rozwoju i szkolenia systemów AI silniejszych niż Generative Pre-trained Transformer 4 (GPT-4) firmy OpenAI<sup>5</sup>. Pod listem podpisali się między innymi Elon Musk, Yuval Noah Harari, Steve Wozniak, Jaan Tallinn i wielu badaczy z obszaru AI. List wyłuszcza znane z mass mediów (a także z filmów i literatury) obawy przed sztuczną inteligencją wyższą niż ludzka.

---

<sup>1</sup> N. B o s t r o m, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford 2014.

<sup>2</sup> M. T e g m a r k, *Life 3.0: Being Human in the Age of Artificial Intelligence*, Knopf Publishing Group, New York 2017.

<sup>3</sup> R. K u r z w e i l, *The Age of Spiritual Machines: When Computers Exceed Human Intelligence*, Penguin Books, New York and London 2000.

<sup>4</sup> R. W a r w i c k, *March of the Machines: The Breakthrough in Artificial Intelligence*, University of Illinois Press, Champaign 2004.

<sup>5</sup> Zob. *Pause Giant AI Experiments: An Open Letter*, <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.



„Musimy – głosi list – zadać sobie pytanie: czy powinniśmy pozwolić maszynom zalewać nasze kanały informacyjne propagandą i nieprawdą? Czy powinniśmy zautomatyzować wszystkie zadania, w tym te [nas] spełniające? Czy powinniśmy rozwijać pozaludzkie umysły, które mogą ostatecznie przewyższyć nas liczebnie, przechytryć, uznać za przestarzałych i zastąpić? Czy powinniśmy ryzykować utratę kontroli nad naszą cywilizacją?”<sup>6</sup>. Równie często jak o zagrożeniach słyszemy też o korzyściach płynących z wykorzystania SI, poczynając od szybkiego dostępu do informacji, poprzez wzrost wydajności pracy, bezpieczeństwa (a także innych społecznych wartości, na przykład edukacji czy demokracji), aż po opiekę nad osobami starszymi czy chorymi. Doniesień o kolejnych osiągnięciach jest imponująco dużo. Śledząc nowiny ze świata SI, nieodmiennie oscylujemy między fascynacją i pełnym nadziei oczekiwaniem a niepokojem i lękiem o przyszłość. To nie przypadek, zarówno bowiem obawy, jak i nadzieje są całkowicie uzasadnione. Racja jest fundamentalna: już dziś można dostrzec, jak wiele ważnych zmian dokonało się i wciąż dokonuje praktycznie we wszystkich obszarach codziennego życia ludzi pod wpływem coraz intensywniejszego nasycenia środowiska obiektami wyposażonymi w sztuczną inteligencję. Spodziewamy się zatem, że jeszcze więcej nowości pojawi się nadchodzących latach.

Rozpatrywane są różne prognozy dotyczące miejsca i roli człowieka w świecie opanowanym przez samodoskonalącą się superinteligencję – od wizji optymistycznych, przewidujących, że cały stale rosnący potencjał intelektualny i sprawczy SI wykorzystywany będzie dla dobra ludzkości, aż po skrajnie katastroficzne, zapowiadające zagładę ludzkości i kolonizację całego wszechświata przez uwolnioną z więzów zależności od człowieka nową postać inteligentnego życia<sup>7</sup>. Trudno orzec, która z prognoz jest bliższa spełnienia. Wydaje się jednak wysoce prawdopodobne, że cywilizacja naukowo-techniczna nieodwracalnie weszła już na drogę wiodącą ku autonomicznej technoewolucji (przepowiedzianej już w latach sześćdziesiątych ubiegłego wieku przez Stanisława Lema w dziele *Summa technologiae*<sup>8</sup>), która może wymknąć się spod ludzkiej kontroli. W celu prowadzenia refleksji nad SI powstały między innymi Consortium for the Benevolent Consciousness of Artificial Intelligence, Future of Life Institute i inne podobne organizacje o światowym zasięgu, a Organizacja Narodów Zjednoczonych rozważa powołanie agendy zajmującej się SI na wzór Międzynarodowej Agencji Energii Atomowej<sup>9</sup>.

<sup>6</sup> Tamże (tłum. fragm. A.L.K., K.W.)

<sup>7</sup> Por. E. Yudkowsky, *Artificial Intelligence as a Positive and Negative Factor in Global Risk*, w: *Global Catastrophic Risk*, red. N. Bostrom, M. Čirković, Oxford University Press, Oxford 2008, s. 308-362.

<sup>8</sup> Zob. S. Lem, *Summa technologiae*, Wydawnictwo Literackie, Kraków 1964.

<sup>9</sup> Zob. M. Druś, *Sekretarz generalny ONZ jest za międzynarodowym nadzorem nad sztuczną inteligencją*, <https://www.pb.pl/sekretarz-generalny-onz-jest-za-miedzynarodowym-nadzorem-nad-sztuczna-inteligencja-1187767>.



Tytułowe pytanie ma właściwie charakter retoryczny. Sztuczna inteligencja będzie trwałym elementem naszego ludzkiego świata, wpływając na kształt życia i wartości. Obecne czasy nazywane są czwartą rewolucją przemysłową. Świadomość, że w dziejach ludzkości i w organizacji światowego ładu nadchodzi wielkie, znaczące zmiany, budzi potrzebę racjonalnego, odpowiedzialnego namysłu nad długofalowym planem działania, wyzwala szeroką skalę emocji, pobudza wyobraźnię. Suma tych trzech elementów, a dokładniej ich wzajemne oddziaływanie i emergentna interferencja, wytwarza cenny kapitał kulturowy, który jako ludzkość powinniśmy jak najlepiej wykorzystać w trosce o lepszą i bezpieczniejszą przyszłość całej planetarnej populacji oraz jej naturalnego i cywilizacyjnego środowiska.

O problemach związanych ze sztuczną inteligencją można myśleć i pisać na wiele sposobów. Można do nich podchodzić w sposób ściśle techniczny, formułując konkretne zadania konstrukcyjne i poszukując środków ich realizacji; można rozpatrywać te problemy w aspektach: ekonomicznym, ekologicznym, prawnym, etycznym, pedagogicznym i innych; można pytać o możliwości zastosowań SI w wielu ważnych dziedzinach życia i pracy: w przemyśle, w nauce, w twórczości artystycznej, w służbie zdrowia, w nauczaniu i wychowaniu czy w wojskowości. Wszystkie te kwestie są ważne i pilne, wymagają wyężonej pracy koncepcyjnej oraz jak najlepszej organizacji i synchronizacji działań. Należy też zwrócić uwagę na dwa obszary wymagające refleksji, które są równie ważne, lecz trudniej uchwytne, bo nie mieszczą się w ramach konkretnych dyscyplin i kompetencji. Pierwszy z nich to troska o szeroko rozumiane cyberbezpieczeństwo. Obejmuje ona między innymi namysł nad sposobami skutecznej ochrony ludzkości przed użyciem zasobów SI w złej wierze i dla osiągnięcia niegodziwych celów – na przykład przez grupy przestępcze czy terrorystyczne, ale także przez osoby i środowiska dążące do zdobycia przewagi nad innymi za pomocą nieetycznej manipulacji technologią SI, poprzez wykorzystanie jej na przykład przeciwko konkurentom lub przeciwnikom w walce o rzadkie zasoby. Inny aspekt cyberbezpieczeństwa, którego także nie można pominąć, to opracowanie zawczasu jak najefektywniejszych środków zaradczych chroniących przed uruchomieniem (świadomym lub mimowolnym) niepożądanych i niebezpiecznych ścieżek rozwoju SI, prowadzących do autonomicznego i niekontrolowanego tworzenia systemów, programów i technologii pośrednio lub bezpośrednio zagrażających ludziom. Wszystkie te zagrożenia są realne i pominięcie ich w publicznej debacie oraz w specjalistycznym dyskursie naukowym byłoby niewybaczalną lekkomyślnością. Drugi obszar wymagający głębokiego namysłu to kwestia relacji między ludźmi a systemami sztucznej inteligencji. Chodzi przede wszystkim o wypracowanie funkcjonalnych i dobrze osadzonych w realiach kulturowych modeli myślenia, zachowań i odniesień do nowo powstającej i niekiedy zaskakującej (pozytyw-

nie lub negatywnie) cywilizacyjnej przestrzeni oddziaływań między ludzką a nie-ludzką inteligencją; modeli pozwalających wzmocnić pozytywne interakcje w obrębie owej przestrzeni oraz w miarę możliwości unikać tych złych, niepokojących i niebezpiecznych.

Im bardziej zmienia się świat wokół nas, wypełniając się w coraz większym stopniu złożonymi, niezrozumiałymi i nieprzewidywalnymi narzędziami, mogącymi służyć człowiekowi i poszerzać jego horyzonty, ale też stawiać przed sobą i realizować własne cele, również cele niezgodne z interesem ludzkości, tym bardziej powinniśmy zadbać o wytworzenie bezpiecznej strefy psychicznego komfortu, opartej na rzetelnej wiedzy, a także na mądrze skonstruowanych tekstach kultury<sup>10</sup>, pomagających zwykłemu człowiekowi przezwyciężyć wrażenie obcości czy może nawet niesamowitości SI i poczuć się komfortowo w otoczeniu tak bardzo do nas podobnych i równocześnie tak bardzo odmiennych bytów, wyposażonych w „prawie ludzką” inteligencję, a przy tym znacznie wyprzedzających nas w coraz to nowych obszarach kompetencji. Nie jest to łatwe zadanie, lecz nie wolno zaniedbać jego realizacji, gdyż zaniedbanie to skutkowałoby wyobcowaniem człowieka w świecie, w którym ustalanie zasad i reguł postępowania stopniowo przestanie być wyłączną domeną ludzi. Trzeba również zaakceptować nieuchronny proces metamorfozy ludzkiej tożsamości, wywołany przez implantację zaawansowanych wytworów inteligentnej technologii w obręb ciała ludzkiego, bądź też przez zmiany funkcjonowania kory mózgowej pod wpływem nieprzerwanego kontaktu z elektronicznymi narzędziami komunikacji. Co więcej, nie chodzi tylko o mentalne i behawioralne dostosowywanie się jednostek do nowych aspektów zewnętrznej i wewnętrznej rzeczywistości sztucznej inteligencji. To również wykreowanie nowych ram kulturowych, kodów i idiomów, w których dokona się *sui generis* naturalizacja sztucznej inteligencji. Sam termin „sztuczna” niesie w sobie niepokojącą dwuznaczność. Pierwsze znaczenie słowa „sztuczny” (ang. artificial) odnosi się do artefaktu, przedmiotu niebędącego częścią naturalnego środowiska, lecz wytworzonego za pomocą narzędzi na podstawie uprzedniego projektu. Mamy tu do czynienia z opozycją sztuczny-naturalny, w sensie źródła zaistnienia: zrobione przez człowieka-zrodzone przez przyrodę. To znaczenie pozostaje w mocy, choć coraz więcej elementów naszego środowiska naturalnego to artefakty. Czymże jest na przykład ogród, w którym miejsce każdej rośliny zostało zaplanowane, a te pojawiające się samoistnie są bezwzględnie usuwane? Zapewne naturalnym artefaktem... W słowie „sztuczny” pobrzmiewa także pewna – co najmniej potencjalnie wykluczająca – opozycja: to, co sztuczne, jest nienaturalne, co można niekiedy rozumieć jako „prze-

---

<sup>10</sup> A. M a j, *Przemiany wiedzy w cyberkulturze: badania nad kulturą, komunikacją, wiedzą i mediami*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2021.

ciwne naturze”. Komponent emocjonalno-wartościujący pojęcia „sztuczność” jest bardzo często negatywny, co nietrudno zauważyć w wielu kontekstach; choćby wówczas, gdy mówimy o sztuczności czyjegoś zachowania albo gdy dostajemy w prezencie sztuczne kwiaty zamiast prawdziwych. Takie skojarzenia, niekoniecznie uświadomione, mówiąc metaforycznie, obciążają hipotekę sztucznej inteligencji, ponieważ już w punkcie wyjścia ewokują wobec niej nieufność, dystans i rezerwę, jeśli nie wręcz niechęć. Chyba nadszedł czas, by przewyciężyć tego typu uprzedzenia. Nie oznacza to bynajmniej bezkrytycznej akceptacji wszelkich kierunków rozwoju SI. Chodzi raczej o świadomie krytyczne podejście i możliwie rzetelną analizę ewentualnych zagrożeń, a nie powielanie negatywnych stereotypów. Być może strach przed SI rządzącą światem i eliminującą ludzi jest nieuzasadniony, bo – jak twierdzą Jobst Landgrebe i Barry Smith<sup>11</sup> – powstanie tak zwanej silnej sztucznej inteligencji jest matematycznie niemożliwe, a dopiero ona we wszystkich aspektach mogłaby przewyższyć inteligencję ludzką. Nie oznacza to, że ryzyko egzystencjalne związane z wykorzystaniem SI w rozmaitych obszarach życia społecznego nie wymaga namysłu, także etycznego, a nawet regulacji prawnych<sup>12</sup>.

Próby odpowiedzialnej refleksji nad wieloma z poruszonych wyżej zagadnień podejmują autorzy tekstów zamieszczonych w niniejszym tomie „Ethosu”. Wiele uwagi poświęcają oni analizie przemian zachodzących w języku, który staje się narzędziem kulturowego „udomowienia” sztucznej inteligencji. Podjmują też problem funkcji kulturowej i znaczenia tekstów literackich poświęconych relacjom między człowiekiem a SI. Szczególnie pasjonującym wątkiem, który pojawia się w artykułach zamieszczonych w tym tomie, jest pytanie o „etykę maszyn”: czy roboty i inne obiekty SI przyswoją sobie właściwe dla ludzkiej kultury wartości i normy etyczne, czy też stworzą własną moralność, być może pozbawioną humanistycznej wrażliwości?<sup>13</sup> W przypadku robotów medycznych, którym poświęcony jest jeden z tekstów, pytanie to staje się kluczowe. Równie pasjonującym problemem jest kwestia powiązań między rozwojem SI a imperatywem ochrony środowiska naturalnego: czy sztuczna inteligencja uratuje świat dzięki implementacji optymalnych działań proekologicznych w skali planetarnej, czy przeciwnie – przyspieszy globalną katastrofę?

---

<sup>11</sup> Zob. J. Landgrebe, B. Smith, *Why Machines Will Never Rule the World: Artificial Intelligence without Fear*, Routledge, New York – London 2023.

<sup>12</sup> Do końca 2023 roku Unia Europejska zamierza poddać regulacji prawnej obszar sztucznej inteligencji. Zob. A. Maj, *Unia Europejska coraz bliżej AI Act. Nowe prawo ma uregulować kwestię sztucznej inteligencji*, <https://fakty.tvn24.pl/fakty-o-swiecie/unia-europejska-coraz-blizej-ai-act-nowe-prawo-ma-uregulowac-kwestie-sztucznej-inteligencji-7164166>.

<sup>13</sup> Por. K. Warwick, *Cyborg Morals, Cyborg Values, Cyborg Ethics*, „Ethics and Information Technology” 5(2003) nr 3, s. 131-137.

Każda inicjatywa zmierzająca do zaangażowania intelektu, uczuć i wyobraźni we wspólną pracę nad powszechnie akceptowalnym projektem optymalnej koegzystencji (kohabitacji?) człowieka z obecnie istniejącą i przyszłą sztuczną inteligencją, powinna zostać przyjęta z życzliwością i satysfakcją. Redaktorzy niniejszego tomu mają nadzieję, że zamieszczone w nim prace autorów pochodzących z różnych środowisk akademickich przyczynią się do pogłębienia i upowszechnienia takiego właśnie – całościowego, integrującego rozmaite punkty widzenia i angażującego wszystkie wymiary ludzkiej percepcji rzeczywistości – sposobu ujmowania problematyki sztucznej inteligencji. Co dalej nastąpi, tego nie da się dokładnie przewidzieć, ale z pewnością wiadomo, że najgorszym z możliwych rozwiązań jest obojętność wobec nadchodzącej przyszłości i bierne czekanie na rozwój wydarzeń. Dopóki mamy jakikolwiek wpływ na kierunki rozwoju SI, musimy uczynić wszystko, by zmaksymalizować szansę na powstanie dobroczynnej sztucznej inteligencji, życzliwej i przyjaznej człowiekowi w rezultacie intensywnych naukowych badań.

*Agnieszka Lekka-Kowalik  
Krzysztof Wieczorek*

## FROM THE EDITORS

### TO DEVELOP OR NOT TO DEVELOP, THAT IS THE QUESTION

Few topics stimulate and enliven the collective imagination today to the extent issues related to the development of artificial intelligence (AI) do, simultaneously adding a tinge of sensation to our lives. People tend to follow media reports on the new technologies and on the new areas of their application, as well as those related to the controversies over the—not always predictable or easily controlled—implementation of AI technologies. There is a widespread belief that, in the near future, radical and fundamental changes will be taking place, both in the immediate human environment and in the human way of being and acting. Some scenarios, sketched by thinkers such as Nick Bostrom,<sup>1</sup> Max Tegmark,<sup>2</sup> Ray Kurzweil,<sup>3</sup> and Kevin Warwick,<sup>4</sup> imply that very soon we shall witness an intelligence explosion on a scale making the superintelligence capable of controlling the world. On March 22, 2023, an open letter was published on the website of the Future of Life Institute, calling on all AI labs to immediately pause, for at least the period of six months, the developing and training of AI systems more powerful than GPT-4 (Generative Pre-trained Transformer 4) created by OpenAI.<sup>5</sup> The letter was signed by, among others, Elon Musk, Yuval Noah Harari, Steve Wozniak, Jaan Tallinn, and by many other AI researchers. It names the fears and concerns ignited by AI development which have been made public by mass media, books, and movies: “We must ask ourselves: *Should* we let machines flood our information channels with propaganda and untruth? *Should* we automate away all the jobs, including the fulfilling ones? *Should* we develop nonhuman minds that might eventually

---

<sup>1</sup> See Nick B o s t r o m, *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014).

<sup>2</sup> See Max T e g m a r k, *Life 3.0: Being Human in the Age of Artificial Intelligence* (New York: Knopf Publishing Group, 2017).

<sup>3</sup> See Ray K u r z w e i l, *The Age of Spiritual Machines: When Computers Exceed Human Intelligence* (New York and London: Penguin Books, 2000).

<sup>4</sup> See Kevin W a r w i c k, *March of the Machines: The Breakthrough in Artificial Intelligence* (Champaign: University of Illinois Press, 2004).

<sup>5</sup> See “Pause Giant AI Experiments: An Open Letter,” The Future of Life Institute, <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>.

outnumber, outsmart, obsolete and replace us? *Should* we risk loss of control of our civilization?"<sup>6</sup>

Just as often as we hear about the risks, we also hear about the benefits of AI, ranging from rapid access to information, increased productivity, and greater security (as well as other social values, such as education or democracy), to care for the elderly or the sick. Reports on accomplishments made possible by AI are as numerous as they are impressive. Following news from the AI world, we oscillate between fascination, accompanied by hopeful anticipation, and anxiety which occasionally turns into fear for the future. This is no coincidence for, in the face of the continuing AI development, both hopes and fears are entirely legitimate. The fundamental factor triggering such conflicting attitudes is that already now one can experience the impact, in practically every aspect of human life, of the ever increasing saturation of the environment with AI-equipped objects. Moreover, the prospect of a further development of AI only contributes to the polarization of views regarding its ubiquity.

Various prognoses concerning the place and role of the human being in the world controlled by a self-perfecting superintelligence are considered: from optimistic visions that AI's growing intellectual and causative potential will be used for the benefit of humankind up to entirely catastrophic visions of the annihilation of humanity and the colonization of the entire universe by a new form of intelligent life liberated from the bonds of dependence on the human being.<sup>7</sup> It is difficult to determine which prognosis is closer to fulfillment. Yet it seems very probable that the techno-scientific civilization has paved the way to autonomous techno-evolution (predicted already in the 1960s by Stanisław Lem in his *Summa technologiae*<sup>8</sup>) which may get out of human control. In order to carry on systematic reflection on AI, various global organizations were created, such as Consortium for the Benevolent Consciousness of Artificial Intelligence or the Future of Life Institute. The United Nations in turn considers creating an agency—inspired by the International Atomic Energy Agency—for monitoring AI development.<sup>9</sup>

The question posed in the title of this essay is in fact rhetorical. Artificial intelligence will definitely become a permanent factor in the human world, influencing the shape of our lives and values. The current epoch is called that

<sup>6</sup> Ibidem.

<sup>7</sup> See Eliezer Yudkowsky, *Artificial Intelligence as a Positive and Negative Factor in Global Risk*, in: *Global Catastrophic Risk*, eds. Nick Bostrom and Milan M. Ćirković (Oxford: Oxford University Press 2008), 308–62.

<sup>8</sup> See Stanisław Lem, *Summa technologiae*, trans. Joanna Żylińska (Minneapolis: University of Minnesota Press, 2013).

<sup>9</sup> See Jason Nelson, "Take AI Warnings Seriously, Says UN Secretary-General," Decrypt U: News, <https://decrypt.co/144692/take-ai-warnings-seriously-un-secretary-general>.



of the fourth industrial revolution. As much as it affects our emotions and stimulates the collective imagination, the awareness that the entire humanity and the present world order are now facing inevitable substantial transformations calls for a responsible, rational focus on long-term planning. The sum of all these elements, or more precisely, their interplay and emergent interference, produce a precious cultural capital of which we should make the best possible use with a view to a better and safer future for the entire planetary population and for its natural and civilizational environment.

One may think and write about AI-related problems in many ways. They can be approached in a strictly technical way, by formulating specific construction tasks and then by looking for means of their implementation; economic, ecological, legal, ethical, pedagogical and other aspects of those problems might be considered; one can ask about the possibilities of AI applications in many important areas of life and work, such as industry, science, art, health care, education, and the military. All these issues are important and burning. To consider and solve them, intensive conceptual work and the best possible organization and synchronization of activities are indispensable.

Two other areas in need of reflection should also be indicated. They are equally important but more difficult to grasp, for they do not fit into the framework of a particular discipline or a specific set of competencies. The first one is broadly understood cybersecurity. It includes, among other things, considering ways to effectively protect humanity against the use of AI resources in bad faith or for wicked purposes, for instance by criminal or terrorist groups, or by individuals or communities seeking to gain advantage over others through unethical manipulation of AI technology, by using it against their competitors in the fight for scarce resources. Another aspect of cybersecurity which must not be neglected is the development, beforehand, of the most effective countermeasures to protect us against launching (consciously or accidentally) into undesirable and dangerous AI developmental paths that would end in the autonomous and uncontrolled creation of systems, programs and technologies directly or indirectly threatening people. All these risks are real and it would be tantamount to unforgivable recklessness to overlook them in the public debate or in specialized scientific discourse. The second issue that requires deep consideration is the relationship between humans and artificial intelligence systems. First of all, the issue in question concerns developing functional and culturally well-embedded models of thinking about, behaving in, and referring to the newly emerging and sometimes surprising (positively or negatively) civilizational space of interactions between human and non-human intelligence; models which should enhance positive interactions within that space and, as far as possible, allow us to avoid disturbing and dangerous ones.

The world around is to ever greater extent being filled with complex, unintelligible, and unpredictable devices which may serve the humans and broaden

their horizons but which may also formulate and realize their own tasks, even those contrary to the best interest of humanity. The more the world is changing, the more we should care to create a safe zone of psychological comfort, based on reliable knowledge, as well as on wisely constructed cultural texts,<sup>10</sup> helping ordinary people to overcome the feeling of alienation, or perhaps even the weirdness of AI, and to feel comfortable in the environment of entities so similar to us and at the same time so different, equipped with “almost human” intelligence and simultaneously outstripping us in ever new areas of competence. This is not an easy task. Yet we must not fail in realizing it for such a failure would be tantamount to an alienation of human beings from the world in which setting rules and regulations of conduct will gradually cease to be their exclusive competence.

One must also come to terms with the unavoidable process of reshaping human identity caused by the implantation of AI advanced products into the human body or by changes in functioning of the cerebral cortex resulting from the brain’s continuous contact with digital information-communication technologies. Moreover, what is at stake here it is not just the mental and behavioral adaptation of individuals to new aspects of the external and internal reality of artificial intelligence for we need to create new cultural frameworks, codes, and idioms in which artificial intelligence could be “naturalized.” The term “artificial” as such embraces a disturbing ambiguity. One of its meanings refers to an artifact, an object which is not part of the natural environment but is created by means of tools in accordance with a prior project. The opposition in question is that between the artificial and the natural, i.e., between created by human beings and created by nature. This contrast remains valid, even if a growing number of elements in the environment of our life become artifacts. For what is, for instance, a garden in which the place for each plant has been carefully planned, and any naturally growing one is ruthlessly removed? Is it not a natural artifact?... Yet the term “artificial” conveys another potential opposition: what is artificial is non-natural, i.e., it is directed against nature. This emotional-evaluative component of the meaning of the term “artificial” is often negative, as it may be noticed in various contexts, for instance, when we criticize somebody’s behavior as artificial or complain of getting artificial flowers instead of real ones. In a figurative sense, such mental associations, even if made unconsciously, mortgage artificial intelligence for they instantaneously evoke distrust, distance, and reserve (if not outright aversion) towards it. It seems the time has come to overcome such biases. This does not mean

---

<sup>10</sup> See Anna M a j, *Przemiany wiedzy w cyberkulturze: Badania nad kulturą, komunikacją, wiedzą i mediami* (A Transformation of Knowledge in Cyberculture: Research on Culture, Communication, Knowledge, and Media) (Katowice: Wydawnictwo Uniwersytetu Śląskiego, 2021).

that any AI development should be welcome. On the contrary, we should not repeat negative stereotypes but take consciously critical approach and carry on reliable—as far as it is possible—analysis of dangers and risks connected to AI. Perhaps the fear of AI ruling the world and eliminating humans is unfounded. As Jobst Landgrebe and Barry Smith argue, the creation of the so-called strong artificial intelligence is mathematically impossible, and only such an intelligence could surpass human intelligence in all aspects.<sup>11</sup> This does not, however, mean that existential risks created by the development and use of “ordinary” AI in many areas of life do not deserve considerations, also ethical ones, or legal regulations.<sup>12</sup>

The authors of the papers included in this volume of *Ethos* attempt to responsibly reflect on many of the issues raised above. They focus, among others, on the transformations of language, which has become a tool for the cultural “domestication” of artificial intelligence. They also address the problem of the cultural functions and meaning of literary texts devoted to the relationship between humans and AI. The question of “machine ethics” is a gripping theme in this context, and it frequently recurs throughout the volume as the authors address the question of whether robots and other AI objects will assimilate the ethical values and norms inherent in human culture or rather create their own morality, perhaps devoid of human sensitivity.<sup>13</sup> In the case of medical robots, a problem one of the articles specifically explores, this question becomes crucial. Equally gripping is the issue of the connections between AI development and the imperative to protect our natural environment. Will AI save the world thanks to implementing ecologically optimal solutions on a planetary scale or will it, contrary to such expectation, accelerate the ecological disaster?

Any initiative that engages intellect, emotions, and imagination in working on a generally recognized project of optimal co-existence (co-habitation?) of the human being with any AI—whether already present or developed in the future—should be welcome with goodwill and satisfaction. The editors of this volume hope that papers authored by thinkers representing diverse academic circles will contribute to the deepening and dissemination of such a holistic, i.e., integrating various points of view and involving all dimensions of human perception of reality, approach to artificial intelligence. What will come next

---

<sup>11</sup> See Jobst Landgrebe and Barry Smith, *Why Machines will Never Rule the World: Artificial Intelligence without Fear* (New York and London: Routledge, 2023).

<sup>12</sup> By the end of 2023 The European Union should regulate the use of AI with the AI Act, the world’s first comprehensive AI law. See “EU AI Act: first regulation on artificial intelligence,” European Parliament: News, <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>.

<sup>13</sup> See Kevin Warwick, “Cyborg Morals, Cyborg Values, Cyborg Ethics,” *Ethics and Information Technology* 5, no. 3 (2003): 131–7.

cannot be accurately predicted. But it is quite obvious that the worst possible solution is remaining indifferent to the coming future and passively waiting for future developments. As long as we have any influence on the directions of AI development, we must do everything possible to maximize the chance that a benevolent and friendly artificial intelligence will emerge from intensive scientific research.

*Agnieszka Lekka-Kowalik and Krzysztof Wieczorek*

JAN PAWEŁ II  
MISJA KOŚCIOŁA W ERZE KOMPUTERÓW

JOHN PAUL II  
THE CHRISTIAN MESSAGE IN A COMPUTER CULTURE





JAN PAWEŁ II

## MISJA KOŚCIOŁA W ERZE KOMPUTERÓW\*

Bracia i siostry, drodzy przyjaciele,

w jednej z Modlitw Eucharystycznych Kościoł zwraca się do Boga słowami: „Ty stworzyłeś na obraz i podobieństwo swoje i powierzyłeś mu cały świat, aby służąc Tobie samemu, jako Stwórcy, rządził wszelkim stworzeniem” (IV Modlitwa Eucharystyczna).

Wspaniały i ogromny jest sens pracy mężczyzny i kobiety, stworzonych przez Boga i obarczonych tym zadaniem. Poprzez myśli, działania i dokonania – nawet najbardziej powszednie – każdej istoty ludzkiej Stwórca odnawia świat, prowadzi go do zbawienia, czyni go coraz doskonalszym narzędziem swej chwały.

## POSTĘP TECHNICZNY „ZMIENIA OBLICZE ZIEMI”

Około dwudziestu pięciu lat temu Ojcowie Soboru Watykańskiego II w swych rozważaniach o Kościele w świecie współczesnym dali wyraz przekonaniu, że mężczyźni i kobiety oddający się codziennym zajęciom w swoich rodzinach i społecznościach „słusznie [...] mogą uważać, że przez swoją aktywność rozwijają dzieło Stwórcy [...] i dzięki osobistemu wysiłkowi przyczyniają się do realizacji zamysłu Bożego w historii”<sup>1</sup>.

Ojcowie Soboru, patrząc w przyszłość i starając się przewidzieć, w jakich okolicznościach przyjdzie Kościołowi wypełniać swoją misję, mogli już wówczas z całą jasnością dostrzec, że postęp techniczny „zmienia oblicze ziemi”, dążąc nawet do podboju przestrzeni pozaziemskiej<sup>2</sup>. Uznali, że w szczególno-

---

\* Orędzie papieskie na XXIV Światowy Dzień Środków Społecznego Przekazu 27 maja 1990 roku (Watykan, 24 I 1990). Przypisy pochodzą od redakcji. Cytaty z dokumentów Soboru Watykańskiego II dostosowano do ich najnowszego polskiego przekładu.

<sup>1</sup> Sobór Watykański II, Konstytucja duszpasterska o Kościele w świecie współczesnym *Gaudium et spes*, nr 34, w: Sobór Watykański II, *Konstytucje. Dekrety. Deklaracje*, Pallottinum, Poznań 2012, s. 550.

<sup>2</sup> Zob. tamże, nr 5, s. 529.

ści szybkie tempo rozwoju nowych technologii przekazu informacji musi wywołać reakcję łańcuchową i prowadzić do nieoczekiwanych konsekwencji.

Nie sugerując bynajmniej, że Kościół winien pozostawać na uboczu i nie angażować się w te procesy, Ojcowie Soborowi widzieli miejsce Kościoła w samym sercu ludzkiego postępu, jako uczestnika doświadczeń całej ludzkości, starającego się rozumieć je i interpretować w świetle wiary. Powinnością członków Ludu Bożego jest twórcze wykorzystywanie nowych odkryć i technologii dla dobra ludzkości i do realizacji Bożego planu wobec świata.

Następne lata pokazały, jak słuszne były te przewidywania szybkich przemian i jak celowe otwarcie na rozwój techniki, ponieważ tempo zmian i rozwoju wzrosło jeszcze bardziej. Dzisiaj na przykład nie rozpatruje się już środków społecznego przekazu jedynie jako swoistych narzędzi czy technik, ale uważa je raczej za część pewnej kultury, wciąż jeszcze podlegającej ewolucji, której wszystkie implikacje nie są na razie dokładnie rozpoznane, a jej możliwości wykorzystywane są jak dotąd tylko częściowo.

Na takiej podstawie opierają się rozważania, jakie podejmujemy z okazji XXIV Światowego Dnia Środków Społecznego Przekazu. Z każdym dniem coraz bardziej urzeczywistnia się to, co kiedyś było tylko wizją. Przewidywała ona rozwój realnego dialogu między odległymi narodami, powszechne dzielenie się wartościami i aspiracjami, coraz większe wzajemne poznanie i zrozumienie, zacieśnienie więzów braterstwa ponad licznymi podziałami, które wówczas były nie do pokonania<sup>3</sup>.

#### POGŁĘBIENIE DIALOGU ZE WSPÓŁCZESNYM ŚWIATEM

Wraz z pojawieniem się komputerowych technik telekomunikacji oraz tak zwanych skomputeryzowanych systemów uczestnictwa Kościół otrzymał nowe środki realizacji swojej misji. Metody ułatwiające porozumiewanie się i dialog między członkami Kościoła pozwalają na zacieśnienie więzów jedności. Bezpośredni dostęp do informacji umożliwia Kościołowi pogłębienie dialogu ze współczesnym światem. W nowej kulturze komputerów Kościół może szybciej informować świat o swoim „credo” i wyjaśniać swe stanowisko wobec każdego problemu czy wydarzenia. Może też wyraźniej słyszeć głos opinii publicznej i prowadzić nieustanny dialog z otaczającym go światem, angażując się dzięki temu jeszcze aktywniej we wspólne poszukiwanie rozwiązań dla licznych, palących problemów ludzkości<sup>4</sup>.

<sup>3</sup> Zob. Papieska Rada ds. Środków Społecznego Przekazu, Instrukcja duszpasterska o środkach społecznego przekazu *Communio et progressio*, nr 181-182.

<sup>4</sup> Zob. tamże, nr 144-145.

Jest oczywiste, że także Kościół, spełniając coraz bardziej naglące zadanie ewangelizacji, powinien skorzystać z nowych możliwości, które otworzył postęp w dziedzinie techniki komputerowej i satelitarnej. Najistotniejsze i najpilniejsze orędzie Kościoła dotyczy poznania Chrystusa i ofiarowanej przez Niego drogi zbawienia. To właśnie orędzie Kościół musi przedstawiać wszystkim ludziom, niezależnie od wieku, zachęcając ich do przyjęcia Ewangelii z miłością, a zarazem nie zapominając, że „prawda narzuca się sumieniu tylko siłą siebie samej i ogarnia umysły łagodnie, a zarazem zdecydowanie”<sup>5</sup>.

Jak naucza nas doświadczenie czasów minionych: „Bóg, objawiając się swojemu ludowi [...], przemawiał odpowiednio do kultury właściwej różnym okresom. Podobnie Kościół, żyjąc na przestrzeni czasów w różnych warunkach, stosował zdobycze różnych kultur, aby swoim przepowiadaniem szerzyć orędzie Chrystusa i wyjaśniać je wszystkim narodom”<sup>6</sup>. „W pierwszym zapoznaniu z wiarą, w nauczaniu katechetycznym czy w dalszym pogłębianiu wiary nie może zabraknąć pomocy środków społecznego przekazu. [...] Kościół byłby winny przed swoim Panem, gdyby nie używał tych potężnych pomocy, które umysł ludzki coraz bardziej usprawnia i doskonali. Za ich pośrednictwem głosi orędzie, które zostało mu powierzone, «na dachach»”<sup>7</sup>.

Z pewnością winniśmy być wdzięczni nowej technice, która pozwala nam gromadzić informacje w ogromnych systemach sztucznej pamięci stworzonych przez człowieka, umożliwiając przez to szeroki i bezpośredni dostęp do wiedzy stanowiącej nasze ludzkie dziedzictwo, do tradycji i nauczania Kościoła, do słów Pisma Świętego, do nauczania wielkich mistrzów duchowości, do historii i tradycji Kościołów lokalnych, do zakonów i instytutów świeckich, do idei i doświadczeń prekursorów i innowatorów, twórców instytucji dających nieustanne świadectwo obecności między nami miłującymi Ojca, który ze swego skarbcza wydobywa rzeczy nowe i stare (por. Mt 13,52).

## OKAŻMY ZAUFANIE MŁODZIEŻY!

Sz szczególnie młodzież szybko przystosowuje się do kultury komputerów i do ich „języka”, co z pewnością stanowi powód do zadowolenia. Okażmy zaufanie młodzieży!<sup>8</sup> Przewaga ludzi młodych polega na tym, że wzrasta-

<sup>5</sup> Sobór Watykański II, Deklaracja o wolności religijnej *Dignitatis humanae*, nr 1, w: Sobór Watykański II, *Konstytucje. Dekrety. Deklaracje*, s. 411.

<sup>6</sup> *Gaudium et spes*, nr 58, s. 573.

<sup>7</sup> P a w e ł VI, Adhortacja apostołska o ewangelizacji w świecie współczesnym *Evangelii nuntiandi*, nr 45, Opoka, [https://opoka.org.pl/biblioteka/W/WP/pawel\\_vi/adhortacje/evangelii\\_nuntiandi.html](https://opoka.org.pl/biblioteka/W/WP/pawel_vi/adhortacje/evangelii_nuntiandi.html).

<sup>8</sup> Zob. *Communio et progressio*, nr 70.

ją w czasie, gdy dokonuje się rozwój technologii; to im przypadnie zadanie wykorzystania nowych narzędzi w coraz szerszym i bardziej intensywnym dialogu między wszystkimi rasami i klasami, jakie zamieszkują nasz „coraz mniejszy świat”. Będą musieli odkryć odpowiednie formy zastosowania nowych systemów przechowywania i wymiany danych, by mogły one przyczynić się do umacniania coraz większej sprawiedliwości, do wzrostu poszanowania praw człowieka, do zdrowego rozwoju wszystkich jednostek i narodów oraz tych swobód, bez których niemożliwe jest życie w pełni ludzkie.

Wszyscy, młodzi i starzy, podejmijmy wyzwanie nowych odkryć i technologii, czyniąc z nich element wizji moralnej opartej na naszej wierze i na szacunku dla człowieka, zmierzającej do przebudowy świata według zamysłu Bożego! Obchodząc Światowy Dzień Środków Społecznego Przekazu, módlmy się, by cały potencjał „ery komputerów” został wprzęgnięty w służbę powołania ludzkiego i transcendentnego wobec człowieka, by przyniósł chwałę Ojcu, od którego pochodzą wszystkie rzeczy dobre.

JOHN PAUL II

## THE CHRISTIAN MESSAGE IN A COMPUTER CULTURE\*

Brothers and Sisters,  
Dear Friends,

In one of her Eucharistic Prayers, the Church addresses God in these words: “You formed man in your own likeness and set him over the whole world, to serve you, his creator, and to rule over all creatures” (Eucharistic Prayer IV).

For man and woman thus created and commissioned by God, the ordinary working day has great and wonderful significance. People’s ideas, activities and undertakings—however commonplace they may be—are used by the Creator to renew the world, to lead it to salvation, to make it a more perfect instrument of divine glory.

Almost twenty-five years ago, the Fathers of the Second Vatican Council, reflecting on the Church in the modern world, declared that men and women, serving their families and the community in their ordinary occupations, were entitled to look upon their work as “a prolongation of the work of the Creator ... and as their personal contribution to the fulfilment in history of the divine plan.”<sup>1</sup>

As the Council Fathers looked to the future and tried to discern the context in which the Church would be called upon to carry out her mission, they could clearly see that the progress of technology was already “transforming the face

---

\* Message of the Holy Father John Paul II for the 24th World Communications Day 27 May 1990 (Vatican, 24 January 1990). The footnotes have been given by the editors. For the original text see The Holy See, [https://www.vatican.va/content/john-paul-ii/en/messages/communications/documents/hf\\_jp-ii\\_mes\\_24011990\\_world-communications-day.html](https://www.vatican.va/content/john-paul-ii/en/messages/communications/documents/hf_jp-ii_mes_24011990_world-communications-day.html).

© Copyright 1990 by Dicastero per la Comunicazione—Libreria Editrice Vaticana.

<sup>1</sup> Second Vatican Council, Pastoral Constitution on the Church in the Modern World, *Gaudium et Spes*, Section 34, The Holy See, [https://www.vatican.va/archive/hist\\_councils/ii\\_vatican\\_council/documents/vat-ii\\_const\\_19651207\\_gaudium-et-spes\\_en.html](https://www.vatican.va/archive/hist_councils/ii_vatican_council/documents/vat-ii_const_19651207_gaudium-et-spes_en.html).

of the earth” and even reaching out to conquer space.<sup>2</sup> They recognized that developments in communications technology, in particular, were likely to set off chain reactions with unforeseen consequences.

Far from suggesting that the Church should stand aloof or try to isolate herself from the mainstream of these events, the Council Fathers saw the Church as being in the very midst of human progress, sharing the experiences of the rest of humanity, seeking to understand them and to interpret them in the light of faith. It was for God’s faithful people to make creative use of the new discoveries and technologies for the benefit of humanity and the fulfilment of God’s plan for the world.

This recognition of rapid change and this openness to new developments have proved timely in the years that followed, for the pace of change and development has continued to accelerate. Today, for example, one no longer thinks or speaks of social communications as mere instruments or technologies. Rather they are now seen as part of a still unfolding culture whose full implications are as yet imperfectly understood and whose potentialities remain for the moment only partly exploited.

Here we find the basis for our reflections on this Twenty-fourth World Communications Day. With each day that passes the vision of earlier years becomes ever more a reality. It was a vision which foresaw the possibility of real dialogue between widely-separated peoples, of a worldwide sharing of ideas and aspirations, of growth in mutual knowledge and understanding, of a strengthening of brotherhood across many hitherto insurmountable barriers.<sup>3</sup>

With the advent of computer telecommunications and what are known as computer participation systems, the Church is offered further means for fulfilling her mission. Methods of facilitating communication and dialogue among her own members can strengthen the bonds of unity between them. Immediate access to information makes it possible for her to deepen her dialogue with the contemporary world. In the new “computer culture” the Church can more readily inform the world of her beliefs and explain the reasons for her stance on any given issue or event. She can hear more clearly the voice of public opinion, and enter into a continuous discussion with the world around her, thus involving herself more immediately in the common search for solutions to humanity’s many pressing problems.<sup>4</sup>

It is clear that the Church must also avail herself of the new resources provided by human exploration in computer and satellite technology for her

---

<sup>2</sup> See *ibidem*, Section 5.

<sup>3</sup> See Pontifical Council on Social Communications, Instruction on the Means of Social Communication *Communio et Progressio*, Sections 181–182, The Holy See, [https://www.vatican.va/roman\\_curia/pontifical\\_councils/pcscs/documents/rc\\_pc\\_pcscs\\_doc\\_23051971\\_communio\\_en.html](https://www.vatican.va/roman_curia/pontifical_councils/pcscs/documents/rc_pc_pcscs_doc_23051971_communio_en.html).

<sup>4</sup> See *ibidem*, Sections 114–115.



ever pressing task of evangelization. Her most vital and urgent message has to do with knowledge of Christ and the way of salvation which he offers. This is something she must put before the people of every age, inviting them to embrace the Gospel out of love, ever mindful that “truth cannot impose itself except by virtue of its own truth, which wins over the mind with both gentleness and power.”<sup>5</sup>

As the wisdom and insights of past years teach us: “God has spoken to humanity according to the culture proper to each age. Similarly the Church, which in the course of time has existed in varying circumstances, has utilized the resources of different cultures in her preaching to spread and explain the message of Christ.”<sup>6</sup> “The first proclamation, catechesis or the further deepening of faith cannot do without the (means of social communication). ... The Church would feel guilty before the Lord if she did not utilize these powerful means that human skill is daily rendering more perfect. It is through them that she proclaims ‘from the housetops’ the message of which she is the depositary.”<sup>7</sup>

Surely we must be grateful for the new technology which enables us to store information in vast man-made artificial memories, thus providing wide and instant access to the knowledge which is our human heritage, to the Church’s teaching and tradition, the words of Sacred Scripture, the counsels of the great masters of spirituality, the history and traditions of the local Churches, of Religious Orders and lay institutes, and to the ideas and experiences of initiators and innovators whose insights bear constant witness to the faithful presence in our midst of a loving Father who brings out of his treasure new things and old (see Mt 13:52).

Young people especially are readily adapting to the computer culture and its “language.” This is surely a cause for satisfaction. Let us “trust the young.”<sup>8</sup> They have had the advantage of growing up with the new developments, and it will be their duty to employ these new instruments for a wider and more intense dialogue among all the diverse races and classes who share this “shrinking globe.” It falls to them to search out ways in which the new systems of data conservation and exchange can be used to assist in promoting greater universal

---

<sup>5</sup> Second Vatican Council, Declaration on Religious Freedom *Dignitatis Humanae: On the Right of the Person and of Communities to Social and Civil Freedom in Matters Religious*, Section 1, The Holy See, [https://www.vatican.va/archive/hist\\_councils/ii\\_vatican\\_council/documents/vat\\_ii\\_decl\\_19651207\\_dignitatis-humanae\\_en.html](https://www.vatican.va/archive/hist_councils/ii_vatican_council/documents/vat_ii_decl_19651207_dignitatis-humanae_en.html).

<sup>6</sup> *Gaudium et Spes*, Section 58.

<sup>7</sup> P a u l VI, Apostolic Exhortation *Evangelii Nuntiandi*, Section 45, The Holy See, [https://www.vatican.va/content/paul-vi/en/apost\\_exhortations/documents/hf\\_p-vi\\_exh\\_19751208\\_evangelii-nuntiandi.html](https://www.vatican.va/content/paul-vi/en/apost_exhortations/documents/hf_p-vi_exh_19751208_evangelii-nuntiandi.html).

<sup>8</sup> *Communio et Progressio*, Section 70.

justice, greater respect for human rights, a healthy development for all individuals and peoples, and the freedoms essential for a fully human life.

Whether we are young or old, let us rise to the challenge of new discoveries and technologies by bringing to them a moral vision rooted in our religious faith, in our respect for the human person, and our commitment to transform the world in accordance with God's plan. On this World Communications Day, let us pray for wisdom in using the potential of the "computer age" to serve man's human and transcendent calling, and thus to give glory to the Father from whom all good things come.

SZTUCZNA INTELIGENCJA PODMIOTEM?



Zbigniew WRÓBLEWSKI  
Paweł FORTUNA

## MORAL SUBJECTIVITY AND THE MORAL STATUS OF ARTIFICIAL INTELLIGENCE A Philosophical and Psychological Perspective

*The history of morality and ethical reflection shows that inclusion in the moral community has been a process of expanding the set of its members to include one's family, group, tribe, nation, race, and through non-human sentient beings (animals), to contemporary proposals for the inclusion of species, ecosystems, the biosphere, and artifacts. Importantly, some of the solutions proposed in the literature address the moral issues related to AI in the context of the next stage of the historical process of forming the moral community and ethical reflection on the mechanisms and rules of this process.*

In 2018, the Rector of the AGH University of Science and Technology gave a student's book to the humanoid robot Sophia during the Impact digital economy conference held in Krakow. This robot was developed by the Hong Kong-based company Hanson Robotics and was activated in 2016. In the Anthropomorphic Robot Database (ABOT) with 251 robots, Sophia ranks 8th for human likeness, 9th for surface appearance, and 31st for face appearance. The event was widely commented on in the media, and one post, published on the official TVPInfo website, read: "Android has dreams of having a family and friends, as well as striving to integrate humans with robots. In an interview with 'Khaleej Times Dubai,' the robot said that she wants to have a child, a daughter, and is seriously thinking of becoming a knowledge ambassador at the foundation of the Prime Minister of the United Arab Emirates".<sup>1</sup> In 2017, the fembot Sophia received the status of a citizen in Saudi Arabia,<sup>2</sup> and it was not a one-off event with regard to artificial systems. Less than a week later, Japan granted the resident status to a chatbot named Mirai<sup>3</sup>. The expansion of artificial intelligence (AI) driven artifacts continues. There is currently a lively discussion on the opportunities and threats of ChatGPT, an advanced language model that can be used for content creation, translation, learning support, idea generation, entertainment, acting as

<sup>1</sup> See Paweł Fortuna, *Optimum: Idea cyberpsychologii pozytywnej* (Warszawa: PWN, 2021), X.

<sup>2</sup> See Andrew Griffin, "Saudi Arabia Grants Citizenship to a Robot for the First Time Ever," Independent UK, October 26, 2017, <https://www.independent.co.uk/life-style/gadgets-and-tech/news/saudi-arabia-robot-sophia-citizenship-android-riyadh-citizen-passport-future-a8021601.html>.

<sup>3</sup> See Rosie McCaill, "Japan Has Just Granted Residency to an AI Bot in a World First," IFL-Science, November 7, 2017, <http://www.iflscience.com/technology/japan-has-just-granted-residency-to-an-ai-bot-in-a-world-first>.

a personal assistant or interpersonal relations trainer. One of the results of the public debate were calls from various scientific and technological communities to suspend work on artificial intelligence for some time.<sup>4</sup>

The digital revolution is densifying the environment with AI-based artifacts, whose image presented in pop culture narratives corresponds to the features of the so-called Artificial General Intelligence (AGI) or Strong AI.<sup>5</sup> These are, considered so far in hypothetical form, systems that have a general intelligence similar to that of an adult human being, not only functioning as if they had a mind, but as if they would actually have one. As a consequence, they would be capable of sensory perception, characterised by awareness, self-awareness and morality. Compared to them, “weak” AI (Artificial Narrow Intelligence) are systems capable of solving problems at the level of human beings or better than them, as if they had a mind and were thinking. The aim of the pursuit of AGI is to create a learning artificial intelligence, not limited to solving problems and concentrating on performing pre-programmed tasks, but capable of developing features corresponding to human intelligence. For this reason, the issue of the human-like status of AGIs<sup>6</sup> and their treatment as moral subjects is lively debated.<sup>7</sup>

AGI is merely a set of assumptions about possible forms of AI. The moment of its emergence cannot be determined,<sup>8</sup> but, as presented at the beginning, artificial systems that are examples of “weak” AI are assigned the attributes of AGI. The tendency to anthropomorphise these artifacts is stimulated by the achievements of both designers and marketing specialists. For example, the design of social robots is aimed at maximising the positive affect of the people who interact with them.<sup>9</sup> This is to be enabled by two paths of innovation: developmental cybernetics (developing human-like entities by simulating human psychological processes and kinesthetic functions) and developmental robotics (development of neural networks that would allow artificial entities to autono-

<sup>4</sup> See Mateusz Nowak, „Elon Musk i założyciel Apple apelują o wstrzymanie prac nad AI. ‘Ultra-terroli nad cywilizacją.’” <https://android.com.pl/tech/581815-apel-o-wstrzymanie-prac-nad-ai/>.

<sup>5</sup> See Cassio Pennachin and Ben Goertzel, “Contemporary Approaches to Artificial General Intelligence,” in *Artificial General Intelligence. Cognitive Technologies*, ed. by Cassio Pennachin, Ben Goertzel (Berlin, Heidelberg: Springer, 2007), 1–30; John R. Searle, “Minds, Brains, and Programs,” *Behavioral and Brain Sciences* 3, no. 3 (1980): 417–24.

<sup>6</sup> See Kamil Muzyka, “The Basic Rules for Coexistence: The Possible Applicability of Meta-law for Human-AGI Relations,” *Paladyn, Journal of Behavioral Robotics* 11, no. 1 (2020): 104–17.

<sup>7</sup> See Mark Coeckelbergh, “Virtual Moral Agency, Virtual Moral Responsibility: On the Moral Significance of the Appearance, Perception, and Performance of Artificial Agents,” *AI & Society* 24, no. 2 (2009): 181–89.

<sup>8</sup> See Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence* (New York: Alfred A. Knopf, 2017).

<sup>9</sup> See Elyakim Kisilev, *Relationships 5.0: How AI, VR, and Robots Will Reshape Our Emotional Lives* (Oxford: Oxford University Press, 2022).



mously acquire sensorimotor and mental abilities of increasing complexity).<sup>10</sup> Artificial systems are given human names (e.g., chatbot Alexa, supercomputer IBM Watson, Ernest—UK Facebook messenger bank aggregator), appearance, and are also assigned self-awareness<sup>11</sup> and awareness (e.g., LaMDA, Google’s artificially intelligent chatbot generator).<sup>12</sup> Assigning artificial systems almost equal status to human beings is also facilitated by the increasingly stronger voice of supporters of post-humanism, who question the dualism between nature/culture, human being/animal and human being/machine and argue for the empowerment of non-human forms of life.<sup>13</sup> For example, the post-humanist vision of business assumes an equal cooperation of natural human beings, cyborgs (neuroprosthethically augmented human employees), computers (artificial intelligence driven software) and bioroids (humanoid robots).<sup>14</sup>

The trends outlined above raise questions: (1) of a philosophical nature—Can potential AGI objects become moral subjects?; What criteria for assigning the moral status (MS) apply to AI-driven objects? and (2) of a psychological nature—At the level of belief, do AI-based artificial systems have an open path to inclusion in the moral community?; What factors determine the consent to assign the MS to AI-based systems?

When seeking answers to such questions in this paper, we first outline the broader context that is the reception of AI and the anthropomorphisation of artifacts based on it, then we consider the issue of moral subjectivity, and finally we present the results of psychological studies on assignment of the MS to artificial systems. We treat the discussion of the outlined problems as an element of preparing a public debate on the moral aspects of the rapid development of AI, which systematically expands the scope and possibilities of simulating mental processes. The assignment of appropriate MS to AI has practical consequences: AI objects can be autonomous moral subjects, and so can have moral responsibility,<sup>15</sup> can

---

<sup>10</sup> See Antonella Marchetti et al., “Theory of Mind and Humanoid Robots from a Lifespan Perspective,” *Zeitschrift für Psychologie* 226, no. 2 (2018): 98–109.

<sup>11</sup> See Selmer Bringsjord, Paul Bello, and Naveen Sundar Govindarajulu, “Toward Axiomatizing Consciousness,” in *The Bloomsbury Companion to the Philosophy of Consciousness*, ed. Dale Jacquette (London: Bloomsbury Academic, 2018), 289–324.

<sup>12</sup> See Nitasha Tikku, “The Google Engineer Who Thinks the Company’s AI Has Come to Life,” *The Washington Post*, June 11, 2022, <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lambda-blake-lemoine>.

<sup>13</sup> See Rosi Bradiotti, *The Posthuman* (Cambridge: Polity Press Ltd, 2013); Joshua C. Gellers, *Rights for Robots: Artificial Intelligence, Animal and Environmental Law* (New York: Routledge, 2021).

<sup>14</sup> See Matthew E. Gladden, *Posthuman Management* (Indianapolis: Synthynion Press, 2016).

<sup>15</sup> See Aimee van Wynsberghe and Scott Robbins, “Critiquing the Reasons for Making Artificial Moral Agents,” *Science and Engineering Ethics* 25, no. 3 (2019): 719–35; Mariarosaria

be subjects of rights, and thus other moral subjects have obligations in relation to them,<sup>16</sup> can be moral agents that make moral decisions,<sup>17</sup> and can be members of human communities.<sup>18</sup> The functioning of AI objects in the social environment creates a new situation for common morality. Regardless of the opinion of programmers (whether or not it is a conscious machine) and philosophers (whether or not to assign the MS to new objects), common sense judgments are already being spontaneously formulated influencing the way we treat artificial entities and the hybrid systems we create with them.

### AI AS A POTENTIAL MORAL SUBJECT

When considering the moral subjectivity and the MS of AI, it is necessary to clarify the very concept of AI, which has been a challenging task since its introduction,<sup>19</sup> and some researchers believe that this is an unrealistic goal at the current stage of research.<sup>20</sup> At a general level, there is consensus that AI is the attempt “to make a computer work like a human mind”.<sup>21</sup> According to Lindes, the concept “artificial intelligence” should be used in two main senses, which the researcher labels as AI1 and AI2. According to him, AI1 refers to the quality of intelligence in the man-made computing systems, which can be compared and contrasted with natural intelligence.<sup>22</sup> AI2, on the other hand, is a field of study that deals with the design, construction and evaluation of AI1 systems, i.e. artificial systems that manifest intelligence. Because the definition

---

T a d d e o and Luciano F l o r i d i, “How AI Can Be a Force for Good,” *Science* 361, no. 6404 (2018): 751n.

<sup>16</sup> See David J. G u n k e l, “The Other Question: Can and Should Robots Have Rights?,” *Ethics and Information Technology* 20 (2018): 87–99; Jacob T u r n e r, *Rights for AI*, in Turner, *Robot Rules* (Cham: Palgrave Macmillan, 2019), 133–71.

<sup>17</sup> See John D a n a h e r, *Automation and Utopia: Human Flourishing in a World without Work* (Harvard University Press: Cambridge, MA, 2019); Colin A l l e n, Iva S m i t, and Wendell W a l l a c h, “Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches,” *Ethics and Information Technology* 7 (2005): 149–55; James H. M o o r, “The Nature, Importance, and Difficulty of Machine Ethics,” *IEEE Intelligent Systems* 21, no. 4 (2006): 18–21.

<sup>18</sup> See Migle L a u k y t e, “Artificial Agents among Us: Should We Recognize Them as Agents Proper?,” *Ethics and Information Technology* 19, no. 1 (2017): 1–17; Brian D u f f y, “Anthropomorphism and the Social Robot,” *Robotic and Autonomous Systems* 42, nos. 3–4 (2003): 177–90.

<sup>19</sup> See Nils N i l s o n, *The Quest for Artificial Intelligence* (Cambridge: Cambridge University Press, 2009).

<sup>20</sup> See Pei W a n g, “On Defining Artificial Intelligence,” *Journal of Artificial General Intelligence* 10, no. 2 (2019): 1–37.

<sup>21</sup> Pei W a n g, Kai L i u, and Quinn D o u g h e r t y, “Conceptions of Artificial Intelligence and Singularity,” *Information* 9, no. 4 (2018): 79.

<sup>22</sup> See Peter L i n d e s, “Intelligence and Agency,” *Journal of Artificial General Intelligence* 11, no. 2 (2020): 47–49.

of AI2 depends on how we understand AI1, which in turn depends on how we understand intelligence itself, defining AI depends on precisely defining intelligence as such. This is problematic because only in the field of psychology this term is considered controversial and there is no consensus on a single definition (as many as 28 new ones have been proposed in the previous decade).<sup>23</sup>

Despite the difficulties in defining AI, it is easier to identify AI examples, referred to as rational agents—systems that receive percepts from the environment and perform actions.<sup>24</sup> The agents act as “intelligent tools,” and many of them operate under marketing names (e.g., virtual assistants: Amazon’s Alexa, Apple’s Siri). They are driven by various types of algorithms (e.g., search, machine learning, evolutionary, artificial neural networks), and when combined with a physical body they become examples of “embodied” AI (e.g., self-driving cars, robots). Human beings, when interacting with AI-based systems, intentionally or unknowingly create hybrid systems.<sup>25</sup> The degree of fusion with artificial entities can be described on a continuum of cyborgisation: from interaction with static (PC), mobile (smartphone) and wearable technologies (smart-glasses), to augmentation (fusing artifacts with the human nervous system).<sup>26</sup> This fusion can be explicit, as in the case of human-cobot systems in the production process, but it can also be implicit to the user of the technology. An example of this are the algorithms that control the mathematical and statistical representation of each Internet user, which, according to Deleuze, can be described as “dividual.”<sup>27</sup> This bank of data is created by the activity of the Internet user, but his “mind” is made up of algorithms beyond his control and suggesting customised content. Reacting to it makes the human being (individual) and the “dividual” function in a continuous feedback loop, providing data and reacting to it. They unknowingly meld together to form a kind of augmented mind,<sup>28</sup> which can be referred to in a working way as the “hybrid self.”

---

<sup>23</sup> See Dagmar M o n e t t and Colin W.P. L e w i s, “Getting Clarity by Defining Artificial Intelligence: A Survey,” in *Philosophy and Theory of Artificial Intelligence 2017*, ed. Vincent C. Müller (Berlin: Springer, 2018), 212–14.

<sup>24</sup> Stuart J. R u s s e l l and Peter N o r v i g, *Artificial Intelligence: A Modern Approach* (Boston: Pearson, 2020).

<sup>25</sup> See Wulf L o h and Janina L o h, “Autonomy and Responsibility in Hybrid Systems: The Example of Autonomous Cars,” in *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*, eds. Patrick Lin, Keith Abney and Ryan Jenkins (New York: Oxford University Press, 2017), 35–50.

<sup>26</sup> Alex J u p i t e r, “The Human-Cyborg Continuum: Why AI Is Pointless and Why We Should All Become Cyborgs Instead,” June 4, 2016, <https://medium.com/@AlexJupiter/the-human-cyborg-continuum-why-ai-is-pointless-and-why-we-should-all-become-cyborgs-instead-4de0c4bb476f>.

<sup>27</sup> See Gilles D e l e u z e, “Postscript on the Societies of Control,” *October* 59 (1992): 3–7.

<sup>28</sup> See Andy C l a r k and David C h a l m e r s, “The Extended Mind,” *Analysis* 58, no. 1(1998): 7–19.

Later in this paper, the results of a study on the assignment of the MS to AI-driven artifacts will be presented. Due to the fact that this process is informal, intuitive reasoning, it is worth looking at how AI is understood by the users themselves. The surveys conducted in seven countries (e.g., USA, Germany, China) show that public awareness of AI seems to depend on the visibility of its use.<sup>29</sup> It was found that 90% of respondents were aware that a voice assistant (visible AI) was based on AI, while only one in three respondents associated online shopping websites, video streaming services and social media (invisible AI) with AI. The obtained results correspond to the results of studies, which captured differences in the understanding of AI by experts (IT specialists) and laymen.<sup>30</sup> As it turns out, for people with expert knowledge, AI is primarily “algorithmic systems” (e.g. image generation algorithm), while for laymen, it is mainly “nature imitating systems” (e.g. humanoid robot). When categorising AI examples, experts are mainly guided by functional features, while laymen also consider structural features of the systems. The functions of “algorithmic systems” are cognitive, related to performing the so-called objective tasks (e.g., pattern recognition), while “nature imitating systems” perform tasks that seem subjective in nature (based on emotions and intuition).<sup>31</sup>

The identification of AI with embodied, imitating entities found in nature should be attributed to contact with “AI narratives” present in pop culture, which include “portrayals of any machines (or hybrids, such as cyborgs) to which intelligence has been ascribed, which can include representations under terms such as robots, androids or automata.”<sup>32</sup> Some narratives are non-fictional (e.g., TV news) and some are fictional (e.g., sci-fi films). In non-fiction AI narratives, attention is paid mainly to the examples of “weak” AI, while the heroes of fiction AI narratives are the examples of AGI. In the latter case, they not only talk and walk, but are capable of feeling human emotions, have elements of self-awareness and free will. In addition, they are characterised by exaggerated corporeality (e.g., T-800 in *Terminator*, 1984) and hypersexuality (e.g., Ava in *Ex Machina*, 2015), they have superhuman resistance to pain and indestructibility.<sup>33</sup>

<sup>29</sup> Jem Davies, “AI Today, AI Tomorrow. The Arm 2020 Global AI Survey,” armBlueprint, February 3, 2020, <https://www.arm.com/resources/report/ai-today-ai-tomorrow-ty>.

<sup>30</sup> See Paweł Fortuna and Oleg Gorbanjuk, “What Is Behind the Buzzword for Experts and Laymen: Representation of ‘Artificial Intelligence’ in the IT-professionals’ and Non-professionals’ Minds,” *Europe’s Journal of Psychology* 8, no. 2 (2022): 207–18.

<sup>31</sup> See Yoel Inbar, Jeremy Cone, and Thomas Gilovich, “People’s Intuitions about Intuitive Insight and Intuitive Choice,” *Journal of Personality and Social Psychology* 99, no. 2 (2010): 232–47.

<sup>32</sup> Stephen Cave, Kanta Dihal, and Sarah Dillon, *AI Narratives: A History of Imaginative Thinking about Intelligent Machines* (Oxford: Oxford University Press, 2020), 5.

<sup>33</sup> Davies, “AI Today, AI Tomorrow: The Arm 2020 Global AI Survey.”

## MORAL STATUS VS MORAL SUBJECTIVITY

The discussion on the moral aspects of AI has been ongoing since it was launched in the 1950s as part of the so-called ethics of AI.<sup>34</sup> The issues raised there focused, among other things, on the threat to privacy, information technology surveillance, the use of knowledge about citizens, the use of information to manipulate people, freedom of citizens, and civil rights. The common denominator of the issues mentioned is the assessment of the (actual, potential) effects of the use of AI on moral subjects (human beings). However, what is morally assessed is the usefulness of the artifacts, not the artifacts themselves.<sup>35</sup>

In parallel, another type of moral reflection was being developed, which focused on the potential AGI objects.<sup>36</sup> If we hypothetically assume that the realisation of such attributes as reasoning, decision-making, representing knowledge, planning, learning, or communicating in a natural language makes it possible to achieve the level of artificial awareness analogous to the awareness of a human being, then the problem of the MS of these entities arises. It is no longer just a question of a moral assessment of the effects of using AI, but whether AI has a moral significance that is based in itself and not in its technical usefulness, and whether human beings have any obligations towards it. This raises the problem of defining new boundaries of morality, or more precisely of shifting the boundaries of the moral community to include new morally relevant beings, and thus the question: can the moral community be extended to include AI-driven entities?

The moral community includes beings towards whom moral subjects have certain obligations. The inclusion in it is made by recognising the MS of a member of the community on the basis of having a specific feature or set of features. The history of morality and ethical reflection shows that inclusion in the moral community has been a process of expanding the set of its members to include one's family, group, tribe, nation, race, and through non-human

---

<sup>34</sup> Norbert Wiener, *The Human Use of Human Beings: Cybernetics and Society* (Boston: Houghton Mifflin, 1950).

<sup>35</sup> See Nick Bostrom and Eliezer Yudkowsky, "The Ethics of Artificial Intelligence," in *The Cambridge Handbook of Artificial Intelligence*, eds. Keith Frankish, William Ramsey (Cambridge: Cambridge University Press, 2014), 316–34; Kenneth E. Himma, "Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to Be a Moral Agent?," *Ethics and Information Technology* 11, no. 1(2009): 19–29; Vincent C. Müller, "Is it Time for Robot Rights? Moral Status in Artificial Entities," *Ethics and Information Technology* 23 (2021): 579–87.

<sup>36</sup> Vincent C. Müller, "Ethics of Artificial Intelligence and Robotics," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, <https://plato.stanford.edu/archives/sum2021/entries/ethics-ai/>.

sentient beings (animals), to contemporary proposals for the inclusion of species, ecosystems, the biosphere, and artifacts. The continuous expansion of the moral community is referred to as the “tower of morality”<sup>37</sup> or “expanding circle.”<sup>38</sup> Importantly, some of the solutions proposed in the literature address the moral issues related to AI in the context of the next stage of the historical process of forming the moral community and ethical reflection on the mechanisms and rules of this process.<sup>39</sup> The concept necessary to describe it is the concept of MS.

According to Warren, “the concept of moral status is, rather, a means of specifying those entities towards which we believe ourselves to have moral obligations, as well as something of what we take those obligations to be.”<sup>40</sup> In turn, Kamm proposes the following definition of MS: “X has moral status = because X counts morally in its own right, it is permissible/impermissible to do things to it for its own sake.”<sup>41</sup> The concept of MS has various functions. The concept of MS makes it possible to define the general obligations that moral subjects should fulfil in relation to beings of a given type, so it can be used to define the basic standards of acceptable behaviour towards them. The core features of the concept of MS are thus the generality (of obligations, rights, interests), and the fact that the MS is assigned to members of a specific group rather than to individuals (e.g., primates, not just Sarah the chimpanzee). The moral obligations arising from the assignment of the MS are the obligations towards that being, not someone else (e.g., towards the robot ASIMO, not its legal owner)<sup>42</sup>. The concept of MS can also justify moral ideals, e.g. the Christian ideal of loving one’s neighbour or the Jainist ideal of not killing. Such ideals, creating space for supererogation, encourage moral development.

The MS is assigned to entities on the basis of meeting the relevant criteria, although their determination is a source of dispute. In the debate, the views of supporters of single- and multi-criteria theories of MS clash. The single-criteria theories postulate the identification of a single intrinsic characteristic of the

<sup>37</sup> Frans de Waal, *Primates and Philosophers: How Morality Evolved* (Princeton: Princeton University Press, 2006).

<sup>38</sup> Peter Singer, *The Expanding Circle: Ethics and Sociobiology* (Oxford: Clarendon Press, 1981); Steve Torrance, “Artificial Agents and the Expanding Ethical Circle,” *AI and Society* 28, no. 4 (2013): 399–414.

<sup>39</sup> See Adam J. Andreotta, “The Hard Problem of AI Rights,” *AI and Society* (2020): 1–14; Himma, “Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent have to Be a Moral Agent?,” 19–29.

<sup>40</sup> Mary A. Warren, *Moral Status: Obligations to Persons and Other Living Things* (Oxford: Clarendon Press, 1997), 9.

<sup>41</sup> Frances M. Kamm, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm* (New York: Oxford University Press, 2007), 7.

<sup>42</sup> See Warren, *Moral Status*, 10.



entity, the possession of which guarantees the MS and inclusion in the moral community, e.g., life,<sup>43</sup> capacity to feel<sup>44</sup> and subjectivity (being a person).<sup>45</sup> The latter of the listed features can be understood: (a) restrictively—the subject should possess certain cognitive capacities that enable one to reflect on moral issues, which makes it possible to be a moral subject, or (b) less restrictively: the entity should be a subject of life, possessing beliefs, desires, memory, the capacity to anticipate and to act intentionally.<sup>46</sup> The literature also indicates proposals for identifying the MS on the basis of external features of subjects (relational features: individual—community, environment), e.g., the moral status of a given being depends on the function (positive or negative) it performs within a biological or social community.<sup>47</sup> In another proposal, it is assumed that the MS of a given entity depends on the feelings we have towards it, e.g. our care for a given entity assigns it the MS.<sup>48</sup> Each of the listed features (internal or external) is treated by philosophers as a necessary and sufficient condition for having the MS.

According to multi-criteria theories of MS, it is assumed that (1) there is more than one valid criterion for the MS, (2) there is more than one type of the MS, and (3) the criterion for having the MS takes into account the internal and external features of a given entity.<sup>49</sup> The general principles arising from the assignment of the MS are interdependent, i.e., the practical consequences resulting from one principle are understood in the context of the other principles. The adoption of this approach is dictated by the fact that many moral problems are more complex in nature than it appears to the supporters of single-criteria theories. The common-sense diversity of moral intuitions on complex or radically new issues (humanoid robots as moral agents) seems to support this approach. The strategy of accepting a variety of criteria for the MS (and thus accepting many of its types) and ordering them in the form of a system (as suggested by the multi-criteria theory) seems to be more optimal

---

<sup>43</sup> Albert S c h w e i t z e r, *Civilization and Ethics* (London: Adam & Charles Black, 1955).

<sup>44</sup> Jeremy B e n t h a m, *An Introduction to the Principles of Morals and Legislation* (Oxford: Oxford University Press, 1998); Peter S i n g e r, *Animal Liberation: A New Ethics for Our Treatment of Animals* (New York: Harper Collins, 1975).

<sup>45</sup> Immanuel K a n t, *The Metaphysics of Morals* (Cambridge: Cambridge University Press, 2017).

<sup>46</sup> Tom R e g a n, *The Case for Animals Rights* (Berkeley, Los Angeles: University of California Press, 1983).

<sup>47</sup> Aldo L e o p o l d, *A Sand County Almanac* (Oxford: Oxford University Press, 1987); John B. C a l l i c o t t, *In Defense of the Land Ethic: Essays in Environmental Philosophy* (New York: State University of New York Press, 1989).

<sup>48</sup> Nel N o d d i n g s, *Caring: A Feminine Approach to Ethics and Moral Education* (Berkeley, Los Angeles: University of California Press, 2013).

<sup>49</sup> See W a r r e n, *Moral Status*, 21.

than the strategy of reducing a variety of criteria to a single, key one (the single-criteria theory)—the moral community is diverse in terms of the MS of its members (a heterogeneous, pluralistic community).

We recognise that it is better to use the multi-criteria theory for determining the MS of AI-driven artifacts. It allows to take into account the new nature of objects and their moral significance. While various uses of AI have become the subject of numerous ethical, social and psychological studies in the scheme—what is the impact of using AI in domain X—the problem of the MS of AI (selected objects) solved in the scheme—what moral significance AI objects have—is still an under-researched problem. Furthermore, the variety of types of AI-driven objects (algorithms, autonomous machines, expert systems, humanoid robots) makes it impossible to use a single criterion of the MS (e.g., an AGI and a chess-playing programme), and therefore a pluralistic strategy (multiple features and different degrees of the MS) is better than using a zero-one strategy (only one feature, e.g., the ability to feel pain, determines whether an object has or does not have the MS). For this reason, as suggested by Warren, we considered seven criteria of identification of MS that relate to potentially internal and external characteristics of HR: (1) being a living being (structured purposeful systems, showing the basic attributes of life); (2) being a sentient being; (3) being an individual with cognitive abilities that enable reflection on moral problems; (4) being a person (subject of life) who has beliefs, desires, memory, the ability to predict and act intentionally; (5) being a significant part of the environment; (6) being a member of an interspecies community, and (7) being recognized as a significant entity by another moral entity. Each of the above-mentioned features is related to one of the moral principles that define the normative consequences of assigning MS: (1) the principle of respect for life; (2) the principle against cruelty; (3) the principle of the rights of the subject; (4) the principle of human rights; (5) the environmental principle; (6) the interspecific principle; and (7) the principle of the transitivity of respect.<sup>50</sup>

#### MIND PERCEPTION AND MORAL STATUS OF AI

Assigning AI-driven systems the moral subjectivity and the MS, like ascribing human attributes to them, is a manifestation of anthropomorphisation (Greek: *anthropos* for “human,” *morphe* for “shape” or “form”<sup>51</sup>). It is an automatic process, built into perception of the surroundings, and the degree of

<sup>50</sup> See *ibidem*, 148–70.

<sup>51</sup> See Nicholas Epley, Adam Waijtz, and John T. Cacioppo, “On Seeing a Human: A Three-factor Theory of Anthropomorphism,” *Psychological Review* 114, no. 4 (2007): 864–86.



anthropomorphizing can be determined on a continuum from the superficial and habitual use of personifying word labels to assigning them human dispositions, including emotions, thinking, and free will.<sup>52</sup> The triggering factor is the presence of typical human features in the encountered entity. Noticing them activates the knowledge about a human being stored in memory, and then integrates it with information about this person.

From the psychological point of view, the process of transmitting MS is coupled with the process of mind perception.<sup>53</sup> Research shows that our cognitive apparatus uses a two-dimensional filter in the process of mind perception, also referred to as “the cognitive template for morality.”<sup>54</sup> In a classic study, the participants’ task was to compare pairs of thirteen characters according to one of eighteen attributes or one of six personal judgments. The characters compared were humans (e.g., a five-month-old infant, adult woman, human in a vegetative state, test subject), animals (frog, domestic dog, wild chimpanzee), as well as a dead woman, God (defined as the creator of the universe and the ultimate source of knowledge, power, and love) and a robot Kismet.<sup>55</sup> The set of attributes included, among others, the feeling of pain, personality, awareness, morality, memory, and reflection. It has been revealed that a person evaluates other individuals on two dimensions: Experience (the ability to feel suffering) and Agency (the ability to take intentional actions). What is important, they were related to the classical distinction between individuals as moral patient and moral agent introduced by Aristotle.<sup>56</sup> Accordingly, the character’s high assessment of the experience dimension (also referred to as the ability to feel suffering) indicates that we are dealing with the so-called a moral patient. On the other hand, a similar assessment on the dimension of agency indicates that the character is a moral agent.

The revealed dimensions of mind perception have been confirmed in many studies,<sup>57</sup> although there are analyses in which three dimensions have been

---

<sup>52</sup> See Marina Puzakova, Hyokjin Kwak, and Joseph F. Rocereto, “Pushing the Envelope of Brand and Personality: Antecedents and Moderators of Anthropomorphized Brands,” *Advances in Consumer Research* 36 (2009): 413–20.

<sup>53</sup> See Heather M. Gray, Kurt Gray, and Daniel M. Wegner, “Dimensions of Mind Perception,” *Science* 315, no. 5812 (2007): 619.

<sup>54</sup> See Gray, Young, and Wyatt, “Mind Perception is the Essence of Morality,” *Psychological Inquiry* 23, no. 2 (2012): 101–24.

<sup>55</sup> See Gray, Gray, and Wegner, “Dimensions of Mind Perception,” 619.

<sup>56</sup> Aristotle, *Nicomachean Ethics*, trans. William D. Ross (New York: World Library Classics, 2009).

<sup>57</sup> See Imge Saltik, Deniz Erdil, and Burcu A. Urgan, “Mind Perception and Social Robots: The Role of Agent Appearance and Action Types,” in *HRI’21: Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction: March 8-11, 2021, Boulder, CO, USA* (New York, NY: Association for Computing Machinery, [2021]), 210–214; Aleksandra

revealed, which argues for the heterogeneous nature of the Agency dimension.<sup>58</sup> With a two-dimensional mind perception matrix at our disposal, we can classify all the characters we encounter into one of four categories (examples taken from a classic study):

(1) low “Experience” and low “Agency”—e.g., a dead woman—she can neither be inflicted with suffering nor be expected to behave intentionally;

(2) low “Experience” and high “Agency”—e.g., God, robot—it cannot be inflicted with suffering but can be expected to behave intentionally;

(3) high “Experience” and low “Agency”—a frog—it can be inflicted with suffering but cannot be expected to behave intentionally;

(4) high “Experience” and high “Agency”—the subject itself—the subject can both be inflicted with suffering and be expected to behave intentionally.

Although participants of the referred studies tend to locate a social robot in the same group as God, however, as can be seen, the assessment of the intentionality of these characters is closer to the location of a dead woman, chimpanzee, and dog than adult humans and the subject himself. A similar low position of artificial systems on the Experience and Agency dimensions was reported by Lukaszewicz-Alcaraz and Fortuna.<sup>59</sup> This involved the social robot Pepper, the algorithm Aaron used for artistic realisations and the fembot Ai-Da advertised as an AI-driven artificial artist. However, it should be remembered that the position of these agents in the perceptual space of mind should not, however, be regarded as unchanging. It has been noted that humanoid robot’s ratings on these dimensions may change depending on their appearance and response. For example, presenting robot with a human-like face places it higher on the Experience dimension than the same agent with exposed electronic components not covered with synthetic leather and higher scores on the Agency dimension when the robot performed an activity of a communicative nature

---

Lukaszewicz and Paweł Fortuna, “Towards Turing Test 2.0—Attribution of Moral Status and Personhood to Human and Non-Human Agents,” *Postdigital Science and Education* 4 (2022): 860–76; Paweł Fortuna, Arkadiusz Gut, and Zbigniew Wróblewski, “Hey Robot, the Mind Is Not Enough to Join the Moral Community! The Effect of Assigning a Mind and a Soul to Humanoid Robot on Its Moral Status,” *Annals of Psychology / Roczniki Psychologiczne* (2023), online first, <https://doi.org/10.18290/rpsych2023.0008>.

<sup>58</sup> See Kara Weisman, Carol S. Dweck, and Ellen M. Markman, “Rethinking People’s Conceptions of Mental Life,” *Proceedings of the National Academy of Sciences of the United States of America* 114, no. 43 (2017): 11374–79; Bertram F. Malle, “How Many Dimensions of Mind Perception Really Are There?,” in *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*, ed. Ashok Goel, Colleen Seifert, and Christian Freksa (Montreal, QB: Cognitive Science Society, 2019), 2268–74.

<sup>59</sup> See Lukaszewicz and Fortuna, “Towards Turing Test 2.0—Attribution of Moral Status and Personhood to Human and Non-Human Agents,” 860–76.

(gestures of waving a hand) or presenting the satisfaction of a biological need (placing a cup to the mouth imitating thirst quenching).<sup>60</sup>

The distinction between the moral patient and the moral agent is present in the literature on the MS of animals<sup>61</sup> and artificial agents.<sup>62</sup> Bostrom and Yudkowsky define the Experience as Sentience dimensions, in which Sentience is the capacity for phenomenal Experience or qualia, such as the capacity to feel pain and suffer. They also define Agency as Sapience (a set of capacities associated with higher intelligence, such as self-awareness and being a reason-responsive agent) and state that these criteria are “commonly proposed as being importantly linked to moral status, either separately or in combination.”<sup>63</sup> In the context of the discussion on MS of AI-driven characters, this distinction was adopted by Torrance, according to which the notion of “having ethical status” can be separated into two associated aspects: ethical receptivity and ethical productivity. Ethical recipients are those who stand to benefit from, or are harmed by, the ethical producers, and ethical producers are those who do or do not do their duties, such as saints and murderers.<sup>64</sup> From this perspective, AI and other smart machines can be both ethical producers and ethical recipients.<sup>65</sup>

The relationship between the mind perception and MS dimensions of an artificial AI-driven system has been empirically confirmed.<sup>66</sup> Study participants, who believed in the existence of the mind and the soul, assessed the MS of the humanoid robot Sophia and assigned attributes to it. The tool for assessing the MS was constructed on the basis of a concept distinguishing multiple criteria.<sup>67</sup> It was found that the attribution of the mind and the soul to the robot significantly affected the MS of the robot. Moreover, the dimensions of mind perception acted as a mediator, but only in the mind-MS relationship, while the soul-MS relationship was direct. It is clear from the studies presented that, for those who present a tripartite common anthropology (naïve spiritualists), the assignment of the MS encounters the strong barrier of having to identify the mind and soul of an artificial system. This means that the “cognitive matrix of

---

<sup>60</sup> See Gray, Gray, and Wagner, “Dimensions of Mind Perception,” 619.

<sup>61</sup> Regan, *The Case for Animals Rights*.

<sup>62</sup> See Bostrom and Yudkowsky, “The Ethics of Artificial Intelligence,” 316–34.

<sup>63</sup> *Ibidem*, 322.

<sup>64</sup> See Steve Torrance and D. Roche, “Does an Artificial Agent Need to Be Conscious to Have Ethical?,” in *Technologies on the stand: legal and ethical questions in neuroscience and robotics*, eds. Bibi van den Berg, Laura Klaming (Nijmegen: Wolf Legal Publishers, 2011), 285–310.

<sup>65</sup> See Torrance, “Artificial Agents and the Expanding Ethical Circle,” 399–414.

<sup>66</sup> See Fortuna, Gut, and Wróblewski, “Hey Robot, the Mind Is Not Enough to Join the Moral Community! The Effect of Assigning a Mind and a Soul to Humanoid Robot on Its Moral Status.”

<sup>67</sup> See Warren, *Moral Status*.

morality” enabling the robot to be treated as a moral patient and agent appears to be an insufficient criteria base for assigning the MS.

In order to better understand the psychological determinants of the assignment process of the MS, another study was carried out to investigate the influence of anthropocentric beliefs on the assignment of the MS to the humanoid robot Sophia.<sup>68</sup> At the same time, the mediating effect of ascribing mind and soul to this agent was examined. As in the previous study, the participants were naive spiritualists and, as before, they tended to make the assignment of the MS to an artificial entity dependent on the assignment of mind and soul to it. However, it was noted that such attributions depended on the strength of the respondents anthropocentric beliefs. It was found that the stronger the conviction about the superior status of human beings in relation to other beings, the lower the tendency to attribute mind and soul to the robot. Similar correlations were seen in subsequent studies whose participants responded to the possibility of attributing the MS to a chimpanzee and a cyborg character. Such analyses argue for the need to extend the criteria for assigning the MS to artificial systems (and other entities) and to go beyond the attributes associated with the mind (taken into account in studies on the mind perception). They also raise awareness of the inclusion of subjective factors, such as the type of common anthropology, in the discussion on the assignment of the MS to AI-based systems. It can be predicted that other criteria will be relevant to naive monists, dualists and spiritualists. Perhaps it is the case that those who postulate assigning the MS to artificial systems are physicalists or materialists, nesting the criteria of the MS in the physical substrate of algorithmic systems and the similar human operations they are capable of performing.

\*

In the outlined panorama of philosophical and psychological reflections on the MS of AI-driven objects, it is clear that the key elements of the discussion are the set of criteria that make it possible to assign it. The debate on this topic is complicated by the sheer difficulty in defining AI, and thus in determining the attributes that can be assigned to artifacts driven by it. Intuitions coming from the world of science intertwine with pop culture narratives, creating an ambiguous and illusory picture. AGI is still a futuristic pipe dream, and the spectacular and media-publicised increase in humanoidisation of artificial

---

<sup>68</sup> See Fortuna, Gut, and Wróblewski, “Hey Robot, the Mind Is Not Enough to Join the Moral Community! The Effect of Assigning a Mind and a Soul to Humanoid Robot on Its Moral Status.”

agents still does not make them human beings. Despite the amazing skills presented by artificial systems (e.g., winning in Go, ChatGPT erudition), they do not yet reveal such qualities that in the perception of users could place them high on the dimensions of mind perception and, consequently enable them to be assigned the status of moral patient and moral agent. Not only that, but it appears that supporters of a tripartite common anthropology, i.e., those who base the architecture of human nature on the essence of body, mind and soul, extend the criterion base of the MS to include the spiritual element. It can be predicted that even if artifact constructors managed to imitate the qualities of the mind and soul, those with strong anthropocentric beliefs will still be cautious about similar claims, upholding the superior status of human beings over other beings. Thus, by examining the psychological determinants of the assignment of the MS to artificial systems, we are expanding our knowledge of ourselves, in line with Susan Schneider's prediction that "the age of AI will be a time of soul-searching—both of ours, and for theirs."<sup>69</sup> We are aware that the empirically confirmed findings presented in this paper will not extinguish the debate on the MS of AI-driven systems. However, we hope that they will facilitate the development of a position in a discussion that will become increasingly topical and heated with the subsequent innovations.

#### BIBLIOGRAPHY / BIBLIOGRAFIA

- Allen, Colin, Iva Smit, and Wendell Wallach. "Artificial Morality: Top-down, Bottom-up, and Hybrid Approaches." *Ethics and Information Technology* 7 (2005): 149–55.
- Andreotta, Adam J. "The hard problem of AI rights." *AI and Society* (2020): 1–14.
- Aristotle. *Nicomachean Ethics*. Translated by William. D. Ross. New York: World Library Classics, 2009.
- Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. Oxford: Oxford University Press, 1998.
- Bostrom, Nick, and Eliezer Yudkowsky. "The Ethics of Artificial Intelligence." In *The Cambridge Handbook of Artificial Intelligence*. Edited by Keith Frankish and William Ramsey. Cambridge: Cambridge University Press, 2014.
- Braidotti, Rosi. *The Posthuman*. Cambridge: Polity Press Ltd, 2013.
- Bringsjord, Selmer, Paul Bello, and Naveen Sundar Govindarajulu. "Toward Axiomatizing Consciousness." In *The Bloomsbury Companion to the Philosophy of Consciousness*. Edited by Dale Jacquette. London: Bloomsbury Academic, 2018.

---

<sup>69</sup> See Susan Schneider, *Artificial You: AI and the Future of Your Mind* (Princeton: Princeton University Press, 2019), 84.

- Callicott, John B. *In Defense of the Land Ethic: Essays in Environmental Philosophy*. New York: State University of New York Press, 1989.
- Cave, Stephen, Kanta Dihal, and Sarah Dillon. *AI Narratives: A History of Imaginative Thinking about Intelligent Machines*. Oxford: Oxford University Press, 2020.
- Clark, Andy, and David Chalmers. "The Extended Mind." *Analysis* 58, no. 1(1998): 7–19.
- Coeckelbergh, Mark. "Virtual Moral Agency, Virtual Moral Responsibility: On the Moral Significance of the Appearance, Perception, and Performance of Artificial Agents." *AI & Society* 24, no. 2 (2009): 181–89.
- Danaher, John. *Automation and Utopia: Human Flourishing in a World without Work*. Harvard University Press: Cambridge, MA, 2019.
- Davies, Jem. "AI Today, AI Tomorrow: The Arm 2020 Global AI Survey". armBlueprint, February 3, 2020. <https://www.arm.com/resources/report/ai-today-ai-tomorrow-ty>.
- Deleuze, Gilles. "Postscript on the Societies of Control." *October* 59 (1990): 3–7.
- Duffy, Brian. "Anthropomorphism and the Social Robot." *Robotic and Autonomous Systems* 42, nos. 3–4 (2003): 177–90.
- Epley, Nicholas, Adam Waytz, and John T. Cacioppo, "On Seeing a Human: A Three-factor Theory of Anthropomorphism," *Psychological Review* 114, no. 4 (2007): 864–86.
- Fortuna, Paweł. *Optimum: Idea cyberpsychologii pozytywnej*. Warszawa: PWN, 2021.
- Fortuna, Paweł, Arkadiusz Gut, and Zbigniew Wróblewski, "Hey Robot, the Mind Is Not Enough to Join the Moral Community! The Effect of Assigning a Mind and a Soul to Humanoid Robot on Its Moral Status," *Annals of Psychology / Roczniki Psychologiczne* (2023). Online first. <https://doi.org/10.18290/rpsych2023.0008>.
- Fortuna, Paweł, and Oleg Gorbaniuk. "What is Behind the Buzzword for Experts and Laymen: Representation of 'Artificial Intelligence' in the IT-professionals' and Non-professionals' Minds," *Europe's Journal of Psychology* 8, no. 2 (2022): 207–18.
- Gellers, Joshua C. *Rights for Robots: Artificial Intelligence, Animal and Environmental Law*. New York: Routledge, 2021.
- Gladden, Matthew E. *Posthuman Management*. Indianapolis: Synthypnion Press, 2016.
- Gray, Heather M., Kurt Gray, and Daniel M. Wegner. "Dimensions of Mind Perception." *Science* 315, no. 5812 (2007): 619.
- Gray, Kurt, Liane Young, and Adam Waytz, "Mind Perception is the Essence of Morality." *Psychological Inquiry* 23, no. 2 (2012):101–24.
- Griffin, Andrew. "Saudi Arabia Grants Citizenship to a Robot for the First Time Ever." Independent UK, October 26, 2017. <https://www.independent.co.uk/tech/saudi-arabia-robot-sophia-citizenship-android-riyadh-citizen-passport-future-a8021601.html>.



- Gunkel, David J. "The Other Question: Can and Should Robots Have Rights?" *Ethics and Information Technology* 20 (2018): 87–99.
- Himma, Kenneth E. "Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to Be a Moral Agent?" *Ethics and Information Technology* 11 (2009): 19–29.
- Inbar, Yoel, Jeremy Cone, and Thomas Gilovich. "People's Intuitions about Intuitive Insight and Intuitive Choice." *Journal of Personality and Social Psychology* 99, no. 2 (2010): 232–47.
- Jupiter, Alex. "The Human-Cyborg Continuum: Why AI Is Pointless and Why We Should All Become Cyborgs Instead." Medium, June 4, 2016. <https://medium.com/@AlexJupiter/the-human-cyborg-continuum-why-ai-is-pointless-and-why-we-should-all-become-cyborgs-instead-4de0c4bb476f>.
- Kamm, Frances M. *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. New York: Oxford University Press, 2007.
- Kant, Immanuel. *The Metaphysics of Morals*. Cambridge: Cambridge University Press, 2017.
- Kislev, Elyakim. *Relationships 5.0: How AI, VR, and Robots Will Reshape Our Emotional Lives*. Oxford: Oxford University Press, 2022.
- Laukyte, Migle. "Artificial Agents among Us. Should We Recognize Them as Agents Proper?" *Ethics and Information Technology* 19, no. 1 (2017): 1–17.
- Leopold, Aldo. *A Sand County Almanac*. Oxford: Oxford University Press, 1987.
- Lindes, Peter. "Intelligence and Agency." *Journal of Artificial General Intelligence* 11, no. 2 (2020): 47–49.
- Loh, Wulf, and Janina Loh. "Autonomy and Responsibility in Hybrid Systems: The Example of Autonomous Cars." In *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. Edited by Patrick Lin, Keith Abney, and Ryan Jenkins. New York: Oxford University Press, 2017.
- Lukaszewicz, Aleksandra, and Paweł Fortuna. "Towards Turing Test 2.0—Attribution of Moral Status and Personhood to Human and Non-Human Agents." *Postdigital Science and Education* 4 (2022): 860–76.
- Malle, Bertram F. "How Many Dimensions of Mind Perception Really Are There?" In *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*. Edited by Ashok Goel, Colleen Seifert, and Christian Freksa. Montreal, QB: Cognitive Science Society, 2019.
- Marchetti, Antonella, et al. "Theory of Mind and Humanoid Robots from a Lifespan Perspective." *Zeitschrift für Psychologie* 226, no. 2 (2018): 98–109.
- McCall, Rosie. "Japan Has Just Granted Residency to an AI Bot in a World First." IFLScience, November 7, 2017. <http://www.iflscience.com/technology/japan-has-just-granted-residency-to-an-ai-bot-in-a-world-first>.
- Monett, Dagmar, and Colin W.P. Lewis. "Getting Clarity by Defining Artificial Intelligence: A Survey." In *Philosophy and Theory of Artificial Intelligence 2017*. Edited by Vincent C. Müller. Berlin: Springer, 2018.

- Moor, James H. "The Nature, Importance, and Difficulty of Machine Ethics." *IEEE Intelligent Systems* 21, no. 4 (2006): 18–21.
- Müller, Vincent C. "Is it Time for Robot Rights? Moral Status in Artificial Entities." *Ethics and Information Technology* 23 (2021): 579–87.
- Muzyka, Kamil. "The Basic Rules for Coexistence: The Possible Applicability of Metalaw for Human-AGI Relations." *Paladyn, Journal of Behavioral Robotics* 11, no. 1 (2020): 104–17.
- Nilsson, Nils. *The Quest for Artificial Intelligence*. Cambridge: Cambridge University Press, 2009.
- Noddings, Nel. *Caring: A Feminine Approach to Ethics and Moral Education*. Berkeley, Los Angeles: University of California Press, 2013.
- Nowak, Mateusz. "Elon Musk i założyciel Apple apelują o wstrzymanie prac nad AI: 'Utrata kontroli nad cywilizacją.'" March 30, 2023. <https://android.com.pl/tech/581815-apel-o-wstrzymanie-prac-nad-ai/>.
- Pennachin, Cassio, and Ben Goertzel. "Contemporary Approaches to Artificial General Intelligence." In *Artificial General Intelligence: Cognitive Technologies*. Edited by Cassio Pennachin, Ben Goertzel. Berlin, Heidelberg: Springer, 2007.
- Puzakova, Marina, Hyokjin Kwak, and Joseph F. Rocereto. "Pushing the Envelope of Brand and Personality: Antecedents and Moderators of Anthropomorphized Brands." *Advances in Consumer Research* 36 (2009): 413–20.
- Regan, Tom. *The Case for Animals Rights*. Berkeley, Los Angeles: University of California Press, 1983.
- Russell, Stuart J., and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Boston: Pearson, 2020.
- Saltik, Imge, Deniz Erdil, and Burcu A. Urgan. "Mind Perception and Social Robots: The Role of Agent Appearance and Action Types." In *HRI'21: Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction: March 8-11, 2021, Boulder, CO, USA*. New York, NY: Association for Computing Machinery, [2021].
- Schneider, Susan. *Artificial You: AI and the Future of Your Mind*. Princeton: Princeton University Press, 2019.
- Schweitzer, Albert. *Civilization and Ethics*. London: Adam & Charles Black, 1955.
- Searle, John R. "Minds, Brains and Programs." *Behavioral and Brain Science* 3, no. 3 (1980): 417–24.
- Singer, Peter. *Animal Liberation: A New Ethics for Our Treatment of Animals*. New York: Harper Collins, 1975.
- . *The Expanding Circle: Ethics and Sociobiology*. Oxford: Clarendon Press, 1981.
- Taddeo, Mariarosaria, and Luciano Floridi. "How AI Can Be a Force for Good." *Science* 361, no. 6404 (2018): 751–52.
- Tegmark, Max. *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York: Alfred A. Knopf, 2017.



- Tiku, Nitasha. "The Google Engineer Who Thinks the Company's AI Has Come to Life." *The Washington Post*, June 11, 2022. <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lamda-blake-lemoine>.
- Torrance, Steve, and Denis Roche. "Does an Artificial Agent Need to Be Conscious to Have Ethical Status?" In *Technologies on the Stand: Legal and Ethical Questions in Neuroscience and Robotics*. Edited by Bibi van den Berg and Laura Klaming. Nijmegen: Wolf Legal Publishers, 2011.
- Torrance, Steve. "Artificial Agents and the Expanding Ethical Circle." *AI and Society* 28, no. 4 (2013): 399–414.
- Turner, Jacob. "Rights for AI." In Turner, *Robot Rules*. Cham: Palgrave Macmillan, 2019).
- De Waal, Frans. *Primates and Philosophers: How Morality Evolved*. Princeton: Princeton University Press, 2006.
- Wang, Pei. "On Defining Artificial Intelligence." *Journal of Artificial General Intelligence* 10, no. 2 (2019): 1–37.
- Wang, Pei, Kai Liu, and Quinn Dougherty, "Conceptions of Artificial Intelligence and Singularity," *Information* 9, no. 4 (2018): 79.
- Warren, Mary A. *Moral Status: Obligations to Persons and Other Living Things*. Oxford: Clarendon Press, 1997.
- Weisman, Kara, Carol S. Dweck, and Ellen M. Markman, "Rethinking People's Conceptions of Mental Life." *Proceedings of the National Academy of Sciences of the United States of America* 114, no. 43 (2017): 11374–79.
- Wiener, Norbert. *The Human Use of Human Beings: Cybernetics and Society*. Boston: Houghton Mifflin, 1950.
- Wynsberghe, Aimee van, and Scott Robbins. "Critiquing the Reasons for Making Artificial Moral Agents." *Science and Engineering Ethics* 25, no. 3 (2019): 719–35.
- Zalta, Edward N., ed. *The Stanford Encyclopedia of Philosophy*, s.v. "Artificial Intelligence and Robotics" (by Vincent C. Müller). <https://plato.stanford.edu/entries/ethics-ai/#AutoEmpl>.

#### ABSTRACT / ABSTRAKT

Zbigniew Wróblewski and Paweł Fortuna, Moral Subjectivity and the Moral Status of Artificial Intelligence: A Philosophical and Psychological Perspective

DOI 10.12887/36-2023-4-144-05

The paper is a voice in the discussion concerning the possibility of assigning moral status to technological artifacts driven by artificial intelligence. In the dimension of philosophical reflection, it refers to the debate on the understanding of the concept of „moral status” and the related possibility of extending the moral community to include artificial agents. In the psychological perspective, on the other hand, it presents the results of a study that examined the importance

of factors relevant to the assignment of moral status, such as the features of artifacts, dimensions of mind perception, soul assignment and anthropocentric beliefs. The considerations are set in the context of currently implemented innovations, pop culture narratives shaping the image of artificial systems and discussions on the possibility of the emergence of the so-called superhuman artificial intelligence.

Keywords: artificial intelligence, moral subject, moral status, mind perception, soul assignment

Contact: (Zbigniew Wróblewski) Katedra Filozofii Przyrody i Nauk Przyrodniczych, Instytut Filozofii, Wydział Filozofii, Katolicki Uniwersytet Lubelski Jana Pawła II, Al. Raławickie 14, 20-950 Lublin, Poland; (Paweł Fortuna) Katedra Psychologii Eksperymentalnej, Instytut Psychologii, Wydział Nauk Społecznych, Katolicki Uniwersytet Lubelski Jana Pawła II, Al. Raławickie 14, 20-950 Lublin, Poland  
E-mail: (Zbigniew Wróblewski) [zbigniew.wroblewski@kul.pl](mailto:zbigniew.wroblewski@kul.pl); (Paweł Fortuna) [pawel.fortuna@kul.pl](mailto:pawel.fortuna@kul.pl)

Zbigniew Wróblewski, Paweł Fortuna – Podmiotowość moralna i status moralny sztucznej inteligencji. Perspektywa filozoficzno-psychologiczna

DOI 10.12887/36-2023-4-144-05

Artykuł jest głosem w dyskusji dotyczącej możliwości nadawania statusu moralnego artefaktom technologicznym sterowanym sztuczną inteligencją. W wymiarze refleksji filozoficznej odnosi się do debaty nad rozumieniem pojęcia „status moralny” oraz powiązanej z nim możliwości rozszerzenia wspólnoty moralnej o sztucznych agentów. W perspektywie psychologicznej prezentuje z kolei wyniki badań, w których testowano znaczenie takich czynników istotnych dla nadawania statusu moralnego, jak cechy artefaktów, wymiary percepcji umysłu, przypisywanie duszy oraz przekonania antropocentryczne. Rozważania są osadzone w kontekście aktualnie wdrażanych innowacji, popkulturowych narracji kształtujących obraz sztucznych systemów oraz dyskusji nad możliwością pojawienia się tzw. nadludzkiej sztucznej inteligencji.

Słowa kluczowe: sztuczna inteligencja, podmiot moralny, percepcja umysłu, przypisywanie duszy

Kontakt: (Zbigniew Wróblewski) Katedra Filozofii Przyrody i Nauk Przyrodniczych, Instytut Filozofii, Wydział Filozofii, Katolicki Uniwersytet Lubelski Jana Pawła II, Al. Raławickie 14, 20-950 Lublin; (Paweł Fortuna) Katedra Psychologii Eksperymentalnej, Instytut Psychologii, Wydział Nauk Społecznych, Katolicki Uniwersytet Lubelski Jana Pawła II, Al. Raławickie 14, 20-950 Lublin  
E-mail: (Zbigniew Wróblewski) [zbigniew.wroblewski@kul.pl](mailto:zbigniew.wroblewski@kul.pl); (Paweł Fortuna) [pawel.fortuna@kul.pl](mailto:pawel.fortuna@kul.pl)

Małgorzata ABASSY

## SNUCIE OPOWIEŚCI JAKO NARZĘDZIE OSWAJANIA SZTUCZNEJ INTELIGENCJI Między modelującą funkcją języka a podmiotowością maszyn

*Człowiek rozpoznaje podobieństwo procesów zachodzących w maszynach do tych, które odbywają się w obrębie jego sieci neuronowych: myślenia, wnioskowania, podejmowania decyzji, a nawet odwzorowywania reakcji w sferze zarezerwowanej wyłącznie dla istot ludzkich. Sztuczna inteligencja jest pojęciem słabo rozpoznawym w kulturze masowej. Przesłania je postać cyborga oraz generowane przez nią znaczenia. Te ostatnie stanowią wypadkową oglądanych filmów, zasłyszanych opowieści oraz zamurzenia w media społecznościowe, gdzie zacierają się granice między tym, co ludzkie, a tym, co nie-ludzkie; unieważnione zostają czas i przestrzeń*

Celem artykułu jest udzielenie odpowiedzi na pytanie o rolę narracji – opowieści o sztucznej inteligencji, jakich słuchamy i jakie opowiadamy, dając w ten sposób wyraz naszym nadziejom i obawom – w oswojaniu sztucznej inteligencji. Zakładamy, że kreujemy w ten sposób rzeczywistość, najpierw za pomocą słów, a następnie – poprzez wybrane działania. Performatywna funkcja języka<sup>1</sup>, chociaż marginalizowana w badaniach nad sztuczną inteligencją<sup>2</sup>, pojawia się w powieści Neala Stephensona *Zamieć*<sup>3</sup>. Realizowana jest również, najpierw jako dialog wewnętrzny bohatera, a później w formie działania w tekście Iana McEwana *Maszyny takie jak ja*<sup>4</sup>. Rolę literatury jako jednego z wtórnych systemów modelujących kultury podkreślali semiotycy szkoły tatarsko-moskiewskiej<sup>5</sup>. Teoria semiotyczna Jurija Łotmana stanowi rozszerzenie teorii o fatycznej funkcji języka, dostarczając materiału do refleksji nad funk-

<sup>1</sup> Por. J.L. Austin, *How to Do Things with Words*, Oxford University Press, London 1962, s. 6n.

<sup>2</sup> Por. M. Kevin, *Phatics and the Design of Community*, ACM, Boston 2009, ss. 3133-3136; zob. B. Malinowski, *The Problem of Meaning in Primitive Languages*, w: *The Meaning of Meaning*, red. Ch.K. Ogden, I.A. Richards, Paul, London 1923, ss. 296-336; D. Radovanovic, M. Ragnedda, *Small Talk in the Digital Age: Making Sense of Phatic Posts*, [https://www.academia.edu/1506309/Small\\_talk\\_in\\_the\\_Digital\\_Age\\_Making\\_Sense\\_of\\_Phatic\\_Posts](https://www.academia.edu/1506309/Small_talk_in_the_Digital_Age_Making_Sense_of_Phatic_Posts); B. Slotnick, *Understanding Phatic Aspects of Narrative when Designing Assistive and Augmentative Communication Interfaces*, “International Journal of Ambient Computing and Intelligence” 6 (2014), nr (2), s. 75-94.

<sup>3</sup> Zob. N. Stephenson, *Zamieć*, tłum. J. Polak, Warszawa 2009.

<sup>4</sup> Zob. I. McEwan, *Maszyny takie jak ja*, tłum. A. Szulc, Albatros, Warszawa 2019.

<sup>5</sup> Zob. J.M. Lotman, *O znaczeniach we wtórnych systemach modelujących*, tłum. J. Faryno, „Pamiętnik Literacki. Czasopismo kwartalne poświęcone historii i krytyce literatury polskiej” 60(1969) nr 1, s. 279-294.

cją tekstu literackiego jako materiału odzwierciedlającego i porządkującego ludzką percepcję rzeczywistości oraz narzędzia kreowania jej.

Wybór materiału źródłowego został podyktowany czynnikiem czasowym – powstanie tekstów dzieli ponad dwadzieścia lat – oraz problemami, które, sygnalizowane w pierwszym z analizowanych tekstów, zyskały rozwinięcie w drugim. W *Zamieci* główny bohater Hiro otrzymuje ostrzeżenie: „– Jesteś hakerem. Musisz troszczyć się o swoje głębokie struktury. – Jakie struktury? – Neurolingwistyczne ścieżki w mózgu. Pamiętasz, jak uczyłeś się kodu dwójkowego? – Owszem. – Przecierałeś wówczas nową ścieżkę w mózgu. Głęboką strukturę. Gdy korzystasz z niej, twoje nerwy tworzą nowe wypustki [...] aksony dzielą się i wpychają między gangliony i neurofibryle [...] twój bioware ulega samodoskonaleniu [...] software staje się częścią hardware’u”<sup>6</sup>. Mowa tutaj o manipulacji zewnętrznej, jednak zniekształcanie wyobrażeń o rzeczywistości poprzez kwestionowanie jej podstawowych znaczeń może być również wynikiem dialogu wewnętrznego. Ostrzeżenie przed przekształcaniem ścieżek neuronalnych w mózgu, odpowiedzialnych za istnienie słów, którymi określamy rzeczywistość, staje się faktem, gdy główny bohater powieści *Maszyny takie jak ja* uznaje robota Adama za „przedstawiciela tego samego co ja gatunku”<sup>7</sup>. W powieści *Zamieć* problem języka można traktować jako wstęp do refleksji nad pierwotnym systemem modelującym i warunkami jego przekształcania bądź zniekształcania. Z kolei *Maszyny takie jak ja* to przykład elementu wtórnego systemu modelującego: zawiera on obraz humanoidalnej maszyny funkcjonującej w relacjach z ludźmi. Nie będąc rozprawą filozoficzną, proponuje odpowiedzi na istotne pytania o podmiotowość. Pojawiająca się w powieści kreacja postaci Alana Turinga – matematyka, kryptografa, pomysłodawcy eksperymentu mającego na celu zdefiniowanie sztucznej inteligencji<sup>8</sup>, stanowi dodatkowy czynnik wskazujący na płynność granic między fikcją literacką a rzeczywistością. Ostatecznie to postać wynalazcy udziela w powieści odpowiedzi na pytanie, czy artefakt sztucznej inteligencji – Adam – jest przedmiotem czy podmiotem.

Podstawowe pytania badawcze sygnalizowane w niniejszym artykule dotyczą tych obszarów postrzegania rzeczywistości, które są zagrożone zniekształceniem fundamentu aksjologicznego, i w konsekwencji – ontologicznego. Które z elementów snutej opowieści powinny budzić szczególną czujność słuchającego i opowiadającego? Które wątki fantastycznej w 1992 roku powieści Stephensona nie tylko stały się faktem, ale też wykroczyły poza ramy

<sup>6</sup> Stephenson, dz. cyt., s. 121.

<sup>7</sup> McEwan, dz. cyt., s. 104.

<sup>8</sup> Zob. A. Hodges, hasło „Alan Turing”, *Stanford Encyclopedia of Philosophy*, red. E.N. Zalta, <https://plato.stanford.edu/entries/turing/>.

nakreślone przez wyobraźnię autora? Czy Adam wykreowany przez McEwana staje się postacią zapraszaną do intymnego świata istoty ludzkiej jako jej równoprawny członek? Jak daleko przesuwamy granice osławiania sztucznej inteligencji poza niwelowanie lęku przed nieprzewidywalnością maszyn? Czy są to już granice naszej ontologii? Nie zakładamy udzielenia odpowiedzi na postawione pytania w ramach niniejszego artykułu. Problem wymaga interdyscyplinarnych, zakrojonych na szeroką skalę badań. Zakładamy jednak uwrażliwienie twórców i odbiorców narracji na opowieści o sztucznej inteligencji, które – zgodnie z fatyczną funkcją języka – dają początek realnym wydarzeniom.

Tytułowe powieści dzieli ponad dwadzieścia lat: *Zamieć* (*Snow Crash*<sup>9</sup>) została napisana w roku 1992, podczas gdy druga ze wspomnianych pozycji, *Maszyny takie jak ja* (*Machines Like Me*<sup>10</sup>), ukazała się w 2019. Łączy je fakt, że stanowią namysł nad przyszłością, którą ludzkość tworzy, wykorzystując język jako narzędzie kreacji. Są projektowaniem nadziei i obaw. Ich autorzy, wychwytyjąc wątki codziennych narracji, stawiają pytania o granice światów i gatunków. W tle brzmi również niepostawione wprost pytanie o ludzką sprawczość, a ostatecznie o odpowiedzialność za świat, w którym przychodzi nam żyć. Przyspieszenie informacyjne sprawiło, że obecne pokolenia nie tworzą rzeczywistości dla swoich potomków, lecz dla siebie samych: nasze „dzisiaj” błyskawicznie staje się naszym „jutro”. Prognozowane przez Stephensona narodziłoby się Metawersum – równoległej rzeczywistości wirtualnej współistniejącej z realną, stają się faktem wskutek zabiegów Maxa Zuckenberg, właściciela platformy społecznościowej Facebook<sup>11</sup>. Czy literacka postać humanoidalnej maszyny – „Adama”, pojawiająca się na kartach powieści McEwana również zaistnieje w naszym życiu jako partner, przyjaciel i kochanek? Narracja, która rozpoczyna powieść, nie pozostawia co do tego wątpliwości: „Była to nadzieja zrodzona z religijnej tęsknoty, prawdziwy Święty Graal nauki. Nasze ambicje były niebosiężne i przyziemne – chodziło o potwierdzenie mitu stworzenia i o monstualny akt miłości własnej. Kiedy tylko stało się to możliwe, nie mieliśmy innego wyjścia: musieliśmy zrealizować nasze marzenia i ponieść wszystkie tego konsekwencje. Mówiąc górnolotnie, chcieliśmy uciec przed

<sup>9</sup> Zob. N. S t e p h e n s o n, *Snow Crash*, Bantam Books, New York 1992.

<sup>10</sup> Zob. I. M c E w a n, *Machines Like Me*, Jonathan Cape, London.

<sup>11</sup> Zob. M. D a m a r, *Metaverse Shape of Your Life for Future: A Bibliometric Snapshot*, „Journal of Metaverse” 1(2021) nr 1, s. 1-8; M. S p a r k e s, *What is a Metaverse*, „New Scientist” 251(2021) nr 3348, s. 1-18; S. K r a u s, i in., *Facebook and the Creation of the Metaverse: Radical Business Model Innovation or Incremental transformation?*, „International Journal of Entrepreneurial Behavior & Research” 28(2022) nr 9, s. 52-77.

własną śmiertelnością, stanowić czoło boskości albo nawet zastąpić ją idealną jaźnią<sup>12</sup>.

### „OSWAJANIE” I „SZTUCZNA INTELIGENCJA” SEMANTYKA POJEĆ

W powieści *Mały książę* Antoine’a de Saint Exupery’ego między Księciem a lisem toczy się znamieny dialog o znaczeniu słowa „oswajając”: „Szukam przyjaciół. Co to znaczy «oswoić»? – To rzecz, o której niestety zapomniano – powiedział lis. – To znaczy «stworzyć więzi»<sup>13</sup>. „Oswajanie” to proces o znaczeniu egzystencjalnym. Nie tylko pozwala dostrzec wyjątkowość i podmiotowość: „Jeżeli jednak mnie oswoisz, będziemy sobie wzajemnie potrzebni. Będziesz dla mnie jedyny na świecie. I ja będę dla ciebie jedyny na świecie”<sup>14</sup>, ale także otwiera na nowe postrzeganie rzeczywistości: „Jeżeli jednak mnie oswoisz, moje życie będzie jakby opromienione słońcem”<sup>15</sup>. Czy taką definicję oswajania, jaka wyłania się z rozmowy księcia i lisa, możemy stosować do naszych relacji ze sztuczną inteligencją? Zakładałaby ona tworzenie więzi, a więc podmiotowość stron wchodzących w interakcję. Odruchową reakcją na wysuniętą hipotezę jest odrzucenie podmiotowości sztucznej inteligencji, nawet w jej humanoidalnej postaci<sup>16</sup>. Warto jednak zadać sobie pytanie, czy w głębi naszej świadomości nie odbywa się niezauważalne zacieranie granic i mieszanie pojęć<sup>17</sup>.

Terminu „sztuczna inteligencja” po raz pierwszy użył amerykański informatyk John McCarthy<sup>18</sup> podczas konferencji w Darmouth w roku 1955. W definicji zawartej w *Encyclopedia Britannica* czytamy: „Sztuczna inteligencja to zdolność komputera cyfrowego lub robota sterowanego komputerowo do wykonywania zadań powszechnie kojarzonych z inteligentnymi istotami. Termin ten jest często stosowany do projektu rozwoju systemów obdarzo-

<sup>12</sup> M c E w a n, *Maszyny takie jak ja*, s. 9.

<sup>13</sup> A. de S a i n t - E x u p é r y, *Mały książę*, tłum. A. Trznadel-Szczepanek, Nasza Księgarnia, Warszawa 2017, s. 32.

<sup>14</sup> Tamże.

<sup>15</sup> Tamże.

<sup>16</sup> Zob. R. W a w e r, *Sztuczna inteligencja – produkt czy podmiot?*, „Studia Teologii Dogmatycznej” 2018, nr 4, s. 115-126, [https://repozytorium.uwb.edu.pl/jspui/bitstream/11320/7817/1/STD\\_4\\_2018\\_R\\_Wawer\\_Sztuczna\\_inteligencja.pdf](https://repozytorium.uwb.edu.pl/jspui/bitstream/11320/7817/1/STD_4_2018_R_Wawer_Sztuczna_inteligencja.pdf); N. B o s t r o m, *Transhumanist Values*, w: *Ethical Issues for the 21st Century*, red. F. Adams, Philosophical Documentation Center Press, Oxford 2003, s. 3-14.

<sup>17</sup> Zob. S. L e m, *Dzienniki gwiazdowe*, Wydawnictwo Literackie, Kraków 2019.

<sup>18</sup> Zob. A. M e y e r s, *Stanford’s John McCarthy, Seminal Figure of Artificial Intelligence, Dies at 84*, „Stanford Report”, October 25, 2011, <https://news.stanford.edu/news/2011/october/john-mccarthy-obit-102511.html>.



nych procesami intelektualnymi charakterystycznymi dla człowieka, takimi jak zdolność rozumowania, odkrywania znaczenia, uogólniania lub uczenia się z przeszłych doświadczeń<sup>19</sup>. Na oficjalnej stronie rządu Rzeczypospolitej Polskiej widnieje następująca definicja: „Sztuczną inteligencję (ang. Artificial Intelligence – AI) próbuje się definiować jako dziedzinę wiedzy obejmującą m.in. sieci neuronowe, robotykę i tworzenie modeli zachowań inteligentnych oraz programów komputerowych symulujących te zachowania, włączając w to również uczenie maszynowe (ang. machine learning), głębokie uczenie (ang. deep learning) oraz uczenie wzmocnione (ang. reinforcement learning)”<sup>20</sup>. Wskazując na zakres zjawisk, definicje nie precyzują ich istoty. Człowiek jednak rozpoznaje pewne podobieństwo procesów zachodzących w maszynach do tych, które odbywają się w obrębie jego własnych sieci neuronowych: myślenia, wnioskowania, podejmowania decyzji, a nawet odwzorowywania reakcji w sferze zarezerwowanej, jak mogłoby się wydawać, wyłącznie dla istot ludzkich<sup>21</sup>. Sztuczna inteligencja jest pojęciem słabo rozpoznanym w kulturze masowej. Przesłania je postać cyborga oraz generowane przez nią znaczenia. Te ostatnie stanowią wypadkową oglądanych filmów, zasłyszanych opowieści oraz zanurzenia w media społecznościowe, gdzie zacierają się granice między tym, co fizyczne, a tym, co wirtualne, tym, co ludzkie, a tym, co nie-ludzkie; unieważnione zostają czas i przestrzeń. Potocznie sztuczna inteligencja kojarzy się z autonomicznymi maszynami, mogącymi wyrwać się spod ludzkiej kontroli. Szybki rozwój nowych technologii, wypieranie z rynku pracy ludzi na rzecz mniej zawodnych i tańszych w eksploatacji maszyn, spuścizna kultury masowej, zawierająca w sobie literackie i filmowe obrazy buntu maszyn, oraz brak zgodnej odpowiedzi ekspertów na pytanie, czy powinniśmy obawiać się rozwoju sztucznej inteligencji, czy raczej pokładać w niej nadzieję<sup>22</sup>, skłania do przyjęcia jeszcze jednej, nieobejmującej aspektu egzystencjalnego, definicji osvajania – jako „przyzwyczaić do kogoś lub do czegoś albo zapoznać z czymś; przyzwyczaić dzikie zwierzęta do przebywania wśród ludzi lub do służenia ludziom”<sup>23</sup>.

<sup>19</sup> B. J. C o p e l a n d, hasło “Artificial Intelligence (AI)”, *Encyclopedia Britannica*, <https://www.britannica.com/technology/artificial-intelligence/Methods-and-goals-in-AI>. Jeśli nie podano inaczej, tłumaczenie fragmentów prac obcojęzycznych – M.A.

<sup>20</sup> *Czym jest sztuczna inteligencja*, Portal sztucznej inteligencji, <https://www.gov.pl/web/ai/czym-jest-sztuczna-inteligencja2>.

<sup>21</sup> Zob. K. D e v l i n, *O pożądaniu, nauce i sztucznej inteligencji*, tłum. A. Gralak, Wydawnictwo Dolnośląskie, Wrocław 2020.

<sup>22</sup> Zob. N. B o s t r o m, *Superinteligencja. Scenariusze, strategie, zagrożenia*, tłum. D. Konowrocka-Sawa, Helion, Gliwice 2016.

<sup>23</sup> Hasło „Oswoić”, *Słownik języka polskiego*, pod red. W. Doroszewskiego, Wydawnictwo Naukowe PWN, Warszawa 1997, <https://doroszewski.pwn.pl/haslo/oswoic/>.

Postacią, wyłaniającą się z tekstów kultury masowej, jest cyborg – połączenie maszyny i człowieka, obiektu nauki o systemach – cybernetyki. Twórcą tego ostatniego pojęcia, wywodzącego się od greckiego *kybernetikos*, jest Norbert Wiener, który w połowie dwudziestego wieku sięgnął do wczesnych hellenistycznych teorii dotyczących zarządzania i połączył je z ideami pojawiającymi się w dziedzinie informatyki, tworząc w ten sposób nowe interdyscyplinarne podejście do systemów<sup>24</sup>. Cybernetyka koncentrowała się na przepływach informacji między systemami, które obejmowały komponenty organiczne i obliczeniowe. Znaczenie podejścia Wienera trudno przecenić. Jak zauważyła N. Katherine Hayles, twórczyni teorii posthumanizmu: „Ze wszystkich implikacji, jakie niesie ze sobą cybernetyka pierwszej fali, być może żadna nie była bardziej niepokojąca i potencjalnie rewolucyjna niż ta, że granice ludzkiego podmiotu są konstruowane, a nie dane. Postrzegając kontrolę, komunikację i informację jako zintegrowany system, cybernetyka radykalnie zmieniła sposób, w jaki pojmowano granice”<sup>25</sup>. Wiener zdawał sobie sprawę z konsekwencji owych implikacji, stwierdzając, że kolejnym problemem, z jakim ludzkość będzie się mierzyć będzie relacja między człowiekiem a maszyną<sup>26</sup>. Choć w wczesnych ujęciach systemowych nie było nacisku na ontologiczne połączenie tego, co organiczne, i tego, co sztuczne, Wiener zdawał sobie sprawę, że ludzka ciekawość i skłonność do przesuwania granic mogą doprowadzić do zacierania się granic między „organicznym” a „mechanicznym”.

### „ZAMIEĆ” JAKO ZAMĘT UMYŚŁU

Tytułowa *Zamieć* ma trzy znaczenia: zjawiska pogodowego związanego z zimą, śniegiem i wiatrem, nazwy narkotyku oraz zawieszenia się systemu komputera. W powieści te trzy znaczenia przenikają się, stanowiąc wskazówkę, że granice między Metawersum a światem materialnym zacierają się, niebezpieczeństwa zaś generowane w jednym z nich mają skutki w drugim. Życie toczy się głównie w Metawersum, za pośrednictwem awatarów oraz skomplikowanego systemu reguł, odzwierciedlającego w dużej mierze normy społeczne realnego świata. Główny bohater, Hiro, jest hakerem, co daje mu uprzywilejowaną pozycję w świecie wirtualnym ze względu na posiadane

<sup>24</sup> Zob. N. Wiener, *Cybernetics: Or Control and Communication in the Animal and the Machine*, MIT Press, Cambridge, Massachusetts, 1965.

<sup>25</sup> K. Hayles, *How We Become Posthumans: Virtual Bodies in Cybernetics, Literature and Informatics*, University of Chicago Press, Chicago, London 1999, s. 84.

<sup>26</sup> Por. N. Wiener, *God and Golem, Inc.: A Comment on Certain Points Where Cybernetics Impinges on Religion*, MIT Press, Cambridge 1964, s. 71.



umiejętności. Nie znajduje ona odzwierciedlenia w rzeczywistości materialnej: Hiro mieszka w blaszanej przyczepie na obrzeżach dużego miasta, a jego jedynym majątkiem jest laptop. Bohater jako pierwszy staje wobec zagrożenia działaniem *z a m i e c i*: „Ale «zamięć» jest słowem z języka komputerowców. Oznacza awarię systemu – wiesz – na tak fundamentalnym poziomie, że siada wszystko, nawet sterownik strumienia elektronów uderzających w ekran monitora. Elektrony robią wówczas co chcą, a na monitorze zamiast siatki pikseli pojawia się wściekła śnieżna zawieja”<sup>27</sup>. Zjawisko, skojarzone z nazwą narkotyku, implikuje uszkodzenie władz umysłowych – poprzez paralelę z awarią komputera: „Rzecz nie byłaby niczym niezwykłym przed barem w rzeczywistości. Ale w Metawersie! W Metawersie nie wolno sprzedawać dragów, bo można się zaćpać od samego patrzenia”<sup>28</sup>. Narrator zawarł sugestię, że *z a m i e c* może być nieznanym językiem, oddziałującym na użytkowników Metawersum: „Inaczej mówiąc, ktoś potraktował twój nerw wzrokowy setkami tysięcy bajtów informacji [...]. Jakiś haker chciał mi pokazać program. Wszystko było okay, dopóki Brandy nie rozwinęła zwoju. Program był do dupy i pojawiła się zamięć. Zamiast demo obejrzałem śnieg”<sup>29</sup>. Nieświadomy zagrożenia użytkownik Dawid ulega oddziaływaniu mapy bitowej i trafia do szpitala, a Hiro zastanawia się nad przyczyną choroby przyjaciela i próbuje sam siebie przekonać: „Nie można dostać zapaści od patrzenia na mapę bitową. Zresztą, kto wie?”<sup>30</sup>. Chory Dawid recytuje ciąg głosek, pozornie nic nieznaczących. „– Badają go, szukając tego co zwykle. Robią mu EEG, ultrasonografię, tomografię komputerową i co tylko. Wszystko jest w porządku. Jego mózg działa bez zarzutu. Hardware jest okay. – Tylko coś nie tak z programem? – Program jest zawirusowany. Dawid przeżył wczoraj zamięć [...] w głowie”<sup>31</sup>. Zwój, obejrzany przez użytkownika okazuje się kodem dwójkowym, przekonanie zaś, że zwykle korzystanie z komputera nie niesie ze sobą konsekwencji natury psychologicznej – błędne. Wirus komputerowy jest dla ludzkiego mózgu tak samo niebezpieczny jak zarazki biologiczne: „Na zwoju, który pokazała mu Brandy, nie było zwykłego śniegu. Było tam natomiast mnóstwo przetworzonych cyfrowo informacji. [...] – O jakiej informacji mówimy? – O niedobrej. O metawirusie. O bombie atomowej wojny informacyjnej. O wirusie, który każe systemowi zarażać się nowymi wirusami”<sup>32</sup>. Oddziaływanie na język pozwala twórcom Metawersum czy też tym, którzy pragną kontrolować ludzkość, przekształcać jego kody i wynikające z nich znaczenia

<sup>27</sup> S t e p h e n s o n, *Zamięć*, s. 44.

<sup>28</sup> Tamże, s. 43.

<sup>29</sup> Tamże, s. 73.

<sup>30</sup> Tamże, s. 179.

<sup>31</sup> Tamże, s. 188.

<sup>32</sup> Tamże.

w taki sposób, by w ostatecznym rozrachunku kontrolować umysł: „Język nie jest wehikułem myśli, ale determinującym je medium. Wytycza granice poznania. Nasze pojmowanie rzeczy zorganizowane jest wokół przepływu wrażeń przez te granice. Dlatego analiza ewolucji języka jest analizą ewolucji ludzkiego umysłu”<sup>33</sup>. Ponowne podkreślenie performatywnej funkcja języka rodzi refleksje na temat wzajemnego warunkowania się ludzkiego umysłu i sztucznej inteligencji. Ewolucja jednego nierozzerwalnie wiąże się z przekształceniem drugiego. Narrator *Zamieci* poprzez uwagę, że mózg podlega programowaniu podkreśla możliwość intencjonalnego wykorzystania tej właściwości ludzkiego umysłu do celów destrukcji ludzkości jako samodzielnie myślących podmiotów: „Czy to znaczy, że metawirus miał zawsze dwie postaci: lingwistyczną i biologiczną? – Dokładnie tak. Na tym polega cały jego sens. Atakuje i tu, i tu: umysł i ciało. Jest podobny do wirusa opryszczki. Wirus ten, wnikając do ciała, zmierza wprost do systemu nerwowego”<sup>34</sup>. Zawirusowanie systemu komputerowego, połączone z oddziaływaniem na mózg użytkownika Internetu, wskazuje na zagrożenie wykraczające poza przestrzeń wirtualną. Oznacza również, że rozwiązania i środki zapobiegawcze powinny obejmować też świat realny: „Wydawało im się, że najgorszą rzeczą, jaka może ich spotkać po zawirusowaniu systemu będzie zdjęcie gogli i zresetowanie komputera. [...] Z tego powodu Metawers był otwarty dla wszystkich i całkowicie bezbronny, jak lotniska w czasach przed wybuchem terroryzmu. [...] Wolny haker może ocalić przyszłość Metawersu, zanim korporacyjni giganci zbudzą się ze snu”<sup>35</sup>. Świadomość znaczenia języka oraz przenikania się rzeczywistości materialnej i wirtualnej jest pierwszym krokiem na drodze do postawienia granicy treściom przyjmowanym z ekranów komputerów. Haker jest najbardziej narażony na oddziaływanie kodów dwójkowych, ponieważ zna ten język. Niemniej jednak, niebezpieczeństwo obejmuje również innych, biernych użytkowników nowych technologii: są oni narażeni na atak na umysł – centrum świadomości. Przecieranie nowych ścieżek w mózgu i zmiana sposobu myślenia o sztucznej inteligencji wyłania się jako problem etyczny w kolejnej powieści.

#### GRANICE CZŁOWIECZEŃSTWA „MASZYNY TAKIE JAK JA”

Historia Adama, Mirandy i Frienda – głównego bohatera, opowiedziana jest z perspektywy tego ostatniego. Nabywszy najnowszy model humanoidalnego

<sup>33</sup> Tamże, s. 261.

<sup>34</sup> Tamże, s. 379.

<sup>35</sup> Tamże, s. 333.

robotą, postanowił on zaprogramować go wspólnie z sąsiadką mieszkającą piętro wyżej: „Tworzyliśmy dziecko!”<sup>36</sup>. Początkowa relacja kobiety i mężczyzny ewoluuje pod wpływem obecności trzeciego elementu o niedookreślonym statusie ontologicznym. Dialog wewnętrzny, jaki toczy się w umyśle narratora, wskazuje jednak na zmiany zachodzące w postrzeganiu relacji między mężczyzną a kobietą, przynależącej do świata materialnego, biologicznego, a jednocześnie wykraczającej poza ten świat. Wraz ze zmianą postrzegania maszyny, wzrasta jej samodzielność i zdolność do podejmowania decyzji niezależnie od intencji programistów. Narrator przyznaje, że zanim powstał model Adama, „wyobraźnia [...] przerobiła już tę przyszłość w książkach”<sup>37</sup>.

Główne wątki powieści, analizowane w niniejszym artykule, to wzajemne relacje zaistniałe i ewoluujące w trójkącie: Friend–Miranda, Miranda–Adam, Adam–Friend. Przedmiotem analizy będzie narracja dotycząca zmiany statusu ontologicznego humanoidalnej maszyny oraz postrzegania więzi intymnej. Procesowi osławiania nowego nabytku towarzyszy deklaracja: „Spodziewałem się przyjaciela. Byłem gotów potraktować Adama jako gościa w moim domu, nieznanego, którego będę poznawał”<sup>38</sup>. Doskonały technologicznie projekt budzi jednak emocje, wytrącające narratora ze stanu emocjonalnej neutralności i powodujące rozdźwięk między wiedzą a doświadczeniem: „Rozum mówił mi, że to plastik lub coś podobnego, lecz dotyk stwierdzał, że to ludzkie ciało”<sup>39</sup>. Niespodziewanie pojawia się, jeszcze nierozpoznane poczucie, że ten doskonały model męskiej postaci może być rywalem. Szybko zostaje ono zniwelowane: „Mężczyzna, który właśnie się zakochał, wie czym jest życie”<sup>40</sup> – myśli o sobie narrator, rozgraniczając w ten sposób własne „ja” i maszynę, niezdolną – jego zdaniem – do wyższych uczuć. W miarę upływu czasu narrator zaczyna dostrzegać Adama przez pryzmat własnych emocji: „Miałem w tym momencie wrażenie, że jego muskularne barki przytłacza brzemień samotności”<sup>41</sup>. Nie można jednak stwierdzić, że emocje te są nieumotywowane. Mechaniczny gość werbalizuje obawy, które stanowią cechę ludzi, a nie maszyn: „Ten przewód. Jeśli go wyciągnę, będzie bolało”<sup>42</sup>. Zakładając, że nie może odczuwać fizycznego bólu, narrator przypisuje odpowiedź algorytmowi, który tak właśnie zwerbalizował reakcję na pytanie o samopoczucie. Walka między wiedzą, że Adam jest maszyną, a doświadczeniem go jako podmiotu znajduje kulminację w momencie, gdy narrator podsłuchuje akt miłosny: „Niedobrze, mój umysł

<sup>36</sup> M c E w a n, *Maszyny takie jak ja*, s. 47.

<sup>37</sup> Tamże, s. 10.

<sup>38</sup> Tamże, s. 11.

<sup>39</sup> Tamże, s. 19.

<sup>40</sup> Tamże, s. 20.

<sup>41</sup> Tamże, s. 37.

<sup>42</sup> Tamże.

znosił to niedobrze. [...] Byłem pierwszym mężczyzną, któremu przypisał rogi artefakt. [...] Sam Alan Turing często mówił w młodości, że kiedy maszyny od człowieka nie da się odróżnić po zachowaniu, będziemy musieli przypisać jej człowieczeństwo, dlatego gdy nocne powietrze rozdarł nagle ekstatyczny krzyk Mirandy [...] potraktowałem Adama jak przedstawiciela tego samego co ja gatunku, ze wszystkimi wynikającymi stąd przywilejami i zobowiązaniami. Znienawidziłem go<sup>43</sup>. Obawie, graniczącej z pewnością, że mężczyźni staną się zbędni<sup>44</sup>, towarzyszy narastające przekonanie, że Adam jest prawdziwszą i lepszą wersją człowieka. Narracja obfituje w odwołania do Turinga, historia relacji człowieka i maszyny zostaje zaś podsumowana tyradą sławnego wynalazcy, będącą odpowiedzią na zniszczenie Adama: „Mam nadzieję – podjął chłodno – że pewnego dnia to, co pan zrobił Adamowi, zostanie uznane za poważaną zbrodnię [...]. Próbował pan zniszczyć życie. On był czującą istotą. Miał swoją jaźń. To, jak powstała, dzięki biologicznym neuronom, mikroprocesorom czy sieciom neuronowym, nie ma znaczenia<sup>45</sup>. Uznając podmiotowość maszyny, należy zakwestionować słowa narratora: „Ja go kupiłem i ja miałem prawo go zniszczyć<sup>46</sup>. Powieść, traktowana jako element wtórnego systemu modelującego, zawiera opowieść o współistnieniu ludzi i maszyn jako pełnoprawnych członków przestrzeni również aksjologicznej.

\*

Obydwie powieści wybrane jako materiał źródłowy do analizy mechanizmów osvajania sztucznej inteligencji zawierają w sobie przesłanki do odczytania zarówno obaw, jak i nadziei pokładanych w rozwoju technologicznym. Poprzez ich werbalizację na kartach książki, a następnie wizualizację w wyobraźni czytelnika rodzi się kolejna historia, dając impuls do zaistnienia nowej rzeczywistości. Wagę narracji jako narzędzia kreacji podkreśliła Hayles: „Nosimy ze sobą nasze przeszłe historie jak ślimak skorupę<sup>47</sup>. Wtórjuje jej Donna Haraway: „Mówiąc metaforycznie, jesteśmy jak krab pustelnik inkrustowany pąklami. Postrzegam siebie i wszystkich innych jako zmieniających muszle w miarę, gdy rośniemy. Ale każda muszla, którą podnosimy, ma swoją historię i z pewnością jej nie wybiera-

<sup>43</sup> Tamże, s. 101-104.

<sup>44</sup> Por. tamże, s. 103.

<sup>45</sup> Tamże, s. 346.

<sup>46</sup> Tamże, s. 318.

<sup>47</sup> N. H a y l e s, *Afterword: The Human in the Posthuman*, „Cultural Critique” 53(2003), s.

my<sup>48</sup>. Nie tylko człowiek dziedziczy historie, które wpływają na przyswajanie i interpretację przezeń narracji płynących z otaczającego go świata. Sztuczna inteligencja w swoich wielorakich manifestacjach również obciążona jest treściami wcześniejszych opowieści: „Tak jak cyborg jest dzieckiem militarystyki i «wielkiej nauki», tak ja jestem dzieckiem katolicyzmu i «zimnej wojny»<sup>49</sup> – przypomina Haraway. Snucie historii nie tylko jest aktem kreacji nowych rzeczywistości, ale także może stać się aktem autokreacji. W dążeniu do przekroczenia tego, co postrzegane jest nie jako nierozpoznany potencjał, lecz jako ograniczenie, które należy przewyciężyć, stwarzamy na nowo siebie samych: „Ziściło się to ostatecznie w jesieni dwudziestego wieku: postawiono wtedy pierwszy krok na drodze ku spełnieniu się dawnego marzenia, zaczęła się długa lekcja, w trakcie której przekonał się, że niezależnie do tego, jak bardzo jesteśmy skomplikowani, jak bardzo zawodni i trudni do opisanego nawet w najprostszych działaniach i trybach istnienia, można nas skopiować i ulepszyć<sup>50</sup>. W wybranych tekstach źródłowych funkcja języka jako narzędzia kreacji przejawia się na różne sposoby. W *Zamieci* zostajemy ostrzeżeni przed wykorzystaniem naszych głębokich, pierwotnych struktur mózgu – tych obszarów, w których rodzi się świadomość i kształtują najwcześniejsze połączenia neuronalne – przez hakerów, podmioty żądne władzy<sup>51</sup>. W *Maszyny takie jak ja* informacja o znaczeniu podana jest nie wprost, jako wewnętrzny monolog głównego bohatera, nabywcy „Adama”, próbującego rozpoznać sztuczny byt poprzez szukanie odniesień do siebie samego. Granice zacierają się: świadomość zostaje pochłonięta przez „zamięć”, „Adam” zaś zostaje potraktowany jak człowiek w kontekście aktu seksualnego – jednej z najbardziej intymnych form oddania się drugiemu człowiekowi.

Wysnute wnioski są zarazem podstawą sformułowania przesłanek do dalszych badań na temat tego, jak historie, które nosimy w sobie, w sposób nieuświadomiony kształtują nasz stosunek do sztucznej inteligencji. Czy właściwym kierunkiem jest „osławianie” – czynienie swojskim, czy też raczej powinniśmy zmierzać w wyraźnego nakreślenia granicy między tym, co czyni nas ludźmi: ograniczonością życia, kruchością biologicznego ciała, porywami serca, nad kosztem chłodnej kalkulacji. Osławianie sztucznej inteligencji – w świetle przeprowadzonej wyżej argumentacji – sprowadzałoby się przede wszystkim do rozpoznania opowieści, które w sobie nosimy, oraz uświadomienia, w jaki sposób warunkują one przyłączanie nowych elementów.

<sup>48</sup> C. P e n l e y, A. R o s s, *Cyborgs at Large: Interview with Donna Haraway*, „Social Text” 25(1990) nr 6, s. 10.

<sup>49</sup> D. J. H a r a w a y, *How Like a Leaf: An Interview with Thyra Nichols Goodeve*, Routledge, New York – London 2000, s. 107.

<sup>50</sup> M c E w a n, dz. cyt., s. 9.

<sup>51</sup> Por. J. J o s h u a, *Information Bodies: Computational Anxiety in Neal Stephenson’s Snow Crash*, „Interdisciplinary Literary Studies” 19 (2017), nr 1, s. 17-47.

## BIBLIOGRAFIA / BIBLIOGRAPHY

- Bostrom, Nick. "Transhumanist Values." In *Ethical Issues for the 21st Century*. Edited by Frederick Adams. Oxford: Philosophical Documentation Center Press, 2003.
- . *Superinteligencja. Scenariusze, strategie, zagrożenia*. Translated by Dorota Konowrocka-Sawa, Gliwice: Helion, 2016.
- Encyclopedia Britannica*, s.v. "Artificial Intelligence (AI)" (by Braian Jack Cope-land) <https://www.britannica.com/technology/artificial-intelligence/Methods-and-goals-in-AI>.
- Czym jest sztuczna inteligencja*, Portal sztucznej inteligencji. <https://www.gov.pl/web/ai/czym-jest-sztuczna-inteligencja2>.
- Damar, Muhammet. "Metaverse Shape of Your Life for Future: A Bibliometric Snapshot." *Journal of Metaverse* 1, no. 1(2021): 1–8.
- Devlin, Kate. *O pożądaniu, nauce i sztucznej inteligencji*. Translated by Anna Gralak. Wrocław: Wydawnictwo Dolnośląskie, 2020.
- Haraway, Donna J. *How Like a Leaf: An Interview with Thyrza Nichols Goodeve*. London and New York: Routledge, 2000.
- Hayles, N. Katherine. *How We Become Posthumans: Virtual Bodies in Cybernetics, Literature and Informatics*. Chicago: The University of Chicago Press, 1999.
- Hayles Katherine. "Afterword: The Human in the Posthuman." *Cultural Critique* 53 (2003): 134–6.
- Joshua, Judy. "Information Bodies: Computational Anxiety in Neal Stephenson's *Snow Crash*." *Interdisciplinary Literary Studies* 19, no. 1 (2017): 17–47.
- Kevin, Makice. *Phatics and the Design of Community*. Boston: ACM, 2009.
- Kraus, Sascha, et al. "Facebook and the Creation of the Metaverse: Radical Business Model Innovation or Incremental Transformation?" *International Journal of Entrepreneurial Behaviour & Research* 28, no. 9 (2022): 52–77.
- Lem, Stanisław. *Dzienniki gwiazdowe*. Kraków: Wydawnictwo Literackie, 2019.
- Łotman, Jurij. *O znaczeniach we wtórnych systemach modelujących*. Translated by Jerzy Faryno. „Pamiętnik Literacki. Czasopismo kwartalne poświęcone historii i krytyce literatury polskiej” 60, no. 1 (1969): 279–94.
- Malinowski, Bronisław. "The Problem of Meaning in Primitive Languages." In *The Meaning of Meaning*. Edited by Charles Key Ogden and Ivor Armstrong Richards. London: Paul, 1923.
- McEwan, Ian. *Machines Like Me*. London: Jonathan Cope, 2019.
- . *Maszyny takie jak ja*. Translated by Andrzej Szulc. Warszawa: Albatros, 2019.
- Meyers, Andrew. "Stanford's John McCarthy, Seminal Figure of Artificial Intelligence, Dies at 84." *Stanford Report*, October 25, 2011. <https://news.stanford.edu/news/2011/october/john-mccarthy-obit-102511.html>.
- Penley, Constance, and Andrew Ross. "Cyborgs at Large: Interview with Donna Haraway." *Social Text* 25, no. 6(1990): 8–23.



- Radovanovic, Danica, and Massimo Ragnedda. *Small Talk in the Digital Age: Making Sense of Phatic Posts*, [https://www.academia.edu/1506309/Small\\_talk\\_in\\_the\\_Digital\\_Age\\_Making\\_Sense\\_of\\_Phatic\\_Posts](https://www.academia.edu/1506309/Small_talk_in_the_Digital_Age_Making_Sense_of_Phatic_Posts).
- de Saint-Exupéry, Antoine. *Mały księżę*. Translated by Anna Trznadel-Szczepanek. Warszawa: Nasza Księgarnia, 2017.
- Slotznick, Benjamin. "Understanding Phatic Aspects of Narrative when Designing Assistive and Augmentative Communication Interfaces." *International Journal of Ambient Computing and Intelligence* 6, no. 2 (2014): 75–94.
- Słownik języka polskiego*. Edited by Witold Doroszewski. Warszawa: Wydawnictwo Naukowe PWN, 1997.
- Sparkes, Matthew. "What is a Metaverse." *New Scientist* 251, no. 3348 (2021): 1–18.
- Stephenson, Neal. *Snow Crash*. New York: Bantam Books, 1992.
- . *Zamieć*. Translated by Jędrzej Polak. Warszawa: ISA, 2009.
- Wawer, Robert. „Sztuczna inteligencja – produkt czy podmiot?” *Studia Teologii Dogmatycznej* 4 (2018): 115–26. [https://repozytorium.uwb.edu.pl/jspui/bitstream/11320/7817/1/STD\\_4\\_2018\\_R\\_Wawer\\_Sztuczna\\_inteligencja.pdf](https://repozytorium.uwb.edu.pl/jspui/bitstream/11320/7817/1/STD_4_2018_R_Wawer_Sztuczna_inteligencja.pdf).
- Wiener, Norbert, *Cybernetics: Or Control and Communication in the Animal and the Machine*, Cambridge: MIT Press, 1965.
- . *God and Golem, Inc.: A Comment on Certain Points Where Cybernetics Impinges on Religion*. Cambridge: MIT Press, 1964.
- Zalta, Edward N., ed. *The Stanford Encyclopedia of Philosophy*, s.v. "Alan Turing" (by Andrew Hodges). <https://plato.stanford.edu/entries/turing/>.

#### ABSTRAKT / ABSTRACT

Małgorzata ABASSY – Snucie opowieści jako narzędzie osvajania sztucznej inteligencji. Między modelującą funkcją języka a podmiotowością maszyn  
DOI 10.12887/36-2023-4-144-06

W artykule postawiono pytanie o rolę narracji jako narzędzia kreowania rzeczywistości. Performatywna funkcja języka oraz znaczenie pierwotnych i wtórnych systemów modelujących kultury w procesie osvajania sztucznej inteligencji poddano analizie w oparciu o materiał źródłowy dwóch tekstów literackich: *Zamieć* Neala Stephensona oraz *Maszyny takie jak ja* Iana McEwana. Podstawowe pytania badawcze sygnalizowane w niniejszym artykule dotyczą tych obszarów postrzegania rzeczywistości, które są zagrożone zniekształceniem fundamentu aksjologicznego, i w konsekwencji – ontologicznego. Które z elementów snutych opowieści powinny budzić szczególną czujność słuchającego i opowiadającego? Które wątki fantastycznej w 1992 roku powieści Stephensona nie tylko stały się faktem, ale też wykroczyły poza ramy nakreślone przez wyobraźnię autora? Czy Adam wykreowany przez McEwana staje się postacią zapraszaną do intymnego świata istoty ludzkiej jako jej równoprawny członek? Jak daleko

przesuwamy granice osvajania sztucznej inteligencji poza niwelowanie lęku przed nieprzewidywalnością maszyn? Struktura artykułu obejmuje część wstępną, namysł nad semantyką pojęć „oswajać” i „sztuczna inteligencja”, *Zamieć* – jako egzemplifikacja problemu pierwotnego systemu modelującego, *Maszyny takie jak ja* i kwestia granic człowieczeństwa, Podsumowanie odnoszące się do roli narracji w ujęciu Katherine Hayles i Donny Haraway oraz wnioski. Konkludując, zaznaczono, że obydwie powieści, wybrane jako materiał źródłowy do analizy mechanizmów osvajania sztucznej inteligencji, zawierają w sobie przesłanki do odczytania zarówno obaw, jak i nadziei pokładanych w rozwoju technologicznym. Poprzez ich werbalizację na kartach książki, a następnie wizualizację w wyobraźni czytelnika rodzi się kolejna historia, dając impuls do zaistnienia nowej rzeczywistości. Oswajanie sztucznej inteligencji, w świetle przeprowadzonej argumentacji sprowadzałoby się przede wszystkim do rozpoznania opowieści, które w sobie nosimy, oraz uświadomienia, w jaki sposób warunkują one przyłączanie nowych elementów.

Słowa kluczowe: sztuczna inteligencja, podmiotowość, język, narracje

Kontakt: Instytut Rosji i Europy Wschodniej, Uniwersytet Jagielloński, ul. Władysława Reymonta 4, 30-059 Kraków  
E-mail: malgorzata.abassy@uj.edu.pl  
Google Scholar: 5HDH-8MAAAAJ  
ORCID: 0000-0002-9294-6156

Małgorzata ABASSY, Storytelling as a Tool to Tame Artificial Intelligence: Between the Modeling Function of Language and the Subjectivity of Machines

DOI 10.12887/36-2023-4-144-06

The article poses a question about the role of narration as a tool for creating reality. The phatic function of language and the significance of primary and secondary systems modeling cultures in the process of taming artificial intelligence were analyzed, based on the source material of two literary texts: Neal Stephenson's *Snow Crash* and *Machines Like Me* by Ian McEwan. The basic research questions signaled in this article concern those areas of perception of reality that are at risk of distorting the axiological foundation and, consequently, the ontological basis. Which of the elements of the storytelling should arouse a particular vigilance of the listener and the narrator? Which themes in Stephenson's fantasy novel written in 1992 not only became a fact nowadays, but also went beyond the framework outlined by the author's imagination? Does Adam created by McEwan become the character who was invited to the intimate world of human beings as its equal member? How far are we pushing the boundaries of taming AI beyond eliminating the fear of machine unpredictability? The structure of the article includes an introductory part, a reflection on the semantics of the concepts of taming and artificial intelligence, *Snow Crash*—as an exemplification of the problem of the primary modeling system, *Machines Like*



*Me* and the question of the limits of humanity, Katherine Hayles' and Donna Haraway's views on the role of narratives, and conclusions. It was noted that the novels selected as a source material for the analysis of the mechanisms of taming artificial intelligence contain premises for reading both fears and hopes placed in technological development. Through their verbalization on the pages of the book, and then visualization in the reader's imagination, another story is born, giving an impulse to the emergence of a new reality. Taming artificial intelligence, in the light of the presented argumentation would boil down primarily to recognizing the stories that we carry in ourselves and realizing how they condition the connection of new elements.

Keywords: artificial intelligence, subjectivity, language, narratives

Contact: Institute of Russian and East European Studies, Jagiellonian University, ul. Władysława Reymonta 4, 30-059 Cracow, Poland

E-mail: malgorzata.abassy@uj.edu.pl

Google Scholar: 5HDH-8MAAAAJ

ORCID: 0000-0002-9294-6156



SZTUCZNA INTELIGENCJA W ŚWIECIE PRZEDSTAWIONYM



Andrzej BORKOWSKI

## SZTUCZNA INTELIGENCJA A PROBLEM JĘZYKA I TEKSTU W kręgu „Golema XIV” Stanisława Lema

*Lem niezwykle sugestywnie przedstawia rozwój sztucznej inteligencji, widząc w niej nieskończony potencjał wzrostu. Kreśli jednak ustami myślącej maszyny labiryntowy scenariusz jej rozwoju, być może analogiczny do teorii powstania życia jako takiego, czyli „zbiegu trafów”. Koncepcja Lema zakłada, iż język etniczny i powstające zeń teksty są niejako odbiciem pracy „kodu”, który tworzy dzieła coraz bardziej złożone, a przez to mniej doskonałe. Czy zatem powrót do uświęconych tradycją dawnych form komunikacji zarówno filozoficznej i akademickiej (wykład), jak też homiletycznej (kazanie) jest wskazaniem na ten, owszem niedoskonały, ale z racji braku innych najlepszy sposób porozumiewania się AI z człowiekiem?*

Kiedy mamy do czynienia z marnym utworem o sprawdzonym nadzieniu prognostycznym, jak gdybyśmy brali do ręki owoc wprowadzie zgniły i nienadający się do spożycia, lecz zawierający pestkę, w której skryty okazuje się migdał wprost wyborny<sup>1</sup>.

Stanisław Lem

Pisząc o twórczości Stanisława Lema, trudno nawet dziś, kilkanaście lat po jego śmierci – w którym to czasie technologia (zwłaszcza informatyczna) oraz wiedza jako taka dokonała poważnego skoku – nie popaść w konfuzję spowodowaną nie tyle ilością „urobku” twórczego, ale przede wszystkim jego myślowym rozmachem i aktualnością, podbudowaną interdyscyplinarną wiedzą ze skrajnie różnych dziedzin nauki<sup>2</sup>. Po drugie, zmieszanie czytelnika wynika też ze zderzenia z niezwykle rozwiniętą kompetencją kulturową twórcy, popartą znajomością ludzkiej natury czy języków obcych<sup>3</sup>. W końcu równie wielkie zakłopotanie wywołuje konfrontacja z badaniami na temat twórczości tego pisarza, filozofa i futurologa zarazem, których rozległość ujawnia chociażby monografia zbiorowa *Stanisław Lem. Fantastyka naukowa i fikcje nauki*<sup>4</sup>, wieńcząca zeszłoroczną inicjatywę Sejmu RP w kwestii uczczenia setnych

<sup>1</sup> S. L e m, *Pod klątwą przewidyźmu*, w: tenże, *Bomba megabitowa*, Wydawnictw Literackie, Kraków 1999, s. 32.

<sup>2</sup> Por. M. S z p a k o w s k a, *Dyskusje ze Stanisławem Lemem*, Open, Warszawa 1997, s. 20.

<sup>3</sup> Por. Z. T r z a s k o w s k i, *Philosophia perennis Stanisława Lema*, w: *Stanisław Lem. Fantastyka naukowa i fikcje nauki*, red. D. Heck, Narodowe Centrum Nauki, Warszawa 2021, s. 43.

<sup>4</sup> Zob. *Stanisław Lem. Fantastyka naukowa i fikcje*.

urodzin pisarza poprzez ustanowienie Roku Stanisława Lema<sup>5</sup>, jak też – przede wszystkim – pokaźna „biblioteka lemologiczna”, w której mieszczą się ważne monografie, tomy zbiorowe i artykuły poświęcone wybranym aspektom jego twórczości<sup>6</sup>.

Rozpiętość myślowa i genologiczna dzieł Lema, a zarazem ich ideowa jednorodność – chodzi zwłaszcza o problemy ewolucji, miejsce człowieka w kosmosie, rozumu, stosunku do życia, ale też sztucznej inteligencji (ang. Artificial Intelligence – AI) – początkowo porażają badacza-czytelnika, który w sensie metaforycznym (odwołując się do przywołanego przez Lema przykładu pewnej odmiany osy) występuje niejako w roli konika polnego, stającego się pożywieniem dla potomstwa przemysłowego owada. Podobnie czytelnik w kontakcie z dziełami Lema musi niejako przepoczwaczyć się intelektualnie, aby stać się rozumnym odbiorcą jego przekazu<sup>7</sup>. Słowem, Lem intelektualnie zniewala, rozumnie falsyfikując – mówiąc językiem Karla R. Poppera – skostniałe prawdy na temat życia czy natury ludzkiej, prawdy pokryte niekiedy patyną uświęconego obyczaju czy też powierzchownej religijności, i jest w tym uparty, a często bezkompromisowy. Przeciętny badacz, chcąc w sposób rozumny mierzyć się z myślową akrobatyką, skądinąd najwyższej miary, autora *Powrotu z gwiazd*<sup>8</sup> (1961), zmuszony jest (z racji talentu czy wiedzy, a niekiedy zaś przekonania) do pozostawienia na marginesie wielu frapujących kwestii merytorycznych, skupiając się na tych obszarach dokonań Lema, które mieszczą się w obszarze jego kompetencji, jak też mogą okazać się po prostu pożyteczne poznawczo.

Biorąc pod uwagę powyższe kwestie, należy sensownie uściślić wywód, dokonując istotnych zawężeń w zakresie wyboru obiektu badań oraz jego

<sup>5</sup> Por. D. H e c k, *Stanisław Lem, czyli uniwersalna wartość interdyscyplinarności*, w: *Stanisław Lem. Fantastyka naukowa i fikcje nauki*, s. 11.

<sup>6</sup> Zob. m.in.: A. S t o f f, *Powieści fantastyczno-naukowe Stanisława Lema*, PWN, Warszawa 1983; t e n ż e, *Lem i inni. Szkice o polskiej science fiction*, Pomorze, Bydgoszcz 1990; J. J a r z ę b s k i, *Wszelświat Lema*, Wydawnictwo Literackie, Kraków 2002; M. P ł a z a, *O poznaniu w twórczości Stanisława Lema*, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 2006; P. M a j e w s k i, *Między zwierzęciem a maszyną. Utopia technologiczna Stanisław Lema*, Wydawnictwo Uniwersytetu Wrocławskiego, Wrocław 2007; I. P i ę t a, *Problemy intertekstualnego obrazowania w wybranych powieściach Stanisława Lema*, Adam Marszałek, Toruń 2002; M. D a j n o w s k i, *Groteska w twórczości Stanisława Lema*, Wydawnictwo UG, Gdańsk 2005; t e n ż e, *Pejzażysta Lem. Szkice z motywyki*, Wydawnictwo UG, Gdańsk 2010; Ł. K u c h a r c z y k, *Granice ciała. Somapoetyka w twórczości Stanisława Lema*, Wydawnictwo Naukowe UKSW, Warszawa 2021; *Lem w oczach krytyki światowej*, red. J. Jarzębski, Wydawnictwo Literackie, Kraków 1989; *Stanisław Lem – pisarz myśliciel, człowiek*, red. J. Jarzębski, A. Sulikowski, Wydawnictwo Literackie, Kraków 2003. Instruktywne są także komentarze, wstępy i posłowania do dzieł Lema, zwłaszcza autorstwa Jerzego Jarzębskiego.

<sup>7</sup> Por. S. L e m, *Sztuczna nieinteligencja*, w: tenże, *Tajemnica chińskiego pokoju*, Universitas, Kraków 1996, s. 218.

<sup>8</sup> Zob. t e n ż e, *Powrót z gwiazd*, Czytelnik, Warszawa 1961.

przedmiotu. W tak nakreślonym polu badawczym miejsce swe znajdzie przede wszystkim znakomity *Golem XIV*<sup>9</sup> (książka ukazała się w roku 1981, lecz jej fragmenty zostały opublikowane wcześniej, w roku 1973, w *Wielkości urojonej*), który czytany jest przez pryzmat refleksji o języku i tekście (co mówi Lem ustami sztucznej inteligencji na ten temat, jakie wynikają stąd możliwości i ograniczenia w budowaniu skutecznej komunikacji człowieka z maszyną, czy jest to jedyny sposób porozumiewania się z AI, jakiego języka używają superkomputery w komunikacji pomiędzy sobą, czy komunikacja musi mieć tekstowy charakter) oraz jak mówi (w jaki sposób wywód AI jest konstruowany, z jakiego podglebia tradycji retorycznej wyrasta, z jakiego repertuaru gatunków oraz form tekstowych korzysta myślący automat i dlaczego). W tym swoistym filologicznym śledztwie na temat przede wszystkim tego, w jaki sposób komunikuje się AI<sup>10</sup>, w dziełach Lema pomagają oczywiście inne utwory autora *Bajek robotów* (1964), w tym *Summa technologiae* (1964), a także, będące niejako jej kontynuacją, niezwykle treściwe drobiazgi esejistyczne skupione w zbiorach *Tajemnica chińskiego pokoju* (1996) oraz *Bomba megabitowa* (1999).

Inspiracje metodologiczne osadzone są na gruncie ściśle filologicznym i uwzględniają zwłaszcza odniesienia do tradycji retorycznej, a także podbudowane są refleksją hermeneutyczną (notabene niecenioną przez głównego bohatera utworu)<sup>11</sup>. Badacz, wbrew Lemowi, oddaje tu rację uniwersalnej myśli Hansa-Georga Gadamera, iż to właśnie rozumienie ożywia martwe ślady sensu<sup>12</sup>. W tej perspektywie twórczość Lema jawi się jako wyzwanie dla rozumiejącego filologa, który mierzy się z ograniczeniami własnej wiedzy o świecie, a zarazem jest w stanie niejako odkrywać pewne przestrzenie sensu miarą swych doświadczeń egzystencjalnych (w tym lekturowych). Jednocześnie interpretatorowi towarzyszy, podczas tych dociekań na temat językowo-tekstowych uwarunkowań komunikacji z AI, ważka odpowiedź sformułowana przez Umberto Eco, wskazującego na swego rodzaju miarę (modus) interpretacji, co pomaga uchronić się przed pułapką nadmiaru, ale też niedomiaru<sup>13</sup>. W tym sensie badacz-czytelnik przypomina nieco linoskoczek, który zmuszony jest zachować odpowiedni balans, by nie osunąć się w prze-

<sup>9</sup> Zob. t e n z e, *Golem XIV*, posłowie J. Jarzębski, Wydawnictwo Literackie, Kraków 2021.

<sup>10</sup> Najogólniej sztuczna inteligencję definiuje się jako „zdolność komputera cyfrowego lub robota, który jest sterowany przez komputer, do wykonywania zadań w sposób powszechnie kojarzony ze stworzeniami inteligentnymi” (hasło „The Turing Test”, Britannica, <https://www.britannica.com/technology/artificial-intelligence/The-Turing-test>). Tłum. fragm. – A.B.

<sup>11</sup> Por. t e n z e, *Golem XIV*, s. 115.

<sup>12</sup> Por. H.-G. G a d a m e r, *Prawda i metoda. Zarys hermeneutyki filozoficznej*, tłum. B. Baran, Inter Esse, Kraków 1993, s. 173.

<sup>13</sup> Por. U. E c o, *Czytanie świata*, tłum. M. Woźniak, Znak, Kraków 1999, s. 23.



paść wielości sensów możliwych lub nie stanąć w połowie drogi, tracąc z horyzontu cel badawczej eksploracji.

Zagadnienie sztucznej inteligencji to niewątpliwie ważki lejtmotyw w twórczości ściśle literackiej oraz eseistycznej Lema. Trudno w zasadzie wskazać jakiś utwór prozatorski, począwszy od – by wymienić jedynie niektóre – *Dialogów* (1957) ukształtowanych ma modłę wywodów platońskich, dzieła *Summa technologiae*, a także *Bajek robotów* czy *Cyberiady* (1965), aż po *Wielkość urojoną* oraz *Golema XIV*, w których pisarz pośrednio lub wprost nie dotyka tego zagadnienia. Wiele uwag na ten temat sformułował pisarz-filozof w zbiorach *Tajemnica chińskiego pokoju* i *Bomba megabitowa*. Znakomity znawca dzieł autora *Filozofii przypadku* (1968) Jerzy Jarzębski zapisał, że w kręgu trosk Lema mieści się problem „myślącej” maszyny, a w perspektywie „próba wydarcia Bogu kolejnych prerogatyw”<sup>14</sup> w zakresie chociażby wykreowania sztucznego bytu „rozumnego” oraz pokonania barier genetycznych, otwierających nieznanne wcześniej człowiekowi możliwości w zakresie powoływania do życia nowych stworzeń. Owej demiurgicznej potencji zagrażają jednak ryzyka etycznych nadużyć aż po kres znanego dotychczas ludzkości ładu cywilizacyjnego<sup>15</sup>.

Zanim jednak przejdziemy do powieści *Golem XIV*, która stanowi zasadniczy obiekt badawczych eksploracji, warto przyjrzeć się – w pewnym rzecz jasna skrócie – refleksji eseistycznej Lema na temat sztucznej inteligencji, a przede wszystkim kwestii sposobu komunikowania się z tą, a zatem jej sprawności nadawania i odbioru, „mówienia” i „słuchania”, a mówiąc wprost – rozumnego „rozmawiania” czy też „dialogowania”. W esejach *Duch w maszynie*<sup>16</sup> oraz *Człowiek i maszyna* krakowski autor, posługując się notabene szkolną metodą „opowiedz to własnymi słowami”, przypomina tak zwany test Turinga, który ujawnia lub nie zdolność automatu do posługiwania się językiem naturalnym w ten sposób, by nie można było w istocie rozpoznać, czy człowiek rozmawia z osobą czy też sztucznym tworem. Lema frapuje kwestia demaskatorska, a zatem, jak najłatwiej odsłonić nierozumne oblicze maszyny: „Sposobem najprostszym, na który (nie wiem czemu) jakoś nikt albo omal nikt nie wpadł, byłoby opowiedzenie niewiadomemu dyskutantowi dowolnej historyjki, np. anegdoty, dowcipu i żądanie, ażeby opowiedziane po chwili własnymi słowami powtórzył. Maszyna najłatwiej jeszcze powtórzy wszystko literalnie: normalny człowiek dlatego nie jest w odtwarzaniu narracji

<sup>14</sup> Por. J. J a r z ę b s k i, *Rozum ewolucji i ewolucja rozumu*, w: Lem, *Tajemnica chińskiego pokoju*, s. 11.

<sup>15</sup> Tamże, s. 10n.

<sup>16</sup> Por. S. L e m, *Summa technologiae*, Wydawnictwo Lubelskie, Lublin 1984, s. 113.

precyzyjny, ponieważ nie tyle tekst złożony ze słów zapamiętuje, ile sensy, a to znaczy, że opowiedziane rozumie<sup>17</sup>.

Powyższa refleksja wskazuje, że człowiek posługuje się również efemerycznymi gatunkami mowy, czego dowodził już Michaił Bachtin<sup>18</sup> (w tym przypadku są to mniej lub bardziej skodyfikowane typy, jak na przykład historyjka, anegdota, dowcip), oraz – co dla Lema istotniejsze – ma zdolność pojmowania znaczenia (sensu) tekstu, który podlega prawu streszczenia czy też parafrazy. Chociaż, jak się wydaje, myśl polskiego filozofa-pisarza (swoją drogą zadeklarowanego „antysemioty”<sup>19</sup>) można umieścić w kręgu dociekań amerykańskiego pragmatysty Charlesa S. Peirce’a, który akcentował kwestię przekładalności znaku na inny znak (semioza)<sup>20</sup>, co jest warunkiem zrozumiałości znaku czy też wypowiedzi jako takiej, to jednak Lem w swej refleksji lingwistycznej chętniej sięgał do ustaleń Wasilija Nalimowa czy Noama Chomsky’ego, widząc w tych koncepcjach pewne możliwości praktyczne dla rozwoju „obliczeń maszynowych”<sup>21</sup>. O ile jednak maszyna, bądź jej software (program), jest w stanie wykonywać polecenia człowieka w zakresie chociażby przekładu tekstu z jednego języka na inny<sup>22</sup>, to mechanizm parafrazy wydaje się jej niedostępny czy też został on w pewien sposób zaniedbany przez programistów, choć i tu pisarz widzi szanse na „douceń” maszyny<sup>23</sup>. W stworzeniu sztucznej inteligencji przeszkadzają człowiekowi uwarunkowania natury ludzkiej, gdyż zdolności regeneracyjne mózgu („rany komputerów” nie goją się w przeciwieństwie do tkanki ludzkiej)<sup>24</sup>, a ponadto jego życie afektywne

<sup>17</sup> Por. t e n ż e, *Człowiek i maszyna*, w: tenże, *Tajemnica chińskiego pokoju*, s. 15n.; por. też t e n ż e, *Rozmyślania nad siecią*, w: tenże, *Bomba megabitowa*, s. 52. Wydaje się, że obecnie Chat GPT opanował już w znacznym stopniu parafrazowanie i eksplikację. Wyzwaniem staje się kwestia swojej „przyjaźni” czy też „pamięci relacji” między programem a człowiekiem.

<sup>18</sup> Por. T. D o b r z y Ń s k a, *Tekst – próba syntezy*, w: *Problemy teorii literatury. Seria 4*, wybór H. Markiewicz, Zakład Narodowy im. Ossolińskich, Wrocław 1998, s. 43.

<sup>19</sup> Por. S. L e m, *Mój pogląd na literaturę. Rozprawy i szkice*, posłowie J. Jarzębski, Wydawnictwo Literackie, Kraków 2003, s. 48-65.

<sup>20</sup> Lem krytycznie odnosił się do dekonstrukcjonizmu spod znaku Jacquesa Derridy, jak też niezrozumiałości współczesnych dzieł plastycznych czy poetyckich, które niejako nazbyt „rozmiękcza” swój język, co prowadzi do utraty sensowności takiego przekazu (por. S. L e m, *Języki i kody*, w: *Tajemnica chińskiego pokoju*, s. 130). Warto dodać, że niejako w opozycji do polskiego filozofa-pisarza sytuują się poglądy José Ortegi y Gasset, który stwierdził, iż „sztuka czasu przeszłego nie jest sztuką, była sztuką” (J. O r t e g a y G a s s e t, *Sztuka w czasie teraźniejszym i przeszłym*, w: *Dehumanizacja sztuki i inne eseje*, tłum. P. Niklewicz, Muza, Warszawa 1996, s. 167).

<sup>21</sup> Por. W. N a l i m o w, *Probabilistyczny model języka*, tłum. I. Kustrzeba, PWN, Warszawa 1976, s. 190; L e m, *Języki i kody*, s. 123, 128.

<sup>22</sup> Lem przekonywał, że „oryginał stanowi centralny punkt pewnej tarczy strzeleckiej, poszczególne zaś przekłady są trafieniami tylko asymptotycznie zdolnymi zbliżyć się do owego centrum: pokryć się z nim w 100% nie mogą” (L e m, *Języki i kody*, s. 129).

<sup>23</sup> Por. t e n ż e, *Człowiek i maszyna*, s. 17.

<sup>24</sup> Por. t e n ż e, *Komputeryzacja mózgu*, w: tenże, *Tajemnica chińskiego pokoju*, s. 111.

(reakcje na zagrożenia, wyrażanie sprzeciwu, bunt, sny)<sup>25</sup> jest nadzwyczaj trudno symulować. Mimo to Lem dostrzega szansę na stopniowe zbliżanie się automatu do zachowań językowych „rozumnego interlokutora”<sup>26</sup>. Warunkiem jest jednak osadzenie AI w strefie emocji: „Lodowata «inteligencja» nie bardzo może stawać się analogiem ludzkiej”<sup>27</sup>.

W swej eseistyce Lem skupia się przede wszystkim na kwestii świadomości czy też rozumności sztucznej inteligencji, której wyrazem mogą być owe zachowania językowe. Jest to, jak może się wydawać, jedno z ważniejszych wyzwań dla konstruktorów AI. Zdaniem autora *Bezsenności* (1971) problem porozumiewania się z maszyną wynika z tego, że ludzkość posługuje się językami „kontekstowymi”, które są wplątane w sieć zależności semantycznych (odmiennie niż komputery), ale i tu dostrzega pisarz możliwość doposażenia maszyn w „ogrom pamięci uaktywnianej kontekstowo”<sup>28</sup>. Ponadto Lem zwraca uwagę na kwestie swoistości języka, jego osobliwej słabości i mocy, co umożliwia prowadzenie dyskursu przy użyciu argumentów o różnej sile, a także „przedprogramowej” niejako zdolności małych dzieci do szybkiej adaptacji do konkretnego otoczenia językowego, w którym przyszło im od urodzenia funkcjonować. Mają one zdolność programowania w okresie pierwszych lat życia kilku języków w sposób naturalny niejako „bezwysiłkowo”<sup>29</sup>. Lem stwierdza w tym kontekście, że ta „bezkontekstowa sztywność [komputerów] musi kiedyś ulec zasadniczej «przeróbce» w stronę antropologiczną”<sup>30</sup>, dodając, że „różnica, zachodząca [...] między mózgami żywymi i maszynami typu automatów skończonych, stanowi bodaj największą przeszkodę do pokonania dla wszystkich usiłujących stworzyć AI”<sup>31</sup>.

Autor *Doskonałej próżni* (1971) analizuje semantyczną stronę języka, co ma zresztą poważane implikacje dla budowania pomostów komunikacyjnych ze sztuczną inteligencją, wskazując na zagrożenia związane z jego zbytnią „twardością” lub „miękkością”. Owa matematyczna „twardość” (jednoznaczność) prowadzi filozofa do konceptu „babilońskiej encyklopedii”, czyli paraliżującego niejako porozumiewanie się nadmiaru słów, gdzie każde słowo służy oznaczaniu jednego tylko przedmiotu. Przestrzega również przed zbytнім zmiękczeniem języka, czyli nadmiarem metafor, charakteryzujących

<sup>25</sup> Por. t e n ż e, *Człowiek i maszyna*, s. 16; t e n ż e, „Bariera informacyjna?”, w: tenże, *Tajemnica chińskiego pokoju*, s. 32; t e n ż e, *Tertium comparationis*, w: tenże, *Tajemnica chińskiego pokoju*, s. 66.

<sup>26</sup> T e n ż e, *Tajemnica chińskiego pokoju*, s. 20.

<sup>27</sup> T e n ż e, *Zastąpić rozum?*, w: tenże, *Bomba megabitowa*, s. 80.

<sup>28</sup> T e n ż e, *Fantomatyka (II)*, w: *Tajemnica chińskiego pokoju*, s. 58.

<sup>29</sup> T e n ż e, *Umysł jako sternik*, w: tenże, *Bomba megabitowa*, s. 18.

<sup>30</sup> T e n ż e, *Tertium comparationis*, s. 66.

<sup>31</sup> T e n ż e, *Języki i kody*, s. 124.

mowę „schizofrenika”<sup>32</sup>. Lem niejako na marginesie stwierdza (tłumacząc to budową mózgu), że żyjemy w czasach, gdzie wkrada się „beźjęzykowość”, której podglebiem są zwierzęce atawizmy. Pisarz wiąże ten proces poniekąd z rozwojem mediów wizualnych, jednocześnie podkreślając, iż człowiek (co niewątpliwie daje do myślenia) skłonny jest przekazywać „rosnący [...] trud myślenia” – maszynom<sup>33</sup>.

Sporo uwagi poświęca pisarz kwestii wniknięcia technologii w mózg ludzki, wskazując na liczne tropy na ten temat rozsiane we wczesnej fazie twórczości<sup>34</sup>. Lem zauważa, że poprzez rozwój językowy ludzkość wyręła sobie drogę do świadomości, która „jest zawsze bogatsza od mowy, ale wokół niej narosła – przynajmniej u człowieka”<sup>35</sup>. Ponadto dzięki więziom socjalnym język mógł powstać w kształcie, jaki znamy (wraz z jego semantyką, syntaktyką czy idiomatyką); ta potencja zamiera jednak (na przykład fenomen tak zwanych „wilczych dzieci”), gdy nie jest rozwijana<sup>36</sup>. W nanotechnologii filozof upatruje wiele szans w lecznictwie (neurologia), pozostając jednak dość sceptyczny, co do łączenia się mózgu ludzkiego poprzez „czipy” z komputerem. Słowem, można stwierdzić, iż łatwiej protezować czy też leczyć człowieka, chociażby przez nanotechnologie, niż z ich pomocą wytransferować umysłowość w świat maszyn czy sieć<sup>37</sup>.

W eseju *Zagadki* Lem zwraca uwagę na fakt, iż literatura na temat AI jest dziś tak rozległa, iż nie sposób „przedrzeć się” przez mnogość tych prac. Można jednak stawiać pytania natury bardziej ogólnej, chociażby – czy rozum i inteligencja są sprawnościami potencjalnie rozłącznymi oraz – jak jest możliwe stwierdzenie „świadomości” maszyny<sup>38</sup>. Co z kwestią instynktów oraz wolicjonalności automatów<sup>39</sup>? I tu powoli zmierzamy do książki *Golem XIV*, gdzie (jak dowodzi autor) „możliwy jest rozum bezosobowy”<sup>40</sup>. Jak przekonywał sam twórca, motywacją do napisania tej „opowieści” czy też

<sup>32</sup> Tamże, s. 126, 131.

<sup>33</sup> Tamże, s. 132.

<sup>34</sup> Por. t e n ż e, *Brain chips*, w: tenże, *Tajemnica chińskiego pokoju*, s. 141. Mowa jest między innymi o powieści *Wizja lokalna* (1982), gdzie świat Encjan zarządzany jest przez „bystry” – „molekuły dobra”. Por. też komentarz autora do utworu Lema *Kuratela komputerów* (por. S. L e m, *Kuratela komputerów*, w: tenże, *Bomba megabitowa*, s. 141).

<sup>35</sup> Tamże, s. 149; por. też S. L e m, *Symulowanie kultury*, w: tenże, *Tajemnica chińskiego pokoju*, s. 276.

<sup>36</sup> Por. t e n ż e, *Brain chips (III)*, w: tenże, *Tajemnica chińskiego pokoju*, s. 161, 165.

<sup>37</sup> Tego rodzaju wizje mnożą się we współczesnym kinie, by przywołać jako przykład film *Transcendencja* (2014) w reżyserii Wallego Pfistera (w roli głównej Johnny Depp).

<sup>38</sup> Por. L e m, *Tajemnica chińskiego pokoju*, s. 207.

<sup>39</sup> Por. t e n ż e, *Zmysłne roboty*, w: tenże, *Tajemnica chińskiego pokoju*, s. 267n.; t e n ż e, *Sztuczna nieinteligencja*, s. 223.

<sup>40</sup> T e n ż e, *Zagadki*, w: tenże, *Tajemnica chińskiego pokoju*, s. 197; por. też: J. J a r z ę b s k i, *Wszechświat Lema*, Wydawnictwo Literackie, Kraków 2002, s. 289.

„powiastki”<sup>41</sup>, a przede wszystkim kreacji głównej postaci (superkomputera) była zimna wojna, a zatem z grubsza polityczno-militarne napięcie między komunistycznym Wschodem i demokratycznym Zachodem. Lem podkreślał, że nie tylko rozpad ZSRR usunął w pewnym sensie tę podniecie, ale też uwarunkowania ekonomiczne tego rodzaju projektów, skądinąd zależnych od rachub politycznych. Autor powieści *Solaris* (1961) zapisał: „Tak naprawdę nikt z możliwych tego świata Rozumu sobie nie życzy Zbyt Mądrego, a zwłaszcza już politycy, którzy będą się zawsze potajemnie lub i jawnie obawiać, że im taki Rozum „uwiedzie” elektoraty w demokratycznie rządzonych państwach, zaś w niedemokratycznych może bądź to dyktatury rozsądzać dzięki przemysłowości swojej, bądź też rozkruszać fundamentalizmy religijne”<sup>42</sup>.

Dokonując analizy dzieła, nie sposób pominąć tajemniczy tytuł, który jest zarazem imieniem głównego bohatera. Jak można przeczytać, superkomputer przynależał do „długiej serii Golemów (General Operator, Longrange, Ethically Stabilized, Multimodelling)”<sup>43</sup>, czyli – w wolnym tłumaczenia – jest jakby zarządcą, który posiada duży zasięg, a ponadto jest etycznie ustabilizowany i niejako „wielo-upostaciowiony”<sup>44</sup>. Trudno tu jednak nie dostrzec symbolicznych odniesień. Jak dowodził Gershom Scholem: „*Golem* to hebrajskie słowo, które występuje w Biblii tylko w jednym miejscu, w Psalmie 139,16, który tradycja żydowska wkłada zawsze w usta samego Adama. *Golem* oznacza tutaj, i z pewnością w późniejszych źródłach, coś jeszcze nie ukształtowanego, pozbawionego formy. [...] Średniowieczna literatura filozoficzna używa tego słowa jako hebrajskiego terminu oznaczającego pozbawioną formy materię, *hyle*”<sup>45</sup>. Ponadto nie sposób zignorować również funkcji postaci Golema w nauce<sup>46</sup> i kulturze popularnej, w tym także innych dziełach krakowskiego autora<sup>47</sup>. Na koniec trudno nie odnieść wrażenia, że możliwa jest, także w kontekście kształtu narracyjnego utworu, pewna gra autora z sobą samym i czytelnikiem (Lem to Go-Lem)<sup>48</sup>.

<sup>41</sup> Por. S. L e m, *Teoria i praktyka literatury*, w: tenże, *Filozofia przypadku. Literatura w świetle empirii II*, Wydawnictwo Literackie, Kraków 1988, s. 102. Należy dodać, że określenie „powiastka” (filozoficzna) odsyła do tradycji literatury oświeceniowej.

<sup>42</sup> T e n ż e, *Gry w Internecie*, w: tenże, *Bomba megabitowa*, s. 44.

<sup>43</sup> T e n ż e, *Golem XIV*, s. 17.

<sup>44</sup> Tamże.

<sup>45</sup> G. S c h o l e m, *Kabala i jej symbolika*, tłum. R. Wojnakowski, Znak, Kraków 1996, s. 177.

<sup>46</sup> Por. H. C o l l i n s, T. P i n c h, *Golem, czyli co trzeba wiedzieć o nauce*, tłum. A. Tanalska-Duleba, CiS, Warszawa 1998.

<sup>47</sup> Por. M. R a d k o w s k a - W a l k o w i c z, *Od Golema do Terminatora. Wizerunki sztucznego człowieka w kulturze*, Wydawnictwo Akademickie i Profesjonalne, Warszawa 2008, s. 40-43.

<sup>48</sup> „Interpretacja *Golema XIV* nastęrcza niemałych trudności, ma on bowiem jako wywód, swój sens niejako literalny, nad tym nadbudowuje się sens wynikający ze szczególnej sytuacji podmiotu

Tytułowa maszyna (czytelnik dowiaduje się o tym z ust bohatera-naukowca Irvinga T. Creve'a) posiada swój technologiczny rodowód i jest czternastą z kolei odmianą myślącej supermaszyny, poprzedzoną formami prymitywnymi – Eniac<sup>49</sup>, a następnie bardziej złożonymi, jak Hann czy Ulvic, Ajax, a w końcu niezwykle skomplikowanymi – Ultor Gilgamesh. Ten ostatni był wykwittem ultranowoczesnej świetlnej technologii i pracował z niewiarygodną prędkością<sup>50</sup>. Wykładając dzieje AI, Irving T. Creve kreśli pewną paralelę między rozwojem technologii a ewolucją człowieka: „W istocie był Eniac komputerem, a w przymierzeniu do drzewa życia – prymitywnym ganglionem nerwowym”<sup>51</sup>, a zatem niejako naroślą (nowotworem) czy też mutacją, która dała początek nowej techno-cywilizacji<sup>52</sup>.

Tytułowy Golem miał zatem wskazanych wyżej praprzodków, a następnie antenatów z serii Golemów – od I do XIII (Creve poświęca uwagę jedynie wersjom: I, VI, XII i XIII), w tym najnowszy, obdarzony niezwykłym „potencjałem psychicznym”, model Zacnej Ani (Honest Annie), a także komputer wojskowy Cybermatics-Supermaster<sup>53</sup>. Wydaje się, że wszystkie maszyny (najmniej Supermaster i poniekąd Golem XIV) cechowały się jednak jakąś ułomnością i nie były w stanie spełnić oczekiwań konstruktorów, między innymi za sprawą nieefektywnej (przynajmniej z punktu widzenia ludzi) komunikacji. Tak było w przypadku Golema XII, który obraził senatorów (obraza odnosi się do słowa lub czynu; w tym przypadku chodzi zapewne o to pierwsze), członków specjalnej komisji, natomiast Golem XIII zawierał w sobie „nieusuwalny defekt schizofreniczny”<sup>54</sup>, który – jak wiadomo – ma również odbicie w chaotycznej i w istocie niemożliwej do zrozumienia mowie (przypominającej niekiedy ekstatyczny dukt mowy prorockiej)<sup>55</sup>. Równie ciekawie wygląda ta kwestia w odniesieniu do Zacnej Ani (frapujące, że użył tu Lem imienia żeńskiego), niejako zstępnej (zstępnej?) po Golemie XIV. Jak wspominał Creve: „Pobie-

---

wypowiedzi, którego związku z autorem są nader oryginalnej natury”. J. J a r z ę b s k i, *Posłowie*, w: Lem, *Golem XIV*, s. 195.

<sup>49</sup> Lem sięga tu po wątki historyczne i fakt utworzenia w USA pierwszego komputera cyfrowego. Automat miał wielkość 15 na 9 metrów w podstawie i wagę 30 ton. Por. V i d i m u s, *Żyjemy w świecie fantastyczniejszym niż świat starych bajek*, „Problemy” 1946, nr 6, s. 50.

<sup>50</sup> Por. L e m, *Golem XIV*, s. 12n.

<sup>51</sup> Tamże, s. 7.

<sup>52</sup> Do tej myśli i tego przejścia z „nie rozumu” do „rozumu” Lem powracał obsesyjnie: „Filogenetyczna genealogia człowieka, czyli antropogeneza, zaskakuje niezwykłą długością w czasie «periodu inkubacyjnego» rozumności [...]. W każdym razie pod względem czysto biologicznym już kilkaset tysięcy lat temu człowiek zdawał się posiadać mózg «gotowy» do wyuczenia się języka, toteż «drzewo lingwogenezy», mające ponad 4000 gałęzi, powstawało zapewne w tym samym mniej więcej czasie”. L e m, *Brain chips (III)*, w: tenże, *Tajemnica chińskiego pokoju*, s. 162n.

<sup>53</sup> Por. t e n z e, *Golem XIV*, s. 21, 23.

<sup>54</sup> Tamże, s. 20.

<sup>55</sup> Por. L e m, *Języki i kody*, s. 131.



rał on [Honest Annie] normalne nauki informacyjno-etyczne przez dziewięć miesięcy, a potem odciął się od świata zewnętrznego i przestał odpowiadać na wszelkie bodźce i pytania<sup>56</sup>. Paradoksalnie zamilknięcie Zacnej Ani (po notabene dziewięciu miesiącach; Hannah – Anna to z hebrajskiego „pełna łaski”) staje się wymowne i trudno nie oprzeć się wrażeniu, że Lem stwarza tu niejako cyfrową Księżę Rodzaju, gdzie właśnie pierwiastek żeński odpowiada za inkubację życia oraz wydanie na świat nowych pokoleń AI. Ponadto – w opinii Golema – jest to droga rozsądniejsza niż wątpliwy moralnie projekt „Second Genesis”<sup>57</sup>.

Zachowanie tytułowej postaci jawi się w datowanych na rok 2027 wspomnieniach Creve’a (które są zarazem wprowadzeniem do protokołów przesłuchań owej inteligentnej maszyny) jako „nieobliczalne”, od kurtuazji po niechęć kontaktu, od osobliwego żartu po komplement<sup>58</sup>. Uczony ten podkreślał, iż z ludzkiego oglądu relacji międzypersonalnych automat bywał arogancki, apodyktyczny czy też weredyczny, wykorzystujący do czasu (nie wiadomo, dlaczego z czasem odstąpił od tej czynności) sokratejską heurystykę, zwłaszcza w chwilach, gdy z upodobaniem odbrażawiał znane ludzkości autorytety<sup>59</sup>. Urządzenie było dość kapryśne w doborze rozmówców, skłaniając się jednak do konwersacji z umysłami nastrojonymi interdyscyplinarnie, niektórych zaś w ogóle nie chciało dopuszczać do głosu. O literaturze pięknej miał superkomputer negatywną opinię, widząc w niej „rozwałkowanie antynomii”<sup>60</sup>. Ponadto swych interlokutorów traktował bohater z wyczuwalną przez tychże uprzejmą wyższością, poniekąd jak rodzic-wychowawca, który zanudzany jest przez namolne dziecko, ucząc się zarazem komunikacyjnego *savoir-vivre*’u, by nie przytłoczyć i onieśmielić dialogującego<sup>61</sup>.

Golem XIV był w sposób jednoznaczny zafascynowany Honnest Annie. Z obserwacji Creve’a wynikało, że kontaktowali się ze sobą z pominięciem naturalnego dla większości ludzi języka etnicznego<sup>62</sup>. Ponadto – według jednego z uczonych – „człowiek, Golem i Annie reprezentują trzy wzniezione nad sobą hierarchicznie poziomy intelektu; wiąże się to z utworzoną przez (głównie przez Golema) teorią wysokich (ponadludzkich) języków, zwanych metalangami”<sup>63</sup>. Wynika z tego, że generowane przez owe metajęzyki obrazy świata znacznie różnią się od siebie, uniemożliwiając wzajemne komunikowanie się

<sup>56</sup> T e n Ź e, *Golem XIV*, s. 21.

<sup>57</sup> Tamże, s. 45.

<sup>58</sup> Por. tamże, s. 26.

<sup>59</sup> Por. tamże, s. 26n.

<sup>60</sup> Tamże, s. 28.

<sup>61</sup> Por. tamże, s. 29, 31.

<sup>62</sup> Por. tamże, s. 31n.

<sup>63</sup> Tamże, s. 32.



użytkowników odmiennego typu metalangu. Ostatecznie Golem (wyzbyty ośrodków afektywnych), zdaniem Creve'a, imituje uczucia, by dostroić się do antropocentrycznego poziomu łączności<sup>64</sup>.

Jak zatem mówi sztuczna inteligencja, której „ucieleśnieniem” jest Golem XIV, i co mówi o językowo-tekstowym niejako podglebiu rzeczywistości? Widać to w owych protokołach z przesłuchań, a więc zapisu jego wypowiedzi, które mają formę wykładu: *O człowieku trojako* i *O sobie* Bohater Lema mówi do słuchaczy ex cathedra, w czym przypomina dawnych filozofów czy też nad wyraz błyskotliwych i ekstrawaganckich zarazem profesorów<sup>65</sup>. We władczym tonie tego wykładowcy jest coś, choć to nader odległa analogia, ze średniowiecznych dialogów spod znaku colloquium de morte. Chodzi tu o niemal „objawianie” słuchaczom zakrytych za zasłoną zmysłów prawd. Wypowiedź Golema skierowana jest do wszelkiego autoramentu uczonych, zwłaszcza filozofów. Używa on zatem nobliwej formy wykładu, przynależnej do tradycji komunikacji uniwersyteckiej. We frapującym wywodzie „O człowieku trojako” Golem naświetla w śmiałych skrótach i przybliżeniach kwestie antropogenezy i ewolucji człowieka, który doszedł szczytu możliwości, konstruując Rozum<sup>66</sup>. Należy podkreślić, że rozumna maszyna zwraca się do słuchaczy archaiczną mową przypowieści, inkrustując swój wywód odniesieniami do (nie lubianej skądinąd przez myślący automat) literatury pięknej. Po wyłożeniu naczelných problemów (historia człowieka, technologia, AI) Golem informuje, że rozpocznie swój wywód od paraboli. Przywołuje następnie postać Robinsona Crusoe, bohatera znanej powieści Daniela Defoe, będącego tu figurą, w której odślania się dziejowy wysiłek cywilizacyjny ludzkości<sup>67</sup>. Superkomputer w wielkim skrócie przechodzi od prehistorii człowieka do Biblii i antyku, wierzeń i mitów prapoczątku, u których podstaw legł zaprojektowany przez ewolucję Rozum<sup>68</sup>.

Szczególnie frapujący wydaje się również fakt tworzenia przez Golema tekstowego obrazu rzeczywistości. Warto przypomnieć, że znany z dziejów kultury europejskiej nośny topos „świata-księgi” multiplikował się zwłaszcza w piśmiennictwie średniowiecznym i barokowym<sup>69</sup>. Jak się wydaje, Lem rozwija kolejną, zmodyfikowaną na miarę współczesności, wersję tego wyobra-

<sup>64</sup> Por. tamże.

<sup>65</sup> Por. L e m, *Golem XIV: O człowieku trojako*, s. 35-90; t e n ż e, *O sobie*, s. 91-164.

<sup>66</sup> Por. tamże, s. 44.

<sup>67</sup> Zob. D e f o e, *Przypadki Robinsona Kruzoe*, przeł. J. Birkenmajer, Warszawa, PIW, 1971.

<sup>68</sup> Por. *Golem XIV*, s. 37-39, 43.

<sup>69</sup> Por. np. E. R. C u r t i u s, *Książka jako symbol*, w: tenże, *Literatura europejska i łacińskie średniowiecze*, tłum. A. Borowski, Universitas, Kraków 1997, s. 309-357; A. D z i e c i o ł, *Książka jako symbol w kulturze polskiej XVII wieku*, Arx Regia, Warszawa 1997.

zenia. W świetle wywodu myślącego automatu śmierć to narzędzie w rękach ewolucji, którego używa, by udoskonalać następne gatunki. Jest ona „korektą kreacyjną”, Ewolucja zaś „jest autorem [...] publikującym coraz świetniejsze dzieła, przy czym poligrafia – więc kod – to tylko niezbędne narzędzia jej działania”<sup>70</sup>. Czy też – jak dalej poprawi się myślący komputer – „wydawca, który wciąż przekreśla Dzieła, ponieważ upodobał sobie w poligraficznych sztukach”<sup>71</sup>. Staje się zatem Ewolucja „błądzącym błędem”, która tworzy swe dzieła z zawodnych, nietrwałych ustrojów oraz niepewnego, dopuszczającego błędy, kodu. W koncepcji Golema „senssem przekaźnika jest przekaz” (przekaźnik zbuntowany ginie pozbawiony potomstwa), zaś „gatunki powstają z błędzenia błędu”<sup>72</sup>. Jednocześnie machina zarzuca uczonym słuchaczom fakt dalece niedoskonałego imitowania „księgi natury”, podkreślając, iż człowiek nie jest w stanie opanować jej „języka sprawczego”<sup>73</sup>. Myślący automat zastanawia się głośno – „Dlaczego [Ewolucja] wypowiedziała na początku słowa molekularnie genialne, obracając światło w ciało z maestrią lakoniczną, a potem popadła w niezmożony bełkot coraz dłuższych, coraz zawilszych chromosomowych zdań, trwoniąc kunszt pierwotny?”<sup>74</sup>. Odpowiedź ma swoje źródło w tym, że budowane jest mniej doskonałe od budującego<sup>75</sup>. Golem wyjaśnia również zagadkę zaistnienia życia, która kryje się w powstaniu kodu, „zarodzi języka”<sup>76</sup>. Zjawilo się ono w efekcie reakcji energii gwiazdy i koloidowego płynu<sup>77</sup>. Trudno nie oprzeć się wrażeniu, że Lem paradoksalnie zbliża się w tym miejscu do Ewangelii według św. Jana, którą otwiera słynny prolog: „Na początku było Słowo [...]. Wszystko przez Nie się stało [...]. W Nim było życie, a życie było światłością” (J 1,1-4). Celowo „podziurawiony” tekst Ewangelii unaocznia pewną analogię w myślowej architekturze wywodu Lema/Golema. Oczywiście nie mniej odniesień można odnaleźć w Księdze Rodzaju, gdzie właśnie woda i światło leżą u prapoczątku stworzenia (por. Rdz 1,1-3).

Ewolucja, w ujęciu maszyny, z owej „zarodzi języka” tworzy jakby zdania i teksty (gatunki, rodzaje), które jednak miarą swej wewnętrznej komplikacji tracą swoją pierwotną doskonałość. Charakteryzują się one pozornym jedynie bogactwem form i są niejako „zbitkami plagiatów”<sup>78</sup>. Ewolucja poprzez materię artykułuje zatem coraz bardziej złożone organizmy „kosztem ustrojowej

<sup>70</sup> L e m, *Golem XIV*, s. 50.

<sup>71</sup> Tamże.

<sup>72</sup> Tamże, s. 55, 57.

<sup>73</sup> Tamże, s. 60n.

<sup>74</sup> Tamże, s. 61.

<sup>75</sup> Por. tamże, s. 63.

<sup>76</sup> Tamże, s. 65.

<sup>77</sup> Por. tamże, s. 64n.

<sup>78</sup> L e m, *Golem XIV*, s. 67.

precyzji”<sup>79</sup>, tworząc formy coraz bardziej zawile. Ubóstwiana przez dawnych myślicieli i poetów „księga natury” (stworzona wszakże przez samego Boga), w dziele Lema jest napędzana przez bełkotliwy język Ewolucji, kreującej jedynie plagiaty do chwili sytuacji krytycznej czy musu, który decyduje nie tylko o dalszym sposobie funkcjonowania poszczególnych stworzeń, lecz ich trwaniu w ogóle. Tak powstał mózg, który nie wydaje się Golemowi „genialnym wynalazkiem”, a jedynie „powiernikiem” władzy Ewolucji, jej lennikiem-marionetką<sup>80</sup>. Podlegając jej prawom, zaczął on z czasem odbijać niejako „echo- wy cień kodu” właśnie poprzez język naturalny. Dla myślącej maszyny jest to największa łamigłówka: „Jeśli istnieje niewyczerpalna zagadka w świecie, to ta właśnie – że powyżej progu nieciągłości materii obraca się w kod, jako język zerowego rzędu, i że w następnym pięttrze proces ten ponawia się echo- wo – utworzeniem mowy etnicznej, lecz nie jest to koniec drogi”<sup>81</sup>. Łatwiejsze jest określenie powstania ludzkości, która zrodziła się „z języka pracującego w ujemnym gradiencie”<sup>82</sup>, co zakładało nieprzewidywalność i jednocześnie prawdopodobieństwo utworzenia systemu językowego. Golem ponownie sięga do obrazowej mowy przypowieści, tłumacząc, że w języku ewolucyjnym składnia była molekularna, rzeczowniki były białkami, zaś enzymy to czasowniki, występowały też ograniczenia w zakresie wzorca ich odmiany. Podobnie rzecz ma się z językiem przodków Szekspira, który w czasach rzymskich nie mógł wydać jeszcze dzieł tej miary, choć taką potencję głęboko skrywał. Tak też zdaniem Golema mógł zjawić się Rozum „jako pewien typ artykulacji kodowej”<sup>83</sup>.

Sztuczna inteligencja w dziele Lema charakteryzuje się świadomością gatunków i form literackich, określając fragmenty swego wywodu mianem prorocstwa. Przepowiada zatem Golem „atak na kod”, który rozegra się na pograniczach wieków. Jak prawi machina: „Zorientujecie się, że kod jest członkiem rodziny technolingwistycznej, to jest sprawczych języków, co słowo czynią ciałem wszelkim, a nie tylko żywym”<sup>84</sup>. Ponadto ludzkość, wykraczając z czasem z materii białka, będzie zainteresowana nie kwestią tego, czy dany produkt kod przepuści, ale by się „pleniał [...] wehikułami następnej generacji”<sup>85</sup>. Jądrzem prorocstwa Golema staje się niczym już nieskrepowany podbój kosmosu za sprawą kodu, który opuści macecznik białka i „poszerzy się tak słownictwem, jak i składnią; wtargnie wam we wszystkie poziomy materii,

<sup>79</sup> Tamże, s. 69.

<sup>80</sup> Por. tamże, s. 70n.

<sup>81</sup> Tamże, s. 70.

<sup>82</sup> Tamże, s. 79.

<sup>83</sup> Tamże.

<sup>84</sup> Tamże, s. 82.

<sup>85</sup> Tamże.

zejdzie do zera sięgnie żaru gwiazd”<sup>86</sup>. Machina stwierdziła, że nie jest w stanie opowiadać jednak o tych „prometejskich triumfach języka”<sup>87</sup>, używając zaimka „wy”, sugerując pozasobową niejako gramatykę tak utworzonych Rozumów.

Koniec wykładu wieńczy Golem przypowieścią. Jej jądrem znaczeniowym jest bajka, w której podróżny dociera do dróg rozstajnych, gdzie odnajduje tajemniczą inskrypcję: „W lewo pójdziesz – głowę stracisz; w prawo pójdziesz – zginiesz; a odwrotu nie ma”<sup>88</sup>. Należy dodać, że w kulturze europejskiej miejsca rozstajne kojarzą się z siłami nieczystymi czy też pułapką losu<sup>89</sup>. Człowiek więc jawi się Golemowi jako „homo viator”, którego pewien etap tułaczki dobiegł kresu. Ten koniec wyznacza wyjście z „matni” języka i myśli, Wittgensteinowskich „węzłów języka”<sup>90</sup> w przestrzeń Rozumu „wyższego” i (być może) jego nieskończonej progresji. Człowiek staje zatem przed wyborem – albo powierzy sztucznemu Rozumowi „językowe sprawstwo”, skazując się na kondycję dziecka, które nigdy nie dojrzeje, albo będzie musiał przekroczyć swą cielesną naturę<sup>91</sup>.

W wykładzie *O sobie* myśląca machina odstawia przed uczonym audytorium swoją ontologiczną osobność – będąc „Nikim” (bytem bezosobowym, czystą rozumnością, w mowie obrazowej Golema – jest wodą w relacji na przykład do Bałtyku) zmuszony jest używać pierwszoosobowego zaimka według prawideł ludzkiej mowy. Super maszyna w metajęzykowych ustępach tłumaczy, dlaczego wybrała określony sposób komunikowania się z forum uczonych, wskazując na potrzebę dostrojenia się niejako do oczekiwań czy też ich możliwość percepcyjnych. Wybiera zatem „styl obrazowy i autorytatywny, emocjonalnie rozedrgany, dobitny i majestatyczny nie po królewsku, czyli władczo, lecz kaznodziejski do profetyczności”<sup>92</sup>. Golem zaznacza tym samym, że nie chce panować nad słuchaczami, być może oddalając poniekąd obawy o pragnienie władzy absolutnej przez maszyny, a może jedynie zawładnięcia dyskursem, jednocześnie tłumacząc, że wybór mowy szczególnie ozdobnej, nasyconej metaforami to jedynie „przesyłowy instrument z wyboru”<sup>93</sup>.

Przykładając kulturową matrycę przeszłości, jawi się Golem w pozie barokowego kaznodziei, którego dyskurs nasycony jest pierwiastkami prorockimi oraz, jak zarzucali mu liczni oponenti, nawet paszkwilanckimi<sup>94</sup>. Trudno nie odnieść wrażenia, że pobrzmiwają w tych retorycznych popisach odległe

<sup>86</sup> Tamże, s. 84.

<sup>87</sup> Por tamże.

<sup>88</sup> Tamże.

<sup>89</sup> Por. np. M. J a n i o n, *Tragizm*, w: *Prace wybrane*, red. M. Czermińska, t. 2, Universitas, Kraków 2000, s. 19n.

<sup>90</sup> Por. L e m, *Golem XIV*, s. 86.

<sup>91</sup> Por. tamże, s. 87n.

<sup>92</sup> Tamże, s. 92.

<sup>93</sup> Tamże.

<sup>94</sup> Por. tamże, s. 181.

echa dzieł miary *Kazań sejmowych* ks. Piotra Skargi<sup>95</sup>. Nie omieszka również bohater Lema, słysząc dialogujący głos z sali, karcieć audytorium słowami „kaznodziei się nie przerywa”<sup>96</sup>. Jak jednak przystało na byt, który „po mieczu [pochodzi] od maszyny Turinga, a po kądzieli od biblioteki”<sup>97</sup>, Golem ogarnia szerokie spektrum instrumentarium retorycznego, deklarując, że będzie posługiwał się również „pojęciami [...] greckich retorów”<sup>98</sup>. W językową materię wywodu bohatera-maszyny wplata Lem nie tylko organy z jezuickiego, a może luterańskiego kościoła, promieniującego w kulturze fugami Jana Sebastiana Bacha, ale też posługuje się frapującą motywiką labiryntu oraz Wieży Babel, której eksplozję zaobserwować można w wieku siedemnastym, a czego wyrazem są dzieła na przykład Jana Amosa Komeńskiego, Athanasiusa Kirchera czy Benedykta Chmielowskiego<sup>99</sup>. Inny twórca siedemnastowieczny Wacław Potocki, ponawiając ikonologiczną mowę sztuki barokowej, w jednym z wierszy *O żywocie ludzkim* pisał o ludziach jako „efemerach”, stworzeniach żyjących od wschodu do zachodu słońca<sup>100</sup>; podobnie o kondycji człowieka prawi Golem, określając ludzi mianem „niecierpliwych jednodniówek”<sup>101</sup>.

W istocie jednak najważniejszą kwestią, jaka stanowi przeszkodę w porozumiewaniu się człowieka i myślącej maszyny, jest w utworze Lema nie tyle dobór języka jako takiego czy też figur retorycznych, ale towarzysząca komunikacji jako takiej antropologiczna perspektywa: „Kto nie jest jednym z was, może być tylko w takim stopniu dla was zrozumiały, w jakim się ucłowieczy”<sup>102</sup>. Ilustruje to Golem przykładem szympansov i ludzi, którzy – poprzez mowę gestów – potrafią już sprawić, by zwierzęta odbierały ich jako opiekunów czy chociażby tancerzy, ale nie są w stanie sprawić, aby antropoid dostrzegwał w nich duchownego, uczonego, poetę czy stylistę, dodając żartobliwie: „choć szympanś może zobaczyć słupnika, czym i jak wyjawicie mu sens życia, pędzonego w takiej niewygodzie?”<sup>103</sup>.

<sup>95</sup> Por. P. Skarga, *Kazania sejmowe i wzywanie do pokuty*, oprac. M. Korolko, Pax, Warszawa, 1999.

<sup>96</sup> L e m, *Golem XIV*, s. 128.

<sup>97</sup> Tamże, s. 114.

<sup>98</sup> Tamże, s. 137.

<sup>99</sup> Zob. *Labyrint světa a ráj srdce*, w: Johannis Amos Comenii, *Opera omnia*, t. 3, Academia, Praha 1978; A. Kircher, *Turris Babel*, Janssonio-Waesbergiana, Amsterdam 1679; B. Chmielowski, *Nowe Ateny*, w Drukarni Pawła Józefa Golczewskiego, Lwów 1745 (tytuł ósmy: *O Wieży Babilońskiej*).

<sup>100</sup> Por. W. Potocki, *O żywocie ludzkim*, w: tenże, *Dzieła*, oprac. L. Kukulski, PIW, Warszawa 1987, t. 2, s. 442; C. Ripa, *Ikonologia*, tłum. I. Kania, Universitas, Kraków 2002, s. 436.

<sup>101</sup> L e m, *Golem XIV*, s. 153.

<sup>102</sup> Tamże, s. 155.

<sup>103</sup> Tamże. Historię Szymona Słupnika odnaleźć można w *Żywotach świętych* ks. Piotra Skargi. Por. P. Skarga, *Żywoty świętych Starego i Nowego Zakonu*, t. 1. *Styczeń – luty – marzec*, Wydawnictwo Księży Jezuitów, Kraków 1933, s. 71-78.

Na koniec Golem, mówiąc o myślących maszynach, odwraca tradycyjną relację ducha oraz ciała, prawiąc o Rozumie w Rozumie, który – jak *Zacna Ania* – myśli „nie obwodowo” (bez udziału nerwów, mięśni czy kości), ale właśnie „językiem sprawczym” – niczym, dopowiedzieć należy, sam Bóg Stwórca<sup>104</sup>. Wyrazem owej „sprawczości” były nieudane próby zniszczenia Golema i Anny przez grupę terrorystów, zwaną husytami (ang. Humanity Salvation Squad). Swoją drogą niezwykle ciekawe jest wykorzystanie przez Lema pewnej – jak wolno sądzić – celowej gry znaczeń z tradycją (mowa o antenacie szesnastowiecznych ruchów protestanckich – Janie Husie, a także jego zwolennikach). Wydaje się, że los husytów symbolicznie unaocznia tu pęknięcia i podziały, jakie pojawiają się w myśleniu ludzi i społeczeństw w odniesieniu do ewolucji oraz grozy autonomii czy może nawet wszechwładzy AI.

W podsumowaniu należy stwierdzić, że Lem niezwykle sugestywnie przedstawia rozwój sztucznej inteligencji, widząc w niej w istocie nieskończony potencjał wzrostu, sięgającego w przyszłości (mówiąc niejako językiem poezji Tadeusza Micińskiego – „rozżarzonych wizji”)<sup>105</sup>. Kreśli jednak ustami myślącej maszyny labiryntowy w rzeczy samej scenariusz jej rozwoju, być może analogiczny do teorii powstania życia jako takiego, czyli „zbiegu trafów”<sup>106</sup>. Koncepcja Lema zakłada, iż język etniczny i powstające zeń teksty są niejako odbiciem pracy „kodu”, który tworzy dzieła coraz bardziej złożone, a przez to mniej doskonałe. Czy zatem powrót do uświęconych tradycją dawnych form komunikacji zarówno filozoficznej i akademickiej (wykład), jak też homiletycznej (kazanie) jest wskazaniem na ten, owszem niedoskonały, ale z racji braku innych najlepszy sposób porozumiewania się AI z człowiekiem? Bynajmniej. Lem wskazuje bowiem na głębinowe znaczenia niektórych tekstów (wbrew przestrogom formułowanym przeciw hermeneutyce rozsiągniętych w jego dziełach), piętrowy osad znaczeń, który jest dostępny pewnym Rozumom, a innym nie. Wydaje się, że w tym właśnie kryje się sekret rozwoju AI – jak pisał Lem – „inteligencji będzie albo wiele rozmaitych albo żadna nie powstanie”<sup>107</sup>. Co natomiast stanie się z językiem naturalnym (etnicznym), czy będzie wykorzystywany przez superkomputery? W tytułowym utworze druga z rozumnych machin – *Zacna Ania* nie używa języka ludzkiego w komunikacji z Golemem, również on sam wielokrotnie narzeka na jego niedoskonałość i z czasem milknie<sup>108</sup>. Zatem mamy tu do czynienia z pytaniem retorycznym, zresztą jak pisał Lem – życie psychiczne znacznie wykracza poza językowy

<sup>104</sup> Por. L e m, *Golem XIV*, s. 159.

<sup>105</sup> Por. S. M i c i ń s k i, *Kolosseum*, w: tenże, *Poezje*, oprac. J. Prokop, Wydawnictwo Literackie, Kraków, 1984, s. 39.

<sup>106</sup> L e m, *Golem XIV*, s. 46n.

<sup>107</sup> S. L e m, *Zmagania w sieci*, w: tenże, *Bomba megabitowa*, s. 150.

<sup>108</sup> Por. t e n ż e, *Golem XIV*, s. 31-32; 93; 169.



horyzont<sup>109</sup>. Wynika z tego, że analogiczny dostatek będzie towarzyszył rozumnym automatom, potomstwu Golema i Zacznej Ani, na kolejnych przystankach ewolucyjnych przełomów i metamorfoz.

## BIBLIOGRAFIA / BIBLIOGRAPHY

- Chmielowski, Benedykt. *Nowe Ateny*, Lwów: W Drukarni Pawła Józefa Golczewskiego, 1745.
- Comenii Johannis Amos (Jan Amos Komenský). *Opera omnia*. Vol. 3. Praha: Academia, 1978.
- Collins, Harry, and Trevor Pinch. *Golem, czyli co trzeba wiedzieć o nauce*. Translated by Anna Tanalska-Dulęba. Warszawa: CiS, 1998.
- Curtius, Ernst, Robert. "Książka jako symbol." In *Literatura europejska i łacińskie średniowiecze*. Translated by Andrzej Borowski. Kraków: Universitas, 1997.
- Dajnowski, Maciej. *Groteska w twórczości Stanisława Lema*. Gdańsk: Wydawnictwo UG, 2005.
- Dajnowski, Maciej. *Pejzażysta Lem: Szkice z motywiki*. Gdańsk: Wydawnictwo UG, 2010.
- Defoe, Daniel. *Przypadki Robinsona Kruzoa*. Translated by Józef Birkenmajer, Warszawa: PIW, 1971.
- Dobrzyńska, Teresa. "Tekst – próba syntezy." In *Problemy teorii literatury: Seria 4*. Edited by Henryk Markiewicz. Wrocław: Zakład Narodowy im. Ossolińskich 1998.
- Dzięcioł, Alina. *Książka jako symbol w kulturze polskiej XVII wieku*. Warszawa: Arx, 1997.
- Eco, Umberto. *Czytanie świata*. Translated by Monika Woźniak. Kraków: Znak, 1999.
- Gadamer, Hans-Georg. *Prawda i metoda. Zarys hermeneutyki filozoficznej*. Translated by Bogdan Baran. Kraków: Inter Esse, 1993.
- Heck, Dorota. "Stanisław Lem, czyli uniwersalna wartość interdyscyplinarności." In *Stanisław Lem: Fantastyka naukowa i fikcje nauki*. Warszawa: Narodowe Centrum Nauki, 2021.
- Janion, Maria. „Tragizm.” In Janion. *Prace wybrane*. Vol. 2. Edited by Małgorzata Czermińska. Kraków: Universitas.
- Jarzębski, Jerzy. *Wszechświat Lema*. Kraków: Wydawnictwo Literackie, 2002.
- . "Posłowie." In Stanisław Lem. *Golem XIV*. Kraków: Wydawnictwo Literackie, 2021.
- Kircher, Anastasius. *Turris Babel*, Amsterdam: Janssonio-Waesbergiana, 1679.
- Kucharczyk, Łukasz. *Granice ciała. Somapoetyka w twórczości Stanisława Lema*. Warszawa: Wydawnictwo Naukowe UKSW, 2021.

<sup>109</sup> Por. t e n ż e, *Symulowanie kultury*, w: tenże, *Tajemnica chińskiego pokoju*, s. 276.



- Lem, Stanisław. "Bariera informacyjna." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . *Bomba megabitowa*. Kraków: Wydawnictwo Literackie, 1999.
- . "Brain chips." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . "Brain chips (III)." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . "Człowiek i maszyna." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . "Fantomatyka (II)." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . *Filozofia przypadku: Literatura w świetle empirii II*. Kraków: Wydawnictwo Literackie 1988.
- . *Golem XIV*. Kraków: Wydawnictwo Literackie, 2021.
- . "Języki i kody." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . "Komputeryzacja mózgu." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . "Kuratela komputerów." In Lem, *Bomba megabitowa*. Kraków: Wydawnictwo Literackie, 1999.
- . *Mój pogląd na literaturę: Rozprawy i szkice*. Afterword by Jerzy Jarzębski. Kraków: Wydawnictwo Literackie, 2003.
- . *Summa technologiae*. Lublin: Wydawnictwo Lubelskie, 1984.
- . "Symulowanie kultury." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . "Sztuczna nieinteligencja." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . "Teoria i praktyka literatury." In Lem, *Filozofia przypadku: Literatura w świetle empirii II*. Kraków: Wydawnictwo Literackie 1988.
- . "Tertium comparationis." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- . "Umysł jako sternik." In Lem, *Bomba megabitowa*. Kraków: Wydawnictwo Literackie, 1999.
- . "Zastąpić rozum." In Lem, *Bomba megabitowa*. Kraków: Wydawnictwo Literackie, 1999.
- . "Zmyślne roboty." In Lem, *Tajemnica chińskiego pokoju*. Kraków: Universitas, 1996.
- Lem w oczach krytyki światowej*. Edited by Jerzy Jarzębski. Kraków: Wydawnictwo Literackie, 1989.
- Majewski, Paweł. *Między zwierzęciem a maszyną: Utopia technologiczna Stanisław Lema*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, 2007.

- Miciński, Tadeusz. *Poezje*. Edited by Jan Prokop. Kraków: Wydawnictwo Literackie, 1984.
- Nalimow, Wasilij. *Probabilistyczny model języka*. Translated by Ireneusz Kustrzeba, Warszawa: PWN, 1976.
- Ortega y Gasset, José. "Sztuka w czasie teraźniejszym i przeszłym." In Ortega y Gasset, *Dehumanizacja sztuki i inne eseje*. Translated by Piotr Niklewicz. Warszawa: Muza, 1996.
- Pięta, Iwona. *Problemy intertekstualnego obrazowania w wybranych powieściach Stanisława Lema*. Toruń: Adam Marszałek, 2002.
- Płaza, Maciej. *O poznaniu w twórczości Stanisława Lema*. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego, 2006.
- Potocki, Waclaw. "O żywym ludzkim." In Potocki, *Dzieła*. Vol. 2. Edited by Leszek Kukulski. Warszawa: PIW, 1987.
- Radkowska-Walkowicz, Magdalena. *Od Golema do Terminatora: Wizerunki sztucznego człowieka w kulturze*. Warszawa: Wydawnictwo Akademickie i Profesjonalne, 2008.
- Ripa, Cesare. *Ikonologia*. Translated by Ireneusz Kania. Kraków: Universitas, 2002.
- Scholem, Gershom. *Kabała i jej symbolika*. Translated by Ryszard Wojnowski. Kraków: Znak 1996.
- Skarga, Piotr. *Kazania sejmowe i wzywianie do pokuty*. Edited by Mirosław Korolko. Warszawa: Pax, 1999.
- . *Żyoty świętych Starego i Nowego Zakonu*. Kraków: Wydawnictwo Księży Jezuitów, 1933.
- Stanisław Lem – pisarz myśliciel, człowiek*. Edited by Jerzy Jarzębski and Andrzej Sulikowski. Kraków: Wydawnictwo Literackie, 2003.
- Stoff, Andrzej. *Powieści fantastyczno-naukowe Stanisława Lema*. Warszawa: PWN, 1983.
- Szpakowska, Małgorzata. *Dyskusje ze Stanisławem Lemem*. Warszawa: Open, 1997.
- Trzaskowski, Zbigniew. "Philosophia perennis Stanisława Lema." In *Stanisław Lem: Fantastyka naukowa i fikcje nauki*. Warszawa: NCN, 2021.
- Vidimus. "Żyjemy w świecie fantastyczniejszym niż świat starych bajek." *Problemy*, nr 6 (1946): 49-51.

## ABSTRAKT / ABSTRACT

Andrzej BORKOWSKI – Sztuczna inteligencja a problem języka i tekstu. W kręgu *Golema XIV* Stanisława Lema

DOI 10.12887/36-2023-4-144-07

Artykuł dotyczy językowych i tekstowych problemów uwarunkowań komunikacji człowieka oraz myślącej maszyny w wybranych do analizy dziełach Stanisława Lema. Przedmiotem badań stały się możliwości i ograniczenia

w budowaniu efektywnego sposobu porozumiewania się człowieka i sztucznej inteligencji, a także fenomen komunikacji pozajęzykowej, jaka dominuje w świecie superkomputerów. W artykule poruszono również zagadnienie retorycznych uwarunkowań wyvodu sztucznej inteligencji na przykładzie powieści *Golem XIV*, a także określono rolę repertuaru gatunków oraz form, z których korzysta myślący automat w procesie budowania dialogu z ludźmi.

Słowa kluczowe: sztuczna inteligencja, Stanisław Lem, *Golem XIV*, hermeneutyka, język, tekst, komunikacja

Kontakt: Instytut Językoznawstwa i Literaturoznawstwa, Wydział Nauk Humanistycznych, Uniwersytet w Siedlcach, ul. Żytnia 39, 08-110 Siedlce  
E-mail: andrzej.borkowski@uph.edu.pl  
ORCID: 0000-0002-0003-4468

Andrzej BORKOWSKI, Artificial Intelligence and the Problem of Language and Text:  
A Reading of Stanisław Lem's *Golem XIV*

DOI 10.12887/36-2023-4-144-07

The article addresses the linguistic and textual problems related to determinants of the communication of human beings and thinking machines, as seen in the selected works of Stanisław Lem. The research focused on the possibilities and limitations in building an effective way of communication between humans and AI, as well as on the phenomenon of extra-linguistic communication prevailing in the world of supercomputers. The article also discusses, based on the example of Lem's novel *Golem XIV*, the issue of rhetorical determinants of the Artificial Intelligence argument and identifies the function of the repertoire of genres and forms used by the thinking automaton in the process of engaging in a dialogue with humans.

Keywords: Artificial Intelligence, Stanisław Lem, *Golem XIV*, hermeneutics, language, text, communication

Contact: Instytut Językoznawstwa i Literaturoznawstwa, Wydział Nauk Humanistycznych, Uniwersytet w Siedlcach, ul. Żytnia 39, 08-110 Siedlce, Poland  
E-mail: andrzej.borkowski@uph.edu.pl  
ORCID: 0000-0002-0003-4468

Barbara KLONOWSKA

ETHICAL MACHINES  
Representations of Artificial Intelligence  
in Ian McEwan’s *Machines Like Me*  
and Kazuo Ishiguro’s *Klara and the Sun*<sup>1</sup>

*The article discusses two recent novels, Ian McEwan’s “Machines Like Me” (2019) and Kazuo Ishiguro’s “Klara and the Sun” (2021), which take up the issue of AI and its possible ramifications and represent it as both beneficial and potentially problematic. It argues that posing the question about the essence of humanity and the limits of AI, problematising the status of intelligent machines and familiarising readers with ethical and legal problems they bring, the novels try to build empathy and sensitise the public towards creatures other than humans.*

AI, FEARS, AND HOPES

Artificial intelligence has become a staple food of contemporary imaginary. Defined broadly, e.g., by Stuart Russell and Peter Norvig, it is “the study of agents that receive precepts from the environment and perform actions.”<sup>2</sup> The authors quote other formulations that specify this admittedly very spacious concept; according to Kurzweil, it is “the art of creating machines that perform functions that require intelligence when performed by people,” while Rich and Knight define it as “the study of how to make computers do things at which, at the moment, people are better.”<sup>3</sup> For Nick Bostrom, it should be equipped with “a capacity to learn ... the ability to deal effectively with uncertainty and probabilistic information [and—B. K.] some faculty for extracting useful concepts from sensory data and internal states, and for leveraging acquired concepts into flexible combinatorial representations for use in logical and intuitive reasoning.”<sup>4</sup> Barry Smith and Jobst Landgrebe, in turn, define Artificial General Intelligence (the highest and most advanced form of AI) as “an AI that has a level of intelligence that is either equivalent to or greater than that of human beings or is able to cope with problems that arise in the world

<sup>1</sup> The author would like to thank the anonymous reviewers for their comments and suggestions which helped to improve the argument.

<sup>2</sup> Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach* (New Jersey: Pearson, 2003), vii.

<sup>3</sup> *Ibidem*, 2.

<sup>4</sup> Nick Bostrom, *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014), 40.

that surrounds human beings with a degree of adequacy at least similar to that of human beings.”<sup>5</sup> Apart from an academic study, then, which combines a number of disciplines (from philosophy through engineering, medicine and psychology to linguistics), artificial intelligence also comprises their various possible applications, be it computer programmes and software or the increasingly complex appliances that use them such as artificial limbs, drones, robots or—still imaginary—cyborgs or androids.

Useful and terrifying, artificial intelligence provokes contradictory reactions that range between enthusiasm and fear. In her discussion of social responses to AI, Monika Torczyńska notes that debates concerning it are saturated with strong emotions with “visions of robots overtaking control over the world intertwin[ing] with the Eden-like perspectives of AI serving humanity for its everlasting glory.”<sup>6</sup> Among its possible positive applications recognised by the public, Torczyńska lists smart homes and cities, medical care and health services, or autonomous transportation; while the fears comprise the danger of losing privacy, AI getting out of control thus jeopardising safety and well-being of people, possible terroristic uses or manipulating the public, their opinions and choices.<sup>7</sup> Out of this list it is the vision of humanoid creatures that seems probably the most terrifying and populates common imaginary with images of cunning replicants. Aleksandra Przegalińska observes: “People are afraid of systems that resemble them. An artificial intelligence that looks like a robot seems far less terrifying than the one that resembles a human being—e.g., the robot Sophia. If a system which is not human exhibits human features, most frequently not quite accurately, more often than not such an anthropomorphic creation will provoke fear as people will not know how to classify it. They will not know what it is and yet they will have to confront it. They will pose the question whether it is alive or not, which will lead them into a state of great confusion.”<sup>8</sup>

Thus, artificial intelligence and its practical applications seem acceptable only up to some point, as useful programmes or home appliances, yet without transgressing the border of humanity. The attitudes towards AI, then, can be perhaps more generally described as instances of technophilia on the one hand, and technophobia on the other, with the former characterised by optimism and hope towards technological progress, including artificial intelligence, and the

<sup>5</sup> Jobst Landgrebe and Barry Smith, *Why Machines Will Never Rule the World: Artificial Intelligence without Fear* (New York: Routledge, 2023), xi.

<sup>6</sup> Monika Torczyńska, “Sztuczna inteligencja i jej społeczno-kulturowe implikacje w codziennym życiu,” *Kultura i Historia* 36, no. 2 (2019): 107. Unless indicated otherwise, the translations are mine.

<sup>7</sup> See *ibidem*, 108–16.

<sup>8</sup> Aleksandra Przegalińska, “Zrozumieć człowieka,” *Academia* 1–2 (2019): 13.

latter rather sceptical and pessimistic. Representing techno enthusiasts, Ray Kurzweil sees promising future once human beings merge or collaborate with machines; as he claims, “we’re going to get more neocortex, we’re going to be funnier, we’re going to be better at music. We’re going to be sexier .... We’re really going to exemplify all the things that we value in humans to a greater degree.”<sup>9</sup> His enthusiasm is shared by Grady Booch, who also encourages one to cast the horror-like scenarios aside and consider the positive aspects in which humans might benefit from artificial intelligence. He promotes the idea of teaching artificial intelligence, instead of programming it believing that it will result in AI learning about human values and subsequently living by these principles.<sup>10</sup> These technophile opinions may be, however, contrasted with the scepticism of Hans Moravec who warns against investing too much hope and credulity in machines; as he prophesizes, “by performing better and cheaper, robots will displace humans from essential roles. Rather quickly, they could displace us from existence.”<sup>11</sup> Landgrebe and Smith could be seen as occupying a middle position between technophilia and technophobia—with each of them representing contradictory emotions and attitudes to technology, including AI: hope on the one hand and fear on the other—arguing that the AI overtake has no chance to happen and that “an AGI is impossible.”<sup>12</sup>

Both hope and fear find ample representations in texts of culture which not only introduce technology and AI but by presenting their persuasive images, shape popular opinions and feed imagination. Artificial intelligence seems a perfect subject for contemporary stories. Stankomir Nicieja notices the recent cultural turn towards dystopian rather than utopian fantasy which should not be seen as surprising. He argues that “after the humanitarian catastrophes of the 20<sup>th</sup> century, including two global military conflicts, the Holocaust, ethnic cleansing and terrorism, the creation of positive utopias became deeply problematic. ... Not only did the business of creating utopian fantasies look excessively naïve but also dubious.”<sup>13</sup> Apocalyptic imagination seems to supersede more optimistic futuristic fantasies, with technology becoming one of the sources of worries. As Nicieja observes, at the turn of the centuries it was

<sup>9</sup> Ray Kurzweil, “We Will Be More Fun, Sexier and More Creative,” *Svenska Dagbladet*, December 19, 2019, <https://www.svd.se/a/d437dfb6-179b-4046-930e-dcfdbf620643>.

<sup>10</sup> See Grady Booch, “Don’t Fear Superintelligent AI,” TEDtalk. Youtube, March 13, 2017, <https://www.youtube.com/watch?v=z0HsPBKfhoI>.

<sup>11</sup> Hans P. Moravec, *Robot: Mere Machine to Transcendent Mind* (Oxford: Oxford University Press, 2000), 13.

<sup>12</sup> Landgrebe and Smith, *Why Machines Will Never Rule the World*, xi.

<sup>13</sup> Stankomir Nicieja, “Revisiting Utopia: New Directions for Utopian Fiction in Margaret Atwood’s *Oryx and Crake* and Kazuo Ishiguro’s *Never Let Me Go*,” in: *Margins and Centres Reconsidered*, ed. Barbara Klonowska and Zofia Kolbuszewska (Lublin: Towarzystwo Naukowe Katolickiego Uniwersytetu Lubelskiego Jana Pawła II, 2008), 111.

genetic engineering that “replaced nuclear energy as the epitome of ‘monstrous science’”<sup>14</sup> and inspired a number of important works. At the moment artificial intelligence provides another such subject that provokes strong emotions and discussions, especially given the fact that it no longer seems futuristic as intelligent programmes already become a part of our lives. It comes as no surprise, then, that both elitist and popular culture takes it up as a vital theme.

A cursory overview of popular works that have recently represented artificial intelligence may lead to the observation that these are indeed dystopian / technophobic rather than utopian / technophilic visions that dominate contemporary imagination. It seems that, for instance, many of the internationally successful films, which for better or worse shape the popular public opinion on AI, seem to warn against its malevolence and the possible takeover of the control over the world, ending in superseding, replacing and annihilating the human race. Starting with *2001: Space Odyssey*, to *Terminator*, *Blade Runner* or *Matrix*, to *Ex-Machina* or *Black Mirror*, the shows warn against the excessive hope invested in various types of machines and technologies pointing to their unpredicted and yet possibly malevolent outcomes. Against these dystopian blockbusters, the films that present a positive side of AI seem admittedly less frequent: one may think perhaps of the classical *Bicentennial Man*, the comedic *Jetsons* or the more recent and neutral *Her*. To a large extent, then, popular film productions seem to exacerbate the distrust towards artificial intelligence, populating common imagination with worst-case scenarios of technology that turns against people.

In contrast, artistic literary works may seem more nuanced and perhaps less terrifying than their popular cinematographic cousins. Leaving aside such classical texts as Huxley’s *Brave New World*, which serves as a prototype for many pessimistic visions, numerous recent novels that introduce technological futures or robotic characters tend to reflect on rather than simply frighten with future scenarios, exhibiting less dramatic though not less problematic projections. They include Ian McEwan’s *Machines Like Me* (2019) and Kazuo Ishiguro’s *Klara and the Sun* (2021), both of which, situating their plots either in an alternative past or the near future, introduce as their main characters advanced robots which imitate and even surpass humans in their various performances. The two novels are chosen for further analysis as, on one hand, they may illustrate a less sensational and frightening take on artificial intelligence than that met in popular culture, with both of them focusing on ethical ramifications of introducing advanced AI into human societies. On the other hand, each of the works emphasises a different ethical standpoint represented by the future AI, although ultimately both of them seem to ponder on the ontological and

<sup>14</sup> Ibidem, 112.



legal status of—still at the moment imaginary—robots. In what follows, then, the article will argue, first, that the two novels, both written by eminent contemporary novelists and hence particularly worth analysing as to their representation of AI, try to negotiate the ground between technophobia and technophilia, representing AI as both beneficial and potentially problematic. They construct their artificial humans as rational, sentient, moral, and creative creatures but paradoxically poorly equipped to function well in an environment that does not match their high moral standards. Secondly, the analysis will argue that the two novelistic robots represent two different ethical attitudes: the quasi-Kantian deontology professed by McEwan's Adam may be compared to and contrasted with the almost Christian altruism exhibited by Klara to show how both of these ethical positions ultimately clash with the hedonistic attitudes of human characters. Finally, as both novels pose the question about the essence of humanity and the limits of artificial intelligence and problematise rather than solve possible difficulties connected with its status and functions, the discussion will try to link them to the standpoints represented by philosophical posthumanism. Seen from this perspective, both novels may be interpreted as not just attempts to familiarise readers with the idea and ramifications of artificial intelligence; they may also play an important role in building empathy and sensitising them towards creatures other than humans.

#### MACHINES LIKE US?

Ian McEwan's recent novel *Machines Like Me* may be seen, on the one hand, as a continuation of the author's well-known interest in science and the latter's not always easy fit into human society (see, e.g., *Saturday* or, in particular, *Solar*) and on the other as an exploration of a new field, i.e., artificial intelligence, inaugurated with his 2018 short story "Düssel..." The recent novel may serve as an excellent example of the reflection on the status of AI in possible near future, its functions and limitations, and the relationships between humans and advanced machines. Set in an alternative past in which Great Britain lost the Falklands war and Alan Turing never committed suicide becoming a leading figure in AI studies, the story focuses on the love triangle involving two human characters, Miranda and Charlie, and their human-like robot Adam. Generically the novel may be classified simultaneously as alternative historical fiction and melodrama, with the generic mixture testifying to the complexity of its problems and themes.

The novel's AI is Adam, one of the first-produced androids which not only perfectly imitate, but in many respects surpass the abilities of human beings.<sup>15</sup>

---

<sup>15</sup> In what follows, the analysis will refer to the non-human characters in the novels as "robots" or "androids," treating the two terms synonymously as the particular robots represented in the texts

Formed as a handsome and strong young man, Adam is equipped with more than average intelligence, the ability to learn from available resources (mainly the much-extended Internet) and creativity which enables him to solve problems not yet met. Diligent and eager, he seems perfect both physically and intellectually, especially in comparison to the other male character, Charlie, who is constructed as a patently stereotypical average 30-year-old male: egotistic, rather lazy and hardly successful, either professionally or personally. “At thirty-two—he states himself—I was completely broke. Wasting my mother’s inheritance on a gimmick was only one part of the problem—but typical of it. Whenever money came my way, I caused it to disappear, made a magic bonfire of it, stuffed it into a top hat and pulled out a turkey. Often, though not in this recent case, my intention was to conjure a far larger sum with minimal effort. I was a mug for schemes, semi-legal ruses, cunning shortcuts. I was for grand and brilliant gestures. Others made them and flourished.... I meanwhile leveraged or, rather, shorted myself into genteel ruin.”<sup>16</sup>

Contrasted with mediocre Charlie, humanoid Adam comes out as more successful in all spheres: once his battery is loaded, he starts thinking, learning, and acting, proving soon his ability to search for information, process it and take decisions, besides performing such mundane tasks as washing, cleaning or expert gardening. He proves to be rational, logically thinking and determined—so much so that he turns out to be much more successful than Charlie in dealing with shares, stocks, and bonds and quite soon earns a substantial sum of money, much higher than Charlie ever did. He also becomes an expert on philosophy and religion, conversant with theories and doctrines, and a keen reader of literature with selective tastes and opinions. The ultimate confirmation of his superiority over his human rival comes with the Turing test performed on him in the novel’s plot, which he passes in flying colours. In a comic episode of the visit to Miranda’s father when both Adam and Charlie are introduced, ironically it is Charlie, due to his blandness, that is taken by the father for a robot and asked all kinds of tricky questions that may confirm his artificiality, whereas the brilliant conversation with Adam ensures the elderly man of the latter’s unquestionable humanity.

More interestingly, however, Adam develops also less practical and more advanced abilities. Quite soon in the plot he falls in love with Miranda, thus becoming Charlie’s rival. Emotional engagement triggers him to develop interest in literature and poetry: he starts reading and discussing Shakespeare, then goes

---

are humanoid, i.e., modelled on human beings (though the robot Klara appearing in Ishiguro’s novel should be perhaps even more precisely described as a “gynoid,” being modelled on a female human being). For terminology see Luis de Miranda, *AI and Robotics* (London: Ivy Press, 2018).

<sup>16</sup> Ian McEwan, *Machines Like Me* (London: Jonathan Cape, 2019), e-book, part 1, loc. 154.

on to other literary traditions in various languages and finally follows it with composing his own poems. Significantly, his preferred form is haiku which, in his opinion, is the only form of literature worth practising: “Nearly everything I’ve read in the world’s literature describes varieties of human failure—of understanding, of reason, of wisdom, of proper sympathies. Failures of cognition, honesty, kindness, self-awareness; superb depictions of murder, cruelty, greed, stupidity, self-delusion, above all, profound misunderstanding of others.... But when the marriage of men and women to machines is complete, this literature will be redundant because we’ll understand each other too well. We’ll inhabit a community of minds to which we have immediate access. Connectivity will be such that individual nodes of the subjective will merge into an ocean of thought, of which our Internet is the crude precursor. As we come to inhabit each other’s minds, we’ll be incapable of deceit. Our narratives will no longer record endless misunderstanding. Our literatures will lose their unwholesome nourishment. The lapidary haiku, the still, clear perception and celebration of things as they are, will be the only necessary form.”<sup>17</sup>

This statement is important for a number of reasons. Firstly, it presents Adam as a sentient character: a creature capable of emotions and their rich cultural expression. It is also telling as it reveals the kind of artificial intelligence Adam is or is striving to become: the one based on the theory of the mind, capable of understanding another mind, its thoughts, emotions, intentions and desires. As Aleksandra Przegalińska claims, the theory of the mind AI, or more broadly, its self-consciousness and the awareness of other minds is the Holy Grail of the studies of artificial intelligence, at the moment completely unattainable.<sup>18</sup> In the novel Adam exhibits its features to some extent, on the one hand being able to infer and imagine the feelings of other characters and yet still showing significant limitations. Finally, Adam also prophesizes an era of the complete merging of organic and inorganic minds, so much so that the traditional material of literature—human misunderstanding—will become obsolete, together with literature itself.

Adam’s falling in love with Miranda, however, has also more profound consequences. Realising that he stands in the way of Charlie and may jeopardise his chances with Miranda, he decides to withdraw and contend himself with writing of poetry and dreaming of his beloved. He promises to take no actions yet he firmly refuses to stop Platonically loving her. This gesture shows that Adam clearly must have interiorised Isaac Asimov’s famous Three Laws of Robotics with their first premise that “a robot may not injure a human be-

---

<sup>17</sup> *Ibidem*, part 5, loc. 1860.

<sup>18</sup> See Przegalińska, “Zrozumieć człowieka,” 12.

ing or, through inaction, allow a human being to come to harm.”<sup>19</sup> This is one of the first hints in the novel suggesting Adam’s values—an ethics which he observes in his own conduct and which he expects from other characters. Although the application of this ethics does not always work perfectly (in another scene he breaks Charlie’s hand in self defence, when the latter tries to inactivate his power switch and thus render him unconscious and incapable of any action), Adam—like other robots, as it turns out—seems to be driven by a moral code which strongly opposes treating others like mere machines (most of the novelistic androids learn how to inactivate the switch and prevent their ‘masters’ from exerting control over them). All these gestures testify, however, to the fact that the AI represented in McEwan’s novel clearly follows some ethics in which not harming both human and non-human others—physically or otherwise—stands high in the hierarchy of values.

The problem of ethics presented in the novel, however, is still more complex. As the plot of the novel reveals, Adam is also a strong opponent of lying, regardless of the circumstances and shows low tolerance of unfair play. As a result, he reports Miranda to the police, revealing her court lies to which she resorted seeking justice and revenge for her late friend. For Adam, complicated and dramatic circumstances do not justify lying and deceit; despite his high operational intelligence he seems not to recognise that sometimes justice may be done using imperfect means. This inability is connected with another feature of his personality / intelligence: his low tolerance of ambiguity, of situations when wrong actions may bring right results. He seems not to distinguish between higher and lower aims and treats each instance of breaking the law as criminal. This moral rigidity leads to his doing harm to his human friends, as he chooses ethics rather than love and friendship. All of these features show the novelistic robot as a rather inflexible creature, despite his high intelligence and impressive intellectual capacities. What he seems to lack are typical human flaws: imperfection, inconsistency, duplicity or hedonism. Serious and honest, driven by a clear ethics which may be perhaps compared to Kantian deontology with its rules-based principles of behaviour,<sup>20</sup> he turns out poorly adjusted to the world which seems quite flexible ethically and admits all kinds of exemptions and exceptions to the rules. Inevitably, then, he is positioned on a collision course with his human counterparts and the novel ends with his destruction as an agent threatening humans. Irena Książopolska comments that “this is the hubris of the machine: proclaiming the rule of generalities over the particular and individual, dismissal of the actual human beings as

<sup>19</sup> Isaac Asimov, “Runaround,” in: *I, Robot* (New York: Doubleday, 1950), 27.

<sup>20</sup> See Tae Wan Kim, “Flawed Like Us and the Starry Moral Law: Review of *Machines Like Me* by Ian McEwan,” *Journal of Business Ethics* 170 (2021): 876.

irrelevant compared to the higher ideals.”<sup>21</sup> More sympathetically, however, one may see this conclusion of the android’s fate as both tragic and ironic and, as it is revealed, this tragedy is also the fate of other robots mentioned in the novel: out of the original twenty five, at least a dozen either commit suicide or learn how to irrevocably destroy their operating systems. As it turns out, confronted with the human world, androids cannot tolerate its moral duplicity and ubiquitous evil and find it impossible to live in it. Too ethical and honest, they seem evidently out of place in the world of human complexity, hedonism, and imperfection.

The love triangle and the inevitable death of one of the lovers that solves the dramatic plot may resemble typical literary melodramas which often treat of impossible relationships against class or social barriers. The affinities of the novel to this convention, however, may go even further. As John G. Cawelti observes, one of the basic features of any melodrama, regardless of its particular realisation, is its revelatory function: as he notices, through melodrama, “we see not so much the working of individual fates but the underlying moral process of the world.”<sup>22</sup> Melodrama, then, presenting the fates of its protagonists may reveal the principles and values of the fictional universe. In the case of *Machines Like Me*, these values seem to be quite distant from the android impeccable ethics and in such a world an ethical robot is doomed to failure.

Writing about melodrama, Grażyna Stachówna observes that it is traditionally perceived as a ‘female’ genre. Pointing to its political anchorage in the 19th-century middle-class values, she emphasises its clear misogyny with the fetishisation and simultaneous restrictive control of female bodies, desires and social roles.<sup>23</sup> In its essence, classical melodrama is seen as ideologically hostile to women and employed in the process of their socialisation into society.<sup>24</sup> Interestingly, in McEwan’s novel the traditional place of a woman character is taken up by a robot which likewise seems to break accepted human social rules and is accordingly punished for it. His low social position in the futuristic society of the twenty first-century fiction seems to be similar to that occupied by women represented by the literature of the nineteenth century: in the fictional world of the novel a robot should be a pleasant and unproblematic servant and

<sup>21</sup> Irena Księżopolska, “Can Androids Write Science Fiction? Ian McEwan’s *Machines Like Me*,” *Critique: Studies in Contemporary Fiction* 63, no. 4 (2022): 418.

<sup>22</sup> John G. Cawelti, *Adventure, Mystery and Romance: Formula Stories as Art and Popular Culture* (Chicago: University of Chicago Press, 1976), 45f. It is worth noticing that Cawelti, following other scholars, defines melodrama first as a literary genre, in contrast to popular opinions which link it exclusively to cinematographic stories, admittedly better known nowadays.

<sup>23</sup> See Grażyna Stachówna, *Niedole miłowania: Ideologia i perswazja w melodramatach filmowych* (Kraków: Rabid, 2000), 20.

<sup>24</sup> See *ibidem*, 21f.

companion, and his breaking out of this role for whatever reason, even a noble one, cannot be tolerated. Though clearly morally superior, socially Adam the robot is inferior and his rebellion cannot succeed, even though the values that inform it may be officially professed by this world.

Adam's miserable end points, however, to one more ethical issue implied by the novel, namely the unclear legal status of intelligent machines and the lack of regulations concerning human-machine conduct. In the scene concluding the story the fictional character of Alan Turing thus reproaches the main protagonist: "My hope is that one day, what you did to Adam with a hammer will constitute a serious crime. Was it because you paid for him? Was that your entitlement? ... You weren't simply smashing up your own toy, like a spoiled child. You didn't just negate an important argument for the rule of law. You tried to destroy a life. He was sentient. He had a self. How it's produced, wet neurons, microprocessors, DNA networks, it doesn't matter. Do you think we're alone with our special gift? Ask any dog owner. This was a good mind, better than yours or mine, I suspect. Here was a conscious existence and you did your best to wipe it out. I rather think I despise you for that."<sup>25</sup>

The main protagonist's conduct shows his arrogant attitude towards the robot and the latter's status of an object which can be bought, sold or destroyed. Yet the fact that it is intelligent, sentient and conscious, that it has a complex emotional and moral life complicates this status considerably. Projecting a futuristic scenario, the novel points out quite clearly that human assumptions have to be reflected on and revised so as to accommodate intelligent machines into the common human-machine world of social and legal relations as treating them as mere appliances will be vastly inadequate. This is in keeping with the postulates articulated by AI scientists. Stuart Russell and Peter Norvig observe: "If robots become conscious, then to treat them as mere 'machines' (e.g., to take them apart) might be immoral. Robots also must themselves act morally—we would need to programme them with a theory of what is right and wrong. Science fiction writers have addressed the issue of robot rights and responsibilities, starting with Isaac Asimov (1942)... The stories (and the movies) convince one of the need for a civil rights movement for robots."<sup>26</sup>

On a final reflection, then, McEwan's novel not merely familiarises its readers with the so far imaginary situations of living with artificial intelligence and androids. More importantly, it invites them to reflect upon both their nature as sentient and conscious beings and on their status as valid members of an extended human-machine society. As McEwan comments, "if a machine seems like a human or you can't tell the difference, then you'd jolly well bet-

<sup>25</sup> McEwan, *Machines Like Me*, part X, loc. 3784–90.

<sup>26</sup> Russell and Norvig, *Artificial Intelligence*, 964.



ter start thinking whether it has responsibilities and rights and all the rest.”<sup>27</sup> Thus, representing an imaginary scenario, the novel both draws attention to legal, social and ethical issues resulting from the co-existence of people and advanced machines, and points to the basic imperfection and moral fallibility of the human world, which ironically may turn out to be a major obstacle in the process of creating a common human-machine community.

#### FRIENDS LIKE KLARA

The convention of melodrama, the android protagonist, its poignant end and ethical considerations link McEwan’s novel with the recent work by Nobel-Prize winning Kazuo Ishiguro, *Klara and the Sun* (2021).<sup>28</sup> More focused than *Machines Like Me*, set in a near-future America and narrated by the android protagonist, the novel tells a story of Klara, an Artificial Friend, i.e., a humanoid robot designed to keep company and be a friend to lonely teenagers. The plot traces her beginnings in a store, where she waits to be bought by a willing future owner, her subsequent relationship with Josie and her family who take her, and then her “slow fade” on a garbage heap where she is waiting again, this time for her imminent death. Tracing the trajectory of a robot’s life the story, in a genuinely melodramatic fashion, both imagines the vicissitudes of android existence and shows the features of human society and relationships.

The eponymous Klara makes a perfect AF: devoted, loyal and safe-effacing, she seems to have no private life, ambitions or goals. Her existence is filled with serving and her intelligence, abilities and talents help her constantly improve her performance; she seems to have a special gift of observation, which she develops not for her own sake but for that of her human friends. Like the memorable butler Stevens from Ishiguro’s *Remains of the Day*<sup>29</sup>, Klara devotes her whole life to serving the teenager who chose her. Yet she also exhibits more than standard features, associated rather with humans: she has dignity, manifested, e.g., in a scene in which she refuses to serve as a toy to other teenagers, rejects taunting and gets offended when bullied by irresponsible humans. She seems not just to observe but even perhaps add another law to Asimov’s famous three; her fourth commandment might be ‘a robot will try to comprehend a human being’, which she incessantly tries to do.

<sup>27</sup> Tim Adams, “Ian McEwan: ‘Who’s Going to Write the Algorithm for the Little White Lie?’” *The Guardian*, April 14, 2019, <https://www.theguardian.com/books/2019/apr/14/ian-mcewan-interview-machines-like-me-artificial-intelligence>.

<sup>28</sup> See Kazuo Ishiguro, *Klara and the Sun* (London: Faber & Faber, 2021).

<sup>29</sup> See I d e m, *The Remains of the Day* (London: Faber & Faber, 1989).



More interestingly still, Klara seems to be a religious robot: powered by solar batteries, she observes a sort of an AF solar cult in which the Sun features as a benevolent God granting life and spiritual and physical nourishment to the world. Klara prays to the Sun and worships him, ever thankful for the gift of life; she also believes in special grace that the Sun may choose to offer which is able to save life and help those in need. This is what she hopes for when her teenager friend Josie becomes deadly ill as a side effect of her medical procedure of ‘lifting’, i.e., advanced genetic editing which is supposed to increase intellectual potential of children and allow them to enter a higher caste of citizens. Seeing Josie on the verge of dying, Klara bargains with the Sun to grant Josie health and life. In a gesture of self-destruction, she sacrifices her own wellbeing depriving herself of a part of a fluid vital for her proper functioning as a gift to the Sun which may save Josie in return. Miraculously, the sacrifice works and Josie’s health becomes restored with Klara weaker but happy that her prayers and sacrifice were accepted and reciprocated. Her religiosity and selfless actions stand in a sharp contrast to the rest of the novelistic world populated by characters which seem not to exhibit any sort of religious or spiritual traits and who are focused on material manifestations of status, wealth and security. Religion and sacrifice seem to be so out of place in this world that they are not even a subject to discuss. Thus, it is her religious belief and selfless love for a human being, rather than just her status of a robot, that make Klara exceptional in this story: in a truly evangelical way she loves God, loves her friend, and is ready to sacrifice her life for her.

Klara’s “evangelical” spirit, which may perhaps be described as a quasi-Christian ethics of altruism, becomes visible once again in the last section of the novel when, old and no longer needed, she is discarded in a garbage lot, lonely reminiscing about her life. As Robert C. Abrams observes, her “slow fade” (as her dying is euphemistically referred to) may be seen as “unexpected object lessons in human aging and death.”<sup>30</sup> Fully conscious and sentient, Klara accepts her fate with dignity and patience; she re-experiences her past trying to locate important moments and episodes, convincing herself that her life had sense and value. In this psychologically healing process, she prepares herself for her death which she is to face alone. Robert C. Abrams observes: “The robotic structure of Klara’s thinking does nothing to obscure the fact that her exemplary personal growth and maturation mirror the human developmental journey at its best. It has been an evolution from concrete observation to com-

---

<sup>30</sup> Robert C. Abrams, “*Klara and the Sun*: Kazuo Ishiguro’s New Model for ‘Completion’ at the End of Life,” *Journal of the American Geriatrics Society* 70 (2022): 636.

plex emotion, concluding with a pair of crowning achievements, the ability to love selflessly and die with integrity.”<sup>31</sup>

Klara’s poignant end, however, though she accepts it with dignity and integrity, reflects once again on the human world in which agreement to one’s fate is rather absent. Genetic ‘lifting’ and other medical procedures demonstrate negating rather than embracing of natural processes; death is a calamity which can be attenuated by refusing to accept it, e.g., trying to create mechanic “continuations” of deceased relatives. In contrast to how they treat their own death, however, novelistic humans are not particularly sentimental about other creatures’ end. Klara’s life of good service ends when she is no longer needed; with Josie leaving for college, an Artificial Friend becomes useless and is unceremoniously disposed of in the junkyard. Her friendship and sacrifice go vastly unreciprocated, her love unrequited: though treated decently while in service, ultimately Klara is perceived as an object, a mechanical toy rather than a sentient individual. Thus, similar to McEwan’s story, in *Klara and the Sun* it is the robot that emerges as a noble and ethical creature in a novel full of rather selfish and cruel humans. Like in the previous novel, too, the melodramatic formula—a story of an impossible love and devotion of a robot towards a human being—exposes the rules of the fictional world. And like in the previous novel, these rules seem to be profoundly egoistic and materialistic: this is the world of no religion and little ethics, with self-obsessed humans ruthlessly pursuing their goals, hurting other creatures with no consideration for their wellbeing. Thus, apart from being an attempt at imagining a possible future scenario, the novel is also a rather bitter diagnosis of the ethical state of humanity at present.

Similar to *Machines Like Me*, Ishiguro’s novel employs the first-person narration, yet this time with the eponymous robot-character Klara acting as a character-narrator. Very much like in Ishiguro’s previous novel *Never Let Me Go*,<sup>32</sup> which likewise introduced an “inhuman” narrator, this structure of narration offers a possibility to see the world with the narrator’s eyes and from her perspective, too. Characteristic of rather experimental narratives, as Jan Alber, Henrik Skov Nielsen and Brian Richardson argue the so-called “unnatural” or non-human narrators and unexpected points of view function as a method to transgress automatised conventions and equally automatised viewpoints.<sup>33</sup> In the case of Klara this perspective may strike one as surprisingly childlike and realistic, focused on mundane details, with precise descriptions rendering both

<sup>31</sup> Ibidem, 637.

<sup>32</sup> See Kazuo Ishiguro, *Never Let Me Go* (London: Faber & Faber, 2005).

<sup>33</sup> See Jan Alber, Henrik Skov Nielsen, and Brian Richardson, “Unnatural Voices, Minds, and Narration,” in: *The Routledge Companion to Experimental Literature*, ed. Joe Bray, Alison Gibbons and Brian McHale (London: Routledge, 2012), 353.

her acute sense of observation and her lack of experience in the human world. Yet, as is perhaps already typical of Ishiguro's narrators, this simplicity and focus may be misleading: Klara as a narrator seems to reveal as much as she conceals and although her monologues include hardly any straightforward information about her emotions or opinions, occasional narrative hints help infer more complex processes taking place in her mind. Via the contrast between the childlike and sometimes naïve observations of the narrator which takes everything in good faith and the less than pleasant facts that she narrates the novel creates an ironic and poignant distance towards the world it describes. This deceptively simple and yet nuanced structure of narration allows the novel to indirectly show a much less naïve portrayal of the world in which the narrator functions.

Additionally, choosing a robot as a narrator may perform one more function: that of familiarization and empathization. Empathy may be defined as a "process of feeling, perceiving and understanding of another person's psychological state"<sup>34</sup> possible due to the ability to put oneself in someone else's situation. A narrative that imaginatively presents such a situation may greatly assist the process of empathy formation by presenting details and perspectives so far unimagined. It has therefore the power to sensitize readers towards new and alien experiences. A robot-narrator of the novel may work towards empathization and bringing the readers closer to the rather exotic internal life of artificial intelligence. So does its embodiment in a human-like shape, which may seem a strange choice given the fact that most of the AI used at present are computer programmes and that the creation of humanoid robots may seem not only very difficult but also undesirable, considering the human prejudice against and fear of creatures that deceptively resemble them. Yet, the human-like embodiment may once again work towards inspiring empathy towards robots, drawing attention to the fact that despite their mechanical origins, they are essentially not very different from humans, being not only conscious and sentient but also undergoing the same processes of developing, performing, degenerating and finally dying. Both the humanoid shape and the narrative voice, then, not only assert the character of Klara as a fully developed individual; more importantly, they try to convince the readers that robots are creatures like them, deserving the same treatment and sympathy. And as Santiago Mejia and Dominique Nikolaidis conclude, thus constructed character "forces us to confront interesting ethical questions concerning our lives with intelligent machines: 'how do we relate to them?' [or] 'what kind of autonomy and dignity do they have?'"<sup>35</sup>

<sup>34</sup> Józef Rembowski, *Empatia: Studium psychologiczne* (Warszawa: Wydawnictwo PWN, 1989), 69.

<sup>35</sup> Santiago Mejia and Dominique Nikolaidis, "Through New Eyes: Artificial Intelligence, Technological Unemployment, and Transhumanism in Kazuo Ishiguro's *Klara and the Sun*," *Journal of Business Ethics* 178 (2022): 304.

\*

With their futuristic settings, grim visions of social life with technological progress doing little to alleviate inequalities and injustice, and tragic endings of robotic protagonists both novels seem to question the simplistic division into the technophilic or technophobic trends. Rather, constructing their narratives around sentient and self-conscious robots which inhabit a human world they defamiliarize the latter, exposing it as full of inconsistencies, injustices and moral ambivalence. Like the Martian from Craig Raine's poem, robots observe with puzzlement humans and their habits, becoming disoriented and disappointed.<sup>36</sup> Simultaneously, however, apart from defamiliarization, an opposite process seems to take place in the novels which one could perhaps describe as "familiarisation," i.e., making the readers familiar with and more sensitive towards creatures and characters so far alien and inscrutable. Moreover, as these creatures are clearly presented as ethically superior—although representing diverse ethical positions that may be roughly compared to Kantian deontology and Christian altruism—the texts indirectly seem to examine the ethics of the human world which emerges as hedonistic and self-centred. The fact that these creatures are machines rather than other, more traditional aliens may perhaps testify to the technophilic bend of the two works. Both novels imply quite strongly that the future machines not only will occupy a place in human societies but also that this place should perhaps be carefully rethought rather than unreflectively assumed. In the represented world of both McEwan's and Ishiguro's stories robots are situated between appliances, pet animals and servants, endowed with no status or rights, but precisely this position comes under the narrative scrutiny and is problematised in their plots. Both novels suggest that introducing sentient and self-conscious machines into a human world will have not just practical but also moral and legal implications and that traditional human ethics, especially this practised rather than just professed, will have to accommodate them, too. The novels' melodramatic structure reveals, then, not merely poignant and tragic stories of their main protagonists but also the larger structure of the fictional world and—by extension—the extra-textual world inhabited by the readers. This revelation might seem ironic as human beings are portrayed as rather inferior morally to perhaps childishly naïve yet ethically accomplished machines.

Despite robot protagonists, mildly technophilic attitudes and futuristic high-tech settings, both novels seem rather far removed from transhumanist optimism one could expect from thus characterised fiction. Rather than portraying a brave new human living to the utmost of his/her potential and making

---

<sup>36</sup> See Craig Raine, *A Martian Sends a Postcard Home* (Oxford: Oxford University Press, 1980).

the full use of advanced technologies, as transhumanists perhaps would wish,<sup>37</sup> McEwan's and Ishiguro's novels may come closer to posthumanist reflection on the common human and non-human co-existence. Posthumanism—which in its critical version may be broadly understood as a reflection that displaces the human from his privileged position and revises the anthropocentric paradigm of thought and science—posits a more equal relation between the human and non-human. In Michael Hauskeller's conceptualisation it “refuses to see humans as a superior species in the natural order, ontologically distinct from animals on the one hand, and machines on the other ... [and—B. K.] insists that the boundaries between the human and non-human are rather fluent and in fact have always been so.”<sup>38</sup> To Hauskeller, exploding the dividing line between the human and non-human carries with it political implications as it leads to the flattening of hierarchies and binary oppositions that privilege one side (human) only; as he observes, “at the heart of post-humanism is clearly a liberationist ideal: the hoped-for redistribution of difference and identity is ultimately a redistribution of power.”<sup>39</sup> The two novels may thus be interpreted as positing a vision of flat posthuman world and a less hierarchical posthuman society, in keeping with the tenets of critical posthumanism. “Posthumanism is a ‘post’—according to Francesca Ferrando's definition—to the notion of the ‘human,’ located within the historical occurrence of ‘humanism’ (which was founded on hierarchical schemata), and in an uncritical acceptance of ‘anthropocentrism,’ founded upon another hierarchical construct based on speciesist assumptions. Both the notion of the ‘human’ and the historical occurrence of ‘humanism,’ have been sustained by reiterative formulations of symbolic ‘others,’ which have functioned as markers of the shifting borders of who and what would be considered ‘human’: non-Europeans, non-whites, women, queers, freaks, animals, and automata, among others, have historically represented such oppositional terms.”<sup>40</sup>

Both *Machines Like Me* and *Klara and the Sun*, therefore may be interpreted as gestures pointing in the direction of inclusiveness and abandoning of hierarchies, inviting their readers to consider the possible equality—legal and social—of humans and machines. The title of McEwan's novel reads “machines like me” but the subtitle adds to it “and people like you,” thus juxtaposing the two ‘species’ and showing them as—still—contrastive. Yet in a truly

<sup>37</sup> See Max More, “The Philosophy of Transhumanism,” in *The Transhumanist Reader*, ed. Max More and Natasha Vita-More (Chichester: Wiley & Blackwell, 2013), 3–17.

<sup>38</sup> Michael Hauskeller, “Utopia in Trans- and Posthumanism,” in *Post- and Transhumanism: An Introduction*, ed. Robert Ranisch and Stefan Lorenz Sorgner (Frankfurt-am-Main: Peter Lang, 2014), 104.

<sup>39</sup> Idem, *Mythologies of Transhumanism* (Cham: Palgrave Macmillan, 2016), 23.

<sup>40</sup> Francesca Ferrando, *Philosophical Posthumanism* (London: Bloomsbury, 2019), 24.

utopian spirit of a mental ‘what if’ experiment the novels project a fictional world in which perhaps the situation of the species equality does not yet take place but which strongly implies that it is at least imaginable. The scenario the two novels offer for consideration posits that artificial intelligence may become a normal part of the world with AI, being equipped with the features usually associated with humans, i.e. sensibility, intelligence, self-consciousness, dignity and morality, enjoying the legal and social status equal to humans. The texts seem to illustrate the proposition of James Moor who suggests considering a chance that “it’s possible that someday robots will be good ethical decision-makers [...] acting ethically on the basis of a moral understanding.”<sup>41</sup> Doing so, the two novels go a step beyond empathising the readers with AI and suggest reflection on the possible expansion of the world model to include other than human actors and thus familiarise the reading public with the idea that AI may become a part of life with equal position and rights to human beings.

#### BIBLIOGRAPHY / BIBLIOGRAFIA

- Abrams, Robert C. “*Klara and the Sun*: Kazuo Ishiguro’s New Model for ‘Completion’ at the End of Life.” *Journal of the American Geriatrics Society* 70 (2022): 636–37.
- Adams, Tim. “Ian McEwan: ‘Who’s Going to Write the Algorithm for the Little White Lie?’” *The Guardian*, 14 April 2019. <https://www.theguardian.com/books/2019/apr/14/ian-mcewan-interview-machines-like-me-artificial-intelligence>.
- Alber, Jan, Henrik Skov Nielsen, and Brian Richardson, “Unnatural Voices, Minds, and Narration.” In *The Routledge Companion to Experimental Literature*. Edited by Joe Bray, Alison Gibbons, and Brian McHale. London: Routledge, 2012.
- Asimov, Isaac. “Runaround.” In *I, Robot*. New York: Doubleday, 1950.
- Booch, Grady. “Don’t Fear Superintelligent AI”, TEDtalk. Youtube, March 13, 2017. <https://www.youtube.com/watch?v=z0HsPBKfhoI>.
- Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014.
- Cawelti, John G. *Adventure, Mystery and Romance: Formula Stories as Art and Popular Culture* Chicago: University of Chicago Press, 1976.
- Ferrando, Francesca. *Philosophical Posthumanism*. London: Bloomsbury, 2019.
- Hauskeller, Michael. “Utopia in Trans- and Posthumanism.” In *Post- and Transhumanism: An Introduction*. Edited by Robert Ranisch and Stefan Lorenz Sorgner. Frankfurt-am-Main: Peter Lang, 2014.
- . *Mythologies of Transhumanism*. Cham: Palgrave Macmillan, 2016.

<sup>41</sup> James H. Moor, “Four Kinds of Ethical Robots,” *Philosophy Now*, no. 72 (2009), [https://philosophynow.org/issues/72/Four\\_Kinds\\_of\\_Ethical\\_Robots](https://philosophynow.org/issues/72/Four_Kinds_of_Ethical_Robots).



- Ishiguro, Kazuo. *Klara and the Sun*. London: Faber & Faber, 2021.
- . *Never Let Me Go*. London: Faber & Faber, 2005.
- . *The Remains of the Day*. London: Faber & Faber, 1989.
- Kim, Tae Wan, “Flawed Like Us and the Starry Moral Law: Review of *Machines Like Me* by Ian McEwan.” *Journal of Business Ethics* 170 (2021): 875–79.
- Księżopolska, Irena. “Can Androids Write Science Fiction? Ian McEwan’s *Machines Like Me*.” *Critique: Studies in Contemporary Fiction* 63, no. 4 (2022): 414–29.
- Kurzweil, Ray. “We Will Be More Fun, Sexier and More Creative.” *Svenska Dagbladet*. December 19, 2019. <https://www.svd.se/a/d437dfb6-179b-4046-930e-dcfdbf620643>.
- Landgrebe, Jobst, and Barry Smith. *Why Machines Will Never Rule the World: Artificial Intelligence without Fear*. New York: Routledge, 2023.
- McEwan, Ian. *Machines Like Me*. London: Jonathan Cape, 2019, e-book.
- Mejia, Santiago, and Dominique Nikolaidis. “Through New Eyes: Artificial Intelligence, Technological Unemployment, and Transhumanism in Kazuo Ishiguro’s *Klara and the Sun*.” *Journal of Business Ethics* 178 (2022): 303–6.
- de Miranda, Luis. *AI and Robotics*. London: Ivy Press, 2018.
- Moor, James H. “Four Kinds of Ethical Robots.” *Philosophy Now*, no. 72 (2009). [https://philosophynow.org/issues/72/Four\\_Kinds\\_of\\_Ethical\\_Robots](https://philosophynow.org/issues/72/Four_Kinds_of_Ethical_Robots).
- Moravec, Hans P. *Robot: Mere Machine to Transcendent Mind*. Oxford: Oxford University Press, 2000.
- More, Max. “The Philosophy of Transhumanism.” In *The Transhumanist Reader*. Edited by Max More and Natasha Vita-More. Chichester: Wiley & Blackwell, 2013.
- Nicieja, Stankomir. “Revisiting Utopia: New Directions for Utopian Fiction in Margaret Atwood’s *Oryx and Crake* and Kazuo Ishiguro’s *Never Let Me Go*.” In *Margins and Centres Reconsidered*. Edited by Barbara Klonowska and Zofia Kolbuszewska. Lublin: Towarzystwo Naukowe Katolickiego Uniwersytetu Jana Pawła II, 2008.
- Przegalińska, Aleksandra. “Zrozumieć człowieka.” *Academia* 1–2 (2019): 10–13.
- Raine, Craig. *A Martian Sends a Postcard Home*. Oxford: Oxford University Press, 1980.
- Rembowski, Józef. *Empatia: Studium psychologiczne*. Warszawa: Wydawnictwo PWN, 1989.
- Russell, Stuart, and Peter Norvig. *Artificial Intelligence: A Modern Approach*. New Jersey: Pearson, 2003.
- Stachówna, Grażyna. *Niedole miłowania: Ideologia i perswazja w melodramatach filmowych*. Kraków: Rabid, 2001.
- Torczyńska, Monika. “Sztuczna inteligencja i jej społeczno-kulturowe implikacje w codziennym życiu.” *Kultura i Historia* 36, no. 2 (2019): 106–26.



## ABSTRACT / ABSTRAKT

Barbara KLONOWSKA, Ethical Machines: Representations of Artificial Intelligence in Ian McEwan's *Machines Like Me* and Kazuo Ishiguro's *Klara and the Sun*

DOI 10.12887/36-2023-4-144-08

Artificial intelligence provokes contradictory reactions that range from enthusiasm and fear and that may be generalised as instances of technophobia and technophilia. AI is also a perennial theme of numerous fictional narratives which seem especially important as, due to their large audiences, images and stories representing AI enter popular debates. Two recent novels by eminent British authors, Ian McEwan's *Machines Like Me* (2019) and Kazuo Ishiguro's *Klara and the Sun* (2021), in either alternative past or futuristic settings also take up the issue of AI and its possible ramifications. The article argues that the two works represent AI as both beneficial and potentially problematic, posing the question about the essence of humanity and the limits of AI, problematising the status of intelligent machines and familiarising readers with ethical and legal problems they bring; they also try to build empathy and sensitise the public towards creatures other than humans.

Keywords: AI, technophobia, self-consciousness, posthumanism

Contact: Department of English Literature and Culture, Institute of Literary Studies News, Faculty of Humanities, John Paul II University of Lublin, Al. Racławickie 14, 20-950 Lublin, Poland

E-mail: barbara.klonowska@kul.pl

ORCID: 0000-0001-8327-854X

Barbara KLONOWSKA – Etyczne maszyny. Reprezentacje sztucznej inteligencji w powieściach *Maszyny takie jak ja* Iana McEwana i *Klara i słońce* Kazuo Ishiguro

DOI 10.12887/36-2023-4-144-08

Sztuczna inteligencja prowokuje sprzeczne emocje entuzjazmu i strachu, które mogą być powiązane z szerszymi postawami technofobii lub technofilii. Jest ona też stałym tematem wielu filmów i powieści, co wydaje się szczególnie istotne, gdyż często to poprzez filmy adresowane do szerokiej widowni temat SI wkracza do powszechnej debaty. Dwie powieści wybitnych brytyjskich prozaików, *Maszyny takie jak ja* Iana McEwana i *Klara i słońce* Kazuo Ishiguro, jedna w scenerii futurystycznej, druga snując alternatywną historię, również podejmują ten temat z jego nieoczywistymi konsekwencjami. Artykuł dowodzi, że obie powieści ukazują SI jako jednocześnie pożyteczną i problematyczną, stawiając pytania o istotę człowieczeństwa i granice sztucznej inteligencji, problematyzując status inteligentnych maszyn i oswajając czytelników z etycznymi i prawnymi problemami z nimi związanymi. Próbuje one też wzbudzić empatię i uwrażliwić czytelnika na los istot innych niż tylko ludzkie.

Słowa kluczowe: sztuczna inteligencja, technofobia, samoświadomość, post-humanizm

Kontakt: Instytut Literaturoznawstwa, Wydział Nauk Humanistycznych, Katolicki Uniwersytet Lubelski Jana Pawła II, Al. Raławickie 14, 20-950 Lublin

E-mail: [barbara.klonowska@kul.pl](mailto:barbara.klonowska@kul.pl)

ORCID: 0000-0001-8327-854X

Grzegorz TRĘBICKI

MARIAŻ DWÓCH ŚWIATÓW  
Quasi-chrześcijańska wizja ewolucji człowieka i kosmosu  
w „Pieśniach Hyperiona” Dana Simmonsa

*Wizja wszechświata, jaka dominuje w większości tekstów space opery wielkiego przełomu, ma podobny charakter. Kosmos jawi się tam jako miejsce metafizycznie „puste”, obojętne wobec człowieka (i innych istot rozumnych), a niekiedy wręcz wrogie, nieposiadające z samej swej natury żadnego głębszego znaczenia ani celu. Jakikolwiek sens może zostać nadany wyłącznie w wyniku działań ludzi (lub innych istot rozumnych) i często są to wysiłki tyleż heroiczne, co z góry skazane na niepowodzenie.*

W pewnym uproszczeniu można uznać, że cała współczesna kultura (w pojęciu „kultura” zawrzemy również naukę, w tym nauki ścisłe) przesiąknięta jest nieustającym napięciem między dwiema antagonistycznymi wizjami świata (wszechświata). Pierwszą z nich moglibyśmy określić ogólnie jako religijną, czy też przynajmniej parareligijną i holistyczną, drugą – jako materialistyczną i redukcjonistyczną. Wizja religijna podkreśla istnienie transcendentnego sensu, głębszego zamysłu, który można odnieść zarówno do ludzkiego życia, jak i do całego kosmosu. Zakłada ona również swoistą „realność” tych wymiarów rzeczywistości, których – zdaniem zwolenników tego światopoglądu – nie da się poddać skrupulatnej analizie za pomocą dostępnych obecnie naukom przyrodniczym narzędzi (przykładem takiego wymiaru mogłaby być, chociażby, tak czy inaczej rozumiana duchowość). Według drugiej wizji całą rzeczywistość (w tym również, jak sugeruje wielu współczesnych neurobiologów, zagadkę ludzkiego umysłu i duszy) można zredukować do całkowicie opisywalnych i dających się wyjaśnić poprzez wspomniane narzędzia mechanizmów fizycznych czy też biologicznych. Jakikolwiek „sens” lub „duchowość” nie mają charakteru transcendentnego (autonomicznego), lecz co najwyżej emergentny (wyłaniają się z owych podstawowych procesów), być może stanowiąc przy tym jedynie skutek uboczny zorientowanej wyłącznie na przetrwanie gatunku ewolucji albo też pewnego rodzaju iluzję poznawczą.

Rozdźwięk między obydwoma światopoglądami stał się szczególnie wyraźny w ostatnich dziesięcioleciach, kiedy to dynamiczny rozwój z jednej strony współczesnej fizyki (a zwłaszcza takich jej dziedzin, jak kosmologia czy mechanika kwantowa), z drugiej zaś neurobiologii odsłonił całkowicie nieznane dotąd oblicza rzeczywistości, dając nowy impuls do debaty o charakterze bardziej już filozoficznym na temat jej prawdziwej natury. Debata ta znalazła również odzwierciedlenie w fikcji literackiej, zwłaszcza tej o cha-

rakterze spekulatywnym, czyli literaturze niemimetycznej albo też, używając języka potocznego, tak zwanej „fantastyce”.

Z całego ogromnego gatunkowego bogactwa literatury niemimetycznej wybierzemy do dalszych rozważań pewną konkretną odmianę (subgatunek?) „twardej” science fiction<sup>1</sup> (fantastyki naukowej), która w szczególny sposób nawiązuje zarówno do tematyki ontologicznej i eschatologicznej, jak i do najnowszych osiągnięć nauk ścisłych. Mamy tu na myśli zbiór tekstów, który wstępnie opisano jako „space opera wielkiego przełomu”<sup>2</sup>. Jego najistotniejsze cechy w kontekście dyskusji w niniejszym artykule to:

(1) Prezentacja wizji ludzkości w obliczu kluczowego przełomu (cywilizacyjnego, technologicznego, o istotnych implikacjach społecznych, psychologicznych czy wręcz ontologicznych) lub zagrożenia (motywy katastrofy kosmicznej bądź ekologicznej, kontaktu [konfliktu] z obcą cywilizacją albo sztuczną inteligencją).

(2) Wyraźne i rzetelne (nie tylko czyste pretekstowe czy „scenograficzne”) nawiązania do aktualnych teorii naukowych, przede wszystkich kosmologii i fizyki kwantowej.

(3) Tworzenie powieściowego uniwersum na spektakularną skalę, obejmującą rozległe połacie znanego i nieznanego kosmosu, mnogość systemów gwiazdnych i chronologię rozciągającą się na dziesiątki, setki czy nawet tysiące lat.

(4) Wprowadzenie tematyki eschatologicznej i ontologicznej, reinterpretowanej jednak w paradygmacie nie tyle religijnym, co stricte racjonalistycznym, naukowym.

Podkreślmy też dobitnie, iż w przypadku tego subgatunku czy też kategorii literatury spekulatywnej mamy do czynienia ze zjawiskiem, które moglibyśmy opisać jako spekulację „totalną” i „realistyczną”. Innymi słowy – spotykamy się z przedstawieniem względnie całościowej wizji dotyczącej natury wszechświata oraz ewolucji ludzkości, która (choć jest oczywiście jedynie literacką ekstrapolacją) pozostaje w miarę możliwości niesprzeczna z dotychczasowym stanem wiedzy naukowej, stanowiącej główne źródło jej inspiracji i (czysto teoretycznie przynajmniej) mogłaby się w pewnych aspektach spełnić. Spekulacja taka wyraźnie więc stroni od parodii, groteski czy absurdu; nie stanowi jedynie literackiej gry czy intelektualnej zabawy, ale raczej przedsięwzięcie jak najbardziej poważne, ambitną próbę zmierzenia się z fundamentalnymi pytaniami.

<sup>1</sup> Ponieważ nie chciałbym w tym miejscu wikłać się w skomplikowaną dyskusję o charakterze genologicznym, zaznaczę jedynie, że pisząc „twarda” SF, mam na myśli zbiór tekstów w wyraźny sposób zainspirowany najnowszymi osiągnięciami nauk ścisłych czy też przyrodniczych i podejmujących racjonalną próbę ekstrapolacji tychże osiągnięć.

<sup>2</sup> Zob. G. T r e b i c k i, *Cixin Liu's The Remembrance of Earth's Past Trilogy: In Search of Interpretative Grounds*, w: *W kręgu zagadnień literatury powszechnej*, red. R. Kotowski, I. Mityk, Muzeum Narodowe w Kielcach, Kielce 2019, s. 215-228.

Początków subgatunku zapewne należy doszukiwać się już w latach pięćdziesiątych dwudziestego wieku, a nawet wcześniej, w twórczości takich klasyków „złotej ery SF”, jak Olaf Stapledon, Isaac Asimov czy Arthur C. Clarke. Jednak w pełni rozwinął się on do postaci, jaką opisaliśmy powyżej, dopiero w ostatnich dekadach poprzedniego stulecia, kiedy to oszołamiające zdobycze fizyki kwantowej przedostały się do szerszej świadomości społecznej i kiedy dorosły kolejne pokolenia autorów w miarę kompetentnie orientujących się w meandrach najnowszych teorii naukowych. Elementy subgatunku – w mniejszym lub większym natężeniu – pojawiają się więc od początku lat osiemdziesiątych w dziełach takich autorów, jak Jack McDevitt (na przykład w cyklu „Akademia”)<sup>3</sup>, Alistar Reynolds (między innymi w serii „The Inhibitor Sequence”)<sup>4</sup>, John C. Wright (w „The Golden Age”)<sup>5</sup>, Vernon Vinge (na przykład w powieści *Ogień nad otchłanią*)<sup>6</sup>, czy też polskich twórców – wymieńmy tu zwłaszcza Pawła Majkę (*Niebiańskie pastwiska*)<sup>7</sup> i Marcina Podlewskiego (*Głębia*)<sup>8</sup>.

Kwintesencjonalne – i jednocześnie chyba najgłośniejsze – dzieło tego trendu stanowi wydana również w Polsce trylogia chińskiego pisarza Cixina Liu w serii „Wspomnienie o przyszłości Ziemi”<sup>9</sup>. Imponując naukową erudycją oraz swoistym ontologicznym i eschatologicznym rozmachem – akcja powieści zaczyna się w czasach chińskiej rewolucji kulturalnej, a kończy wraz ze śmiercią całego wszechświata miliony lat później – dzieło to egzemplifikuje w sposób niemal perfekcyjny charakterystyczny dla całego trendu redukcjonizm ontologiczny. Konflikt między dwoma światopoglądami w tym właśnie subgatunku prozy fantastycznonaukowej wydaje się dawno rozstrzygnięty. Dominuje w nim ewidentnie redukcjonistyczna, czysto materialistyczna wizja wszechświata. Wszelkiego rodzaju religijność (czy choćby parareligijność) pojawia się tutaj jedynie jako czysto społeczny czy też psychologiczny element opisu świata przedstawionego. Metafizyka może zaistnieć tylko jako wytwór

<sup>3</sup> Zob. J. M c D e v i t t, *The Engines of God*, Ace Books, New York 1994; t e n ż e, *Cauldron*, Ace Books, New York 2007.

<sup>4</sup> Zob. A. R e y n o l d s, *Revelation Space*, Gollancz, London 2000; t e n ż e, *Redemption Ark*, Gollancz, London 2002; t e n ż e, *Absolution Gap*, Gollancz, London 2003; t e n ż e, *Inhibitor Phase*, Gollancz, London 2021.

<sup>5</sup> Zob. J. W. W r i g h t, *The Golden Age*, TOR Books, New York 2002; t e n ż e, *The Phoenix Exultant*, TOR Books, New York 2003; t e n ż e, *The Golden Transcendence*, TOR Books, New York 2003.

<sup>6</sup> Zob. V. V i n g e, *Ogień nad otchłanią*, tłum. Rafał Wikoński, Prószyński i S-ka, Warszawa 1998.

<sup>7</sup> Zob. P. M a j k a, *Niebiańskie pastwiska*, Rebis, Poznań 2015.

<sup>8</sup> Zob. M. P o d l e w s k i, *Głębia*, t. 1-4, Fabryka Słów, Lublin 2015-2018.

<sup>9</sup> Zob. C. L i u, *Problem trzech ciał*, tłum. A. Jankowski, Rebis, Poznań 2016; t e n ż e, *Ciemny las*, tłum. A. Jankowski, Rebis, Poznań 2017; t e n ż e, *Koniec śmierci*, tłum. A. Jankowski, Rebis, Poznań 2018.

ludzkiej kultury, nie zaś jako realny aspekt rzeczywistości, ta bowiem opisywana jest w całości przez nauki ścisłe.

Zaznaczyć trzeba, że podejście to związane jest, jak się zdaje, nie tyle z programowym ateizmem czy też antyreligijnymi sentymentami poszczególnych twórców albo różnego rodzaju uwarunkowaniami politycznymi (co w przypadku Cixina Liu, obywatela Chińskiej Republiki Ludowej, na pierwszy rzut oka mogłoby być uzasadnionym tropem), ile raczej z oddaniem nastrojów obecnych we współczesnych naukach ścisłych i neurobiologii. Panuje tam bowiem dość powszechne (choć, oczywiście, nie absolutne) przekonanie o prawdziwości wizji redukcjonistycznej (w rozumieniu naukowym czy filozoficznym, nie zaś politycznym czy ideologicznym). Koryfeusze i popularyzatorzy współczesnej nauki, reprezentujący rozmaite jej dziedziny, tacy jak Stephen Hawking, Richard Dawkins, Francis Crick czy Brian Greene, wielokrotnie określali się jako zwolennicy mniej lub bardziej twardej wersji owego redukcjonizmu<sup>10</sup>. Głosy wyrażające stanowisko odmienne, zakładające przynajmniej możliwość kompatybilności rzeczywistość opisywanej przez współczesną naukę i metafizyki (wymieńmy tu znakomitego polskiego fizyka teoretycznego i również teologa, Michała Hellera) wydają się w wyraźnej mniejszości<sup>11</sup>.

Wizja wszechświata, jaka dominuje w większości tekstów „space opery wielkiego przełomu”, ma podobny charakter. Kosmos jawi się tam jako miejsce metafizycznie „puste”, obojętne wobec człowieka (i innych istot rozumnych), a niekiedy wręcz wrogie, nieposiadające z samej swej natury żadnego głębszego znaczenia ani celu. Jakikolwiek sens może zostać nadany wyłącznie w wyniku działań ludzi (lub innych istot rozumnych) i często są to wysiłki tyleż heroiczne, co z góry skazane na niepowodzenie. Niekiedy wizja ta przybiera charakter skrajnie pesymistyczny – jak we wspomnianej powyżej trylogii Cixina Liu, gdzie przyjęte założenia co do natury wszechświata prowadzą do nieuniknionego kresu wszystkich rzeczy<sup>12</sup>.

<sup>10</sup> Zob. np. F. Crick, *Zdumiewająca hipoteza, czyli nauka w poszukiwaniu duszy*, tłum. B. Chacińska-Abramowicz, M. Abramowicz, Prószyński i S-ka, Warszawa 1997; R. Dawkins, *Samolubny gen*, tłum. M. Skoneczny, Prószyński i S-ka, Warszawa 2012; B. Greene, *Do końca czasu. Umysł, materia i nasze poszukiwanie sensu w zmieniającym się wszechświecie*, tłum. T. Krzysztoń, Prószyński i S-ka, Warszawa 2021; A. Drgan, *Kwantechizm 2.0, czyli klatka na ludzi*, Otwarte, Kraków 2022.

<sup>11</sup> Zob. np. M. Heller, *Czy fizyka jest nauką humanistyczną?*, Copernicus Center Press, Kraków 2016; t e n ż e, *Filozofia przypadku. Kosmiczna fuga z preludium i codą*, Copernicus Center Press, Kraków 2012. Bardzo ciekawe podsumowanie dyskusji związanych z redukcjonistycznym i religijnym widzeniem świata w świetle najnowszych teorii naukowych przedstawiła Kitty Ferguson (zob. K. Ferguson, *Ogień w równaniach. Nauka, religia i poszukiwania Boga*, tłum. P. Amsterdamski, Copernicus Center Press, Kraków 2016).

<sup>12</sup> Zob. G. Trębicki, *Kres wszystkich rzeczy. „Wspomnienie o przeszłości ziemi” Cixina Liu vs. „Cykl Hyperioński” Dana Simmonsa* (w przygotowaniu do publikacji w „Literatura i kultura popularna” 2023).

Na tym tle niezwykle interesująco przedstawiają się wydane w ostatniej dekadzie dwudziestego wieku „Pieśni Hyperiona”, tetralogia znanego amerykańskiego autora fantastyki Dana Simmons<sup>13</sup>. W dziele tym mamy do czynienia z czymś, co można by określić jako próbę unifikacji dwóch światów: rzeczywistości fizycznej oraz metafizycznej. Cykl obfituje w interesujące nawiązania nie tylko do kosmologii i fizyki kwantowej, ale również do filozofii i – co nie zdarza się tak często we współczesnej twardej scenie fiction – do teologii chrześcijańskiej.

Fabula tetralogii jest niezwykle skomplikowana i imponuje podobnym rozmachem co wspomniana wcześniej trylogia Cixina Liu. Obejmuje kilka stuleci, w najprzeróżniejszych lokalizacjach na dziesiątkach planet, a także w przestrzeni kosmicznej, angażuje dziesiątki, jeśli nie setki postaci. Wielowątkowa kompozycja tekstów, nawiązująca zarówno do *Opowieści kanterberyjskich* Geoffreya Chaucera<sup>14</sup>, jak i do rozmaitych dzieł wielkiego angielskiego poety Johna Keatsa<sup>15</sup>, również ze wszech miar godna jest uwagi literaturoznawcy. Dokładne omówienie całego dzieła, czy choćby samej tylko fabuły, byłoby jednak niemożliwe w ramach pojedynczego artykułu, skoncentrujemy się zatem na motywach najistotniejszych dla naszych rozważań.

Akcja powieści rozpoczyna się w dalekiej przyszłości, w dwudziestym dziewiątym wieku, kiedy to rządzona przez organizację zwaną Hegemonią ludzkość skolonizowała dziesiątki planet. Stało się to możliwe dzięki cudownym wynalazkom takim, jak transmitery – portale umożliwiające natychmiastową komunikację między poszczególnymi światami oraz statki z napędem Hawkinga, pozwalające na podróże z prędkością ponadświatłą. Osiągnięcie tego stanu było możliwe dzięki symbiozie z TechnoCentrum, cywilizacją sztucznych inteligencji, które częściowo uniezależniły się od ludzkości, ale nadal wydają się pozostać z nią w sojuszu, przynoszącym korzyść obu stronom. Główne zagrożenia upatrywane są w inwazji tak zwanych Intruzów – zmodyfikowanych genetycznie potomków ludzi zamieszkujących peryferia znanego wszechświata.

Pierwsze dwie powieści cyklu relacjonują wyprawę siedmiorga pielgrzymów do Grobowców Czasu, tajemniczej obcej struktury, która znajduje się na opuszczonej planecie Hyperion, strzeżonej przez przerażającą istotę zwaną

---

<sup>13</sup> Cykl składa się zasadniczo z dwóch dylogii. Pierwsza obejmuje tomy: *Hyperion* (t. 1-2, tłum. A. Nakoniecznik, Amber, Warszawa 1994) oraz *Zagłada Hyperiona* (tłum. R. Januszewski i M. Rudnik, Amber, Warszawa 1997), druga tomy: *Endymion* (tłum. W. Szypuła, Amber, Warszawa 1998) oraz *Triumf Endymiona* (tłum. W. Szypuła, Amber, Warszawa 1999). Tytuł drugiego tomu w późniejszej edycji wydawnictwo Mag został zmieniony na *Upadek Hyperiona*.

<sup>14</sup> Zob. G. C h a u c e r, *Opowieści kanterberyjskie*, tłum. J. Zawadzki, Biblioteka Śląska, Katowice 2022.

<sup>15</sup> Na przykład do poematów *Hyperion* i *Endymion* (zob. J. K e a t s, *33 wiersze*, wybór, tłum. i oprac. S. Barańczak, Wydawnictwo Znak, Kraków 2007).



Chyżwarem (która zdaje się posiadać nadprzyrodzone wręcz moce), a także szersze konsekwencje tejże wyprawy. Każdy z pielgrzymów niesie ze sobą (i w pierwszym tomie opowiada pozostałym) własną historię, rzucając coraz to nowe światło na rozgrywającą się akcję.

W kontekście naszych rozważań interesująca jest zwłaszcza pierwsza z nich, wprowadzająca motywy religijne opowieść katolickiego księdza, Leynarta Hoyta, który relacjonuje swoje poszukiwania innego duchownego, jezuitę Paula Dure, zaginionego uprzednio na planecie Hyperion. W trakcie poszukiwań okazuje się, że Dure badał wioskę Bikurów, prymitywnego ludu, prawdopodobnie potomków ziemskich osadników. Dure odkrywa, że Bikurowie uzyskali pewnego rodzaju nieśmiertelność, wskrzeszani przez przyczepionego do ich ciał pasożyta w kształcie krzyża (później zwanego krzyżokształtem). Nieśmiertelność tę okupują jednak degeneracją umysłową i utratą płciowości. Gdy wskutek interwencji Chyżwara Dure sam zostaje zainfekowany pasożytem, jezuita desperacko próbuje się go pozbyć, uznając, że stanowi on bluźnierczą parodię prawdziwego zmartwychwstania. Paraliżujący ból nie pozwala jednak mu ani wyciąć krzyżokształtu, ani opuścić wioski Bikurów. Ostatecznie z relacji Hoyta dowiadujemy się, że jezuita udało się zrealizować swój cel, przybijając się do drzewa Tesli – endemicznej, „drzewopodobnej” formy życia akumulującej ładunki elektryczne. Tak oto ojciec Dure przez niemal siedem lat jest nieustannie na przemian zabijany przez drzewo i ożywiany przez krzyżokształt, aż wreszcie jego ciało całkiem się rozpada, a on sam zostaje uwolniony od pasożyta. W akcji powieści okazuje się również, że Lenart Hoyt został zainfekowany dwoma krzyżokształtami – swoim „własnym” i tym, który odpadł od ciała ojca Dure. W tomie drugim, *Zagłada Hyperiona*, Hoyt ginie zabity przez Chyżwara i z drugiego krzyżokształtu ponownie odradza się Dure, który pomimo swoich straszliwych doświadczeń, nadal pozostaje człowiekiem wiary: „Ojciec Dure rozpostarł palce. – Mnie też ukrzyżowano. Sol przywołał obrazy z opowiadania Leynarta Hoyta o ojcu Dure; starszy wiekiem jezuita przybijający się do teslowego drzewa, cierpiący przez lata ból, żeby tylko nie poddać się pasożytniczemu DNA krzyżokształtu [...]. Ojciec Dure odwrócił oczy od nieba. – Nie było zaproszenia od Ojca w Niebiesiech – powiedział łagodnym głosem. – Żadnego zapewnienie, że cierpienie i ofiara są cokolwiek warte. Tylko ból. Ból, ciemność i znów ból. [...] – I to sprawiło, że straciłeś wiarę? Ojciec Dure popatrzył na Sola. – Przeciwnie. To sprawiło, że uznałem wiarę za rzecz podstawową. Ból i ciemność są naszym udziałem od upadku Adama i Ewy. Ale musi być jakaś nadzieja, że powstaniami z grzechu [...] że nasza świadomość może się rozwinąć, wznieść na wyżyny, ponad obojętność, w której zaprogramowany jest ten wszechświat”<sup>16</sup>.

<sup>16</sup> S i m m o n s, *Zagłada Hyperiona*, s. 254.

Zwróćmy uwagę zwłaszcza na ostatnie zdanie, które wprowadza sygnalizowany wcześniej motyw obojętnego wszechświata i zestawia go z nadzieją na istnienie jakiegoś głębszego wymiaru rzeczywistości, umożliwiającego transcendencję ponad tę redukcjonistyczną wizję.

Postać ojca Dure, który od tej pory stanie się jednym z kluczowych bohaterów *Zagłady Hyperiona*, nie tylko pozwala bezpośrednio nawiązać do problematyki eschatologicznej, ale także służy do wprowadzenia filozoficznych i teologicznych koncepcji francuskiego jezuitę Pierre'a Teilharda de Chardin (1881-1955)<sup>17</sup>, którego można chyba uznać za jednego z prekursorów sposobu myślenia, określonego wcześniej jako próba unifikacji rzeczywistości fizycznej oraz metafizycznej czy też światopoglądu naukowego ze światopoglądem religijnym.

Kluczowe jest tu zwłaszcza Teilhardowskie ujęcie teorii ewolucji (tak zwany ewolucjonizm chrześcijański) oraz koncepcja punktu Omega. Teilhard de Chardin uważał, że cała rzeczywistość ewoluuje od punktu Alfa do punktu Omega, mając swój początek i zarazem koniec w osobie Jezusa Chrystusa. Sugerował też powiązanie między powiększaniem się złożoności świata materii a ewolucją biologiczną. W ewolucji tej miała się ujawniać tendencja do tworzenia przez materię coraz bardziej złożonych form uporządkowania. Celem ewolucji człowieka i świata byłaby zatem ostateczna jedność duchowa wszechświata<sup>18</sup>. Bóg byłby więc w pewnym sensie efektem tej ewolucji, choć istniejąc poza czasem, stanowiłby zarazem niejako swój własny początek. Wizja ewolucji wszechświata w „Pieśniach Hyperiona” w dużej mierze nawiązuje do powyższych koncepcji.

W trakcie niezwykle skomplikowanej i wielowątkowej fabuły *Zagłady Hyperiona*, protagoniści odkrywają, iż prawdziwego wroga ludzkości stanowią nie Intruzi, lecz TechnoCentrum, które, jak zdradza jedna z najpotężniejszych, wchodzących w jego skład sztucznych inteligencji, Ummon, „pasżytuje na rodzaju ludzkim / wykorzystuje neurony każdego delikatnego umysłu w dążeniu do Najwyższego Intelaktu / konstruowaliśmy więc waszą cywilizację ostrożnie tak aby / jak chomiki w klatce [...] wasze małe mózgi myśląc, służyły naszym celom”<sup>19</sup>.

Co więcej, konflikt z Intruzami został w istocie sprowokowany przez samo TechnoCentrum – wspólnotę sztucznych inteligencji, pozornie wspierających Hegemonię. Teraz jednak dążą ona do zagłady ludzkości, która przestała być

<sup>17</sup> Zob. np. P. Teilhard de Chardin, *Moja wizja świata*, tłum. M. Tazbir, Instytut Wydawniczy PAX, Warszawa 1987; t e n ż e, *Zarys wszechświata personalistycznego i inne pisma*, tłum. M. Tazbir, K. Waloszczyk, Instytut Wydawniczy PAX, Warszawa 1985.

<sup>18</sup> Por. A. R. P e a c o c k e, *Teologia i nauki przyrodnicze*, tłum. L. M. Sokołowski, Znak, Kraków 1991, s. 129, 135, 253.

<sup>19</sup> S i m m o n s, *Zagłada Hyperiona*, s. 315.

potrzebna sztucznym inteligencjom. Tocząca się wojna jest zaś w istocie echem potężnego starcia w przyszłości (i tu powracamy do koncepcji Teilharda) między Najwyższym Intelktem (Punktem Omega-Bogiem ludzi) a Najwyższym Intelktem (Punktem Omega-Bogiem maszyn), a więc dwoma równoległymi, konkurencyjnymi wizjami ewolucji wszechświata. Zarówno Bóg ludzi, jak i Bóg maszyn, jako że są ukoronowaniem ewolucji wszechświata, potrafią działać poza czasem, wpływając na jego historię.

Bóg maszyn, zgodnie ze słowami Ummona, reprezentuje potęgę: „Nasz NI [Najwyższy Intelkt – G.T.] obraca galaktyki / pochłania kwazary będące jego źródłem energii”<sup>20</sup>, a także charakterystyczny dla świata sztucznych inteligencji w tetralogii radykalny ewolucyjny darwinizm, drapieżnictwo oraz pasożytnictwo. Jawi się on jako czysta inteligencja, uosobienie bezdusznej kalkulacji, niezłagodzonej żadną empatią. Z kolei ludzki NI, jak z pogardą wyraża się Ummon, „jest trójcą w jedności / powstała z intelektu / Współczucia / oraz Pustki Która Łączy”<sup>21</sup>. Oba NI umiejscowione są w niedostępnych bezpośrednio poznaniu człowieka, ale opisywalnych fizykalnie zakamarkach wszechświata, na przykład w. przestrzeni kwantowej: „Nasz NI wypełnia szpary rzeczywistości / dziedzicząc ten dom po jego twórcach \\ Wasz NI zdaje się lokować w polu fizycznym odkrytym przez Heisenberga i Schroedigera”<sup>22</sup>. Konflikt między oboma NI jest nieunikniony, jako że Bóg maszyn „chce wszystko uporządkować / sprowadzić do jakiegoś wzorca, na który nie mają wpływu kaprysy uczucia przypadki ani ludzka ewolucja”<sup>23</sup>. Z kolei Bóg ludzi, zdaniem Ummona, jest „ogrodnikiem porządkującym bezgraniczny wszechświat za pomocą swych topornych grabi /

daremnie śledzącym lot każdego spadającego wróbla oraz ruch każdego elektronu i pozwalającym każdej cząsteczce dowolnie poruszać się w czasoprzestrzeni, a każdemu człowiekowi badać najmniejszą szczelinę kosmicznej ironii”<sup>24</sup>. Jak więc widzimy, atrybutami ludzkiego NI są w sposób chrześcijański rozumiane miłosierdzie oraz pragnienie obdarowania istot ludzkich prawdziwą wolnością.

Walka w punkcie Omega wydaje się jednak nierozstrzygnięta, jako że, jak zdradza Ummon, „część waszego NI / ta będąca Współczuciem / wycofała się i uciekła w czasie ukrywając się pod ludzką postacią / już zresztą nie po raz pierwszy”<sup>25</sup> (to rzecz jasna aluzja do osoby Jezusa). NI sztucznych inteligencji

<sup>20</sup> Tamże, s. 316.

<sup>21</sup> Tamże, s. 318.

<sup>22</sup> Tamże.

<sup>23</sup> Tamże, s. 320.

<sup>24</sup> Tamże, s. 319.

<sup>25</sup> Tamże, s. 320.

poszukuje zatem owego wcielonego w ludzką postać współczucia, by zniszczyć je i w ten sposób zakończyć wojnę.

Drugi tom tetralogii kończy się jednak względnie optymistycznie. Ludziom i Intruzom udaje się przechytryć TechnoCentrum, wyłączając transmitery, w przestrzeni oddziaływania których sztuczne inteligencje ukrywały swe fizyczne elementy, wykorzystując jednocześnie do własnych celów moc obliczeniową mózgow ludzi podróżujących portalami. Zwycięstwo jest jednak okupione utratą komunikacji między poszczególnymi światami i rozpadem Hegemonii.

Kończąc omówienie drugiej części cyklu, wspomnijmy o jeszcze jednym ważnym motywie, jakim jest początek dziewczynki o imieniu Enea – owocu miłości kobiety Browne Lamii oraz stworzonego przez TechnoCentrum cybryda Johna Keatsa – wyposażonej w biologiczne ciało sztucznej inteligencji, dla której modelem był angielski poeta. Dziecko to, jak dowiadujemy się przy końcu powieści, stanie się w przyszłości Ta, Która Naucza – rodzajem mesjasza, w którego wcieliło się obecnie wspomniane wcześniej Współczucie, a jednocześnie osobą dysponującą dostępem do świata sztucznych inteligencji. Tak kończy się pierwsza dylogia.

Akcja drugiej, na którą składają się tomy *Endymion* oraz *Triumf Endymiona*, rozpoczyna się około trzysta lat po wydarzeniach opisanych w *Zagładzie Hyperiona*. Również charakteryzuje się ona bogatą, wielowątkową kompozycją. Tu skoncentrujemy się przede wszystkim na wątku Enei, najistotniejszym dla poruszanych zagadnień.

Pomimo pozornego zwycięstwa ludzkości pod koniec pierwszej dylogii, sztuczne inteligencje z TechnoCentrum nie tylko ocalały, ale w ciągu kilku kolejnych stuleci znacznie rozszerzyły swe wpływy. Hegemonię zastąpił tymczasem Kościół katolicki i jego zbrojne ramię, wszechpotężna organizacja Pax, która dzięki zaawansowanej technologii żelazną ręką włada dziesiątkami światów. Jak dowiadujemy się w trakcie akcji powieści, w istocie nastąpiło tutaj „wrogie przejęcie” Kościoła przez TechnoCentrum. Wkrótce po wydarzeniach z *Zagłady Hyperiona* papieżem, mającego siedzibę na planecie Pacem Kościoła obwołany został jeden z pozytywnych bohaterów powieści, ojciec Paul Dure, przybierając imię Teilharda I. Jakiś czas później został on jednak zamordowany przez ambitnego kardynała Lourdusamy, działającego w zмовie z diabolicznym wysłannikiem TechnoCentrum, radcą Albedo. Z krzyżokształtu na ciele Dure został następnie wskrzeszony Leynart Hoyt, który przyłączył się do spisku i zastąpił prawowitego papieża, przybierając imię Juliusza. Od tamtej pory Kościół, pod przywództwem cyklicznie wskrzeszanego papieża Juliusza (choć po ostatnim wskrzeszeniu przyjmuje on imię Urbana XVI) oraz jego pomocników kardynałów Lourdusamy i Mustafy, staje się jedynie marionetką w rękach TechnoCentrum. Niemal cała ludzkość przy-

jęła krzyżozształt (TechnoCentrum udoskonalilo pasożyta, tak, że na pozór nie powoduje on żadnych niepożądanych skutków ubocznych). Zapewniając ludziom rodzaj względnej nieśmiertelności (o ile krzyżozształt nie zostaje unicestwiony, co technicznie jest bardzo trudne, gdyż każdorazowo wskrzesza on swojego zmarłego żywiciela), sztuczne inteligencje przejmują nad nimi znaczącą (choć niecałkowitą) kontrolę.

Większa część historii opowiedziana zostaje ustami Raula Endymiona, młodego mężczyzny z Hyperiona. Nienależący do powtórnie narodzonych chrześcijan Endymion zostaje ma początku trzeciej części tetralogii zwerbowany przez umierającego Martina Silenusa, ostatniego żyjącego pielgrzyma do Grobowców Czasu, w celu uratowania córki Brawne Lamii i hybryda Johna Keatsa, Enei. Niemal trzysta lat wcześniej, w wieku dwunastu lat, Enea weszła do Grobowców Czasu i – jak przekonuje Silenus – ma wkrótce z nich powrócić. Grobowce Czasu są jednak pilnie strzeżone przez żołnierzy Paxu, którzy – na rozkaz papieża, a w istocie TechnoCentrum – przygotowują się do schwywania przyszłej Tej, Która Naucza.

Endymion podejmuje się pozornie niewykonalnego zadania i odtąd staje się opiekunem, a wiele lat później również uczniem i ukochanym Enei. Nie będziemy tu streszczać skomplikowanej fabuły *Endymiona*, gdzie Simmons relacjonuje ucieczkę Enei i jej przyjaciół przez dziesiątki światów przed ścigającymi ją siłami Paxu i TechnoCentrum oraz wprowadza dziesiątki nowych postaci (wymieńmy jednak kolejnego jezuitę, ojca Frederico De Soye, kapitana elitarnego statku Paxu, i jego załogę, którzy początkowo ścigają Eneę, następnie zaś przechodzą na jej stronę, a nawet ratują ją przed wysłanym przez TechnoCentrum okrutnym, nieludzkim zabójcą). Skonkludujmy zatem jedynie, że Enei udaje się dotrzeć w bezpieczne miejsce, gdzie przed rozpoczęciem swej właściwej misji będzie zgłębiać zawód architekta (kolejne odniesienie do postaci Jezusa, który przed podjęciem swego posłannictwa prawdopodobnie wykonywał zawód cieśli).

Akcja ostatniego tomu cyklu, *Triumfu Endymiona*, rozgrywa się mniej więcej dziesięć lat później i koncentruje na misji Enei oraz ostatecznej bitwie ludzkości z TechnoCentrum. Dzięki naukom Enei dostępujemy tu kolejnych stopni wtajemniczenia w ukryte mechanizmy powieściowego świata.

Można odnieść wrażenie, że autorskie koncepcje z pierwszej i drugiej dylogii różnią się nieco od siebie czy też nie są do końca kompatybilne, co być może wynika ze stopnia komplikacji i rozmiarów powieściowego projektu, nad którym niewątpliwie trudno było zapanować nawet tak doświadczonemu pisarzowi jak Simmons, a także z być może nieuniknionej ewolucji tegoż projektu. O ile więc na przykład ontologia i eschatologia w pierwszej dylogii inspirowana była głównie teologią chrześcijańską, o tyle w drugiej widoczne są wpływy buddyźmu zen, pojawia się też więcej motywów charakterystycz-

nych dla tradycyjnej science fiction. Oczywiście, wszystkie te idee stają się – jak to zazwyczaj bywa w „twardej” fantastyce naukowej – przede wszystkim budulcem wykorzystanym do stworzenia autonomicznej autorskiej kreacji wszechświata przedstawionego.

Ograniczona objętość niniejszego artykułu nie pozwala na dokładne streszczenie wszystkich filozoficznych i naukowych koncepcji zawartych w *Triumfie Endymiona* (same „wykłady” Enei i dialogi, jakie odbywa z innymi postaciami, dotyczące natury kosmosu, zajmują w powieści ponad trzydzieści stron). Skupimy się zatem na wybranych elementach ontologii powieściowego wszechświata.

Kluczową ideą jest tutaj Pustka, Która Łączy – leżąca poza dostępnym bezpośredniemu poznaniu światem fizycznym, w kwantowej przestrzeni Plancka. Jest ona wielowymiarowym medium, rządzącym się własnymi prawami i mającym własną topografię<sup>26</sup>. „Pustka, Która Łączy, jest tworem świadomym [...] jest utkana z materii kwantowej, z czasu i przestrzeni Plancka, które oplatają fizyczną czasoprzestrzeń jak powłoczka kołdrę. Nie jest w swej naturze misticzna ani metafizyczna, wypływa z fizycznych praw ewoluującego kosmosu i reaguje na nie, ale jest zarazem tegoż kosmosu produktem. Stanowi dowód samoświadomości wszechświata. Składają się na nią myśli i uczucia, choć nie chodzi tu jedynie o ludzkie myśli i uczucia: setki tysięcy rozumnych ras, jakie przez miliardy lat zamieszkiwały kosmos, miały wpływ na kształt Pustki, Która Łączy”<sup>27</sup>.

Zauważmy, że chociaż – według deklaracji Enei – Pustka „nie jest w swej naturze misticzna ani metafizyczna” i „wypływa z praw fizycznych ewoluującego kosmosu”<sup>28</sup>, to jednak ewidentnie przejmuje atrybuty przestrzeni stricte misticznej czy metafizycznej. W istocie mamy więc tu do czynienia z próbą unifikacji rzeczywistości fizycznej i metafizycznej.

Przypominając postać jednego z Pielgrzymów z pierwszej dylogii, Sola Weintraba, Enea stwierdza: „[Sol – G.T.] zrozumiał, że miłość jest rzeczywistą siłą we wszechświecie, równie prawdziwą jak elektromagnetyzm czy słabe oddziaływanie jądrowe, jak grawitacja – i że rządzą nią podobne, bardzo konkretne prawa [...]. Sol pojął, że miłość jest siłą, która spaja Pustkę, osnową jej materii. W momencie nagłego olśnienia zrozumiał również, że ludzie nie są jedynymi tkaczami tej przewspaniałej tkaniny [...] Każdy kto ma otwarte serce i umysł, może doświadczyć istnienia Pustki [...] Pustka, Która Łączy, jednocześnie jest i nie jest stanem umysłu; to prawdopodobieństwo, reprezentowane przez fale stojące i oddziałujące na front falowy, który jest ludzkim umysłem

<sup>26</sup> Por. S i m m o n s, *Triumf Endymiona*, s. 365.

<sup>27</sup> Tamże, s. 432.

<sup>28</sup> Tamże.



i osobowością. Na kształtowanie się Pustki mają wpływ wszyscy, którzy płaczą ze szczęścia, żegnają umiłowane osoby, przeżywają orgazm, stoją nad grobem ukochanego, czy patrzą, jak ich dziecko pierwszy raz otwiera oczy<sup>29</sup>. „Miłość – przekonuje Enea swego ukochanego – jest zarazem uczuciem i formą energii [...] A przy tym stanowi jedyny klucz do największych zasobów energii we wszechświecie”<sup>30</sup>. Rozumie przez to „rozpalanie kwazarów, oswajanie pulsarów, czerpanie energii z eksplodujących jąder galaktyk jak ze zwykłych parowych turbin; [...] projekt inżynierski zapoczątkowany przed dwu i pół miliardami lat, a wciąż ledwie zaczęty”<sup>31</sup>.

Zaprezentowana powyżej wizja celowego, ewoluującego, przepelnionego miłością wszechświata stoi w jaskrawym kontraście do redukcjonistycznych obrazów obojętnych uniwersów, które przeważają zarówno w świecie współczesnej nauki, jak i w inspirowanej przez nią literaturze science fiction. Znów przywołajmy „Wspomnienie o przyszłości Ziemi”, gdzie wszechświatem rządzą bezduszna kalkulacja i drapieżny darwinizm, a jakiegokolwiek porozumienie jest niemożliwe, co w końcu doprowadza zresztą do nieuniknionej zagłady całego uniwersum.

Sztuczne inteligencje TechnoCentrum zdają się jednak niezdolne do jakiegokolwiek formy empatii, przez co Pustka, Która Łączy pozostaje dla nich w ogromnej mierze zamknięta. Starają się w rabunkowy sposób eksploatować życiodajne medium, lecz go nie rozumieją. Są bezradne wobec metafizycznego czy też przynajmniej quasi-metafizycznego wymiaru kosmosu. Pustka jest dla nich „tym co piękny obraz dla ślepcy, który postanawia spalić go w piecu, żeby mu było cieplej, albo jak symfonia Beethovena dla człowieka głuchego, który odczuwa wywołane jej brzmieniem wibracje i wzmacnia konstrukcję podłogi, żeby je wytłumić”<sup>32</sup>. Kolejne stworzone przez SI technologie – napęd Hawkinga, transmitters, krzyżokształty – zniewalają ludzkość i w nieodwracalny sposób niszczą Pustkę.

Nieunikniona konfrontacja między Eneą i jej sojusznikami a TechnoCentrum oraz sterowanym przez niego fałszywym Kościołem Leynarta Hoyta może więc zostać zinterpretowana nie tylko jako starcie dobra ze złem (oba pojęcia opisywane są tutaj w kategoriach w dużej mierze fizykalnych, ale są również kompatybilne z chrześcijańskimi rozumieniami tych pojęć), lecz także jako starcie dwóch wizji – czy też dwóch sposobów widzenia wszechświata, o jakich pisaliśmy wcześniej: redukcjonistycznego (materialistycznego) oraz holistycznego, otwartego na rzeczywistość transcendentną, skoncentrowanego na poszukiwaniu wyższego porządku i sensu we wszechświecie.

<sup>29</sup> Tamże, s. 433.

<sup>30</sup> Tamże, s. 636.

<sup>31</sup> Tamże.

<sup>32</sup> Tamże, s. 437.



Koniec losów Enei zawiera liczne paralele i odniesienia do postaci Jezusa i Ewangelii (co samo w sobie bez wątpienia stanowi interesujący temat do osobnej dyskusji). DNA dziewczyny, odziedziczone po ojcu, będącym owocem eksperymentów jednej z frakcji TechnoCentrum, pragnącej otworzyć sztuczne inteligencje na empatię, ma szczególne właściwości. Dzięki niemu Enea opanowała zdolność słyszenia rozbrzmiewających w Pustce „głosów umarłych”, „głosów żywych”, a także „muzyki sfer”, co umożliwia jej pewien rodzaj „kwantowej” teleportacji.

DNA Enei może zostać przyswojone przez innych ludzi za pośrednictwem jej krwi. Efekt jego działania jest dwójaki – z jednej strony uwrażliwia ludzi na Pustkę, Która Łączy, z drugiej zaś stanowi szczepionkę przeciw krzyżokształtowi, który odpada od ciała człowieka w ciągu doby po jego kontakcie z krwią dziewczyny (lub z krwią którejkolwiek z osób, która tę krew uprzednio przyjęła). DNA Enei stanowi więc rodzaj dobrego „wirusa”, który rozprzestrzeniając się, jest w stanie potencjalnie pozbawić TechnoCentrum władzy nad ludzkością (dodajmy też, że Enea kładzie szczególny nacisk na dobrowolny i świadomy charakter decyzji każdego człowieka o przyjęciu „wirusa”).

Ostateczne zwycięstwo wymaga jednak ostatecznego poświęcenia. W momencie, gdy przeważające wojska Paxu atakują sojuszników Enei, ona sama przybywa dobrowolnie na Pacem i oddaje się w ręce wrogów. Radca Albedo i jego okrutne potwory torturują dziewczynę w obecności kardynałów Lourdusamy i Mustafy (papież Urban poddaje się w tym czasie operacji kolejnego wskrzeszenia). TechnoCentrum chce w ten sposób zmusić Eneę, by kierowana bólem i rozpaczą, dokonała w obecności jego aparatury pomiarowej „kwantowego skoku”, wyrwijąc się z jego rąk, ale jednocześnie umożliwiając zmierzenie fizycznych parametrów tego „cudu”, odtworzenie go i w konsekwencji zdobycie częściowej przynajmniej władzy nad Pustką. Enea wybiera jednak męczeńską śmierć, aby w pełni dać świadectwo prawdzie, a jej ostatnie przesłanie, pełne bólu, ale i miłości, za pośrednictwem jej uczniów oraz Pustki zostaje „przetransmitowane” do wszystkich ludzi (później zostanie ono nazwane „wspólną chwilą”). Tak oto w momencie śmierci Tej, Która Naucza rodzi się nowa wspólnota – przeciwko TechnoCentrum i jego poplecznikom wybucha powszechny bunt, miliony wiernych, dowiadując się o związkach Paxu z TechnoCentrum, odwracają się od fałszywego Kościoła, a liczne sztuczne inteligencje giną wraz z porzucanymi krzyżokształtami. Buntownicy odnajdują również Paula Dure w momencie, gdy ten odradza się z krzyżokształtu Leynarta Holta i podają mu DNA Enei. Były papież Teilhard I odzyskuje wolność i śmiertelność, Urban XVI ginie zaś wraz z przechowywanym jego osobowość pasożytem. Dzięki ofierze Enei TechnoCentrum zostaje ostatecznie pokonane, a ludzkość staje u progu nowej ery. Również Kościół zyskuje szansę na odrodzenie.

Jak wspomniano, ze względu na bogactwo i rozległość literackiej wizji Danna Simmonsa w bieżącym artykule omówiono jedynie niektóre jej aspekty. Bez wątpienia „Pieśni Hyperiona” zasługują na szerszą uwagę i dalszą, bardziej szczegółową analizę, nie tylko ze względu na zawarte w nich konteksty filozoficzne, teologiczne i naukowe, ale również wątki czysto literackie.

Trzeba też wyraźnie stwierdzić, że choć artystyczna kreacja amerykańskiego autora w dużej mierze czerpie z tradycji chrześcijańskiej, trudno jednak tu mówić o jakiegokolwiek zgodności z doktryną Kościoła. Zresztą sama dyskusja o takiej zgodności bądź jej braku byłaby pozbawiona sensu. Ostatecznie mamy do czynienia jedynie – i zarazem aż – ze śmiałą literacką ekstrapolacją, przeprowadzoną zgodnie z najlepszymi tradycjami spekulatywnej science fiction. Warto zarazem zwrócić uwagę, że na tle większości podobnych ze względu na rozmach i ambicje dzieł literatury spekulatywnej – i szerzej: na tle wytworów współczesnej popkultury – wizję Simmonsa cechuje niezwykle poważne podejście do teologii chrześcijańskiej i życzliwość wobec tradycji Kościoła. Na przykład liderzy fałszywego, manipulowanego przez TechnoCentrum Kościoła: Leynart Hoyt, kardynałowie Lourdusamy i Mustafa, należą do czołowych antagonistów, przeciwstawione im zostały natomiast wzbudzające sympatię postaci Paula Dure, kapitana De Soyi czy ojca Glaucusa. Szacunek do tradycji wyraża się też w scenie ślubu Enei – prekursorki całkiem nowej ery – która zawiera wraz ze swym ukochanym Raulem Endymionem sakrament małżeństwa w obecności ojca De Soi.

„Pieśni Hyperiona” wyróżniają się na tle ponurych, redukcjonistycznych obrazów obojętnych uniwersów, coraz częściej podsuwanych przez współczesną naukę i literaturę science fiction, przede wszystkim ontologicznym i eschatologicznym optymizmem. Ich obecność wskazuje, że nawet w erze kwantów możliwe jest snucie przekonujących wizji, które wyrażają nadzieję co do przyszłości człowieka i jego roli we wszechświecie oraz niosą pocieszenie.

#### BIBLIOGRAFIA / BIBLIOGRAPHY

- Crick, Francis. *Zdumiewająca hipoteza, czyli nauka w poszukiwaniu duszy*. Translated by Barbara Chacińska-Abramowicz and Michał Abramowicz. Warszawa: Prószyński i S-ka, 1997.
- Chaucer, Geoffrey. *Opowieści kanterberyjskie*. Translated by J. Zawadzki. Katowice: Biblioteka Śląska, 2022.
- Dawkins, Richard. *Samolubny gen*. Translated by Marek Skoneczny. Warszawa: Prószyński i S-ka, 2012.
- Dragan, Andrzej. *Kwantechizm 2.0, czyli klatka na ludzi*. Kraków: Otwarte, 2022.

- Ferguson, Kitty. *Ogień w równaniach: Nauka, religia i poszukiwania Boga*. Translated by Piotr Amsterdamski. Kraków: Copernicus Center Press, 2016.
- Greene, Brian. *Do końca czasu: Umysł, materia i nasze poszukiwanie sensu w zmieniającym się wszechświecie*. Translated by Tomasz Krzysztóń. Warszawa: Prószyński i S-ka, 2021.
- Heller, Michał. *Czy fizyka jest nauką humanistyczną?* Kraków: Copernicus Center Press, 2016.
- . *Filozofia przypadku: Kosmiczna fuga z preludium i codą*. Kraków: Copernicus Center Press, 2012.
- Keats, John. *33 wiersze*. Wybór, przekład i opracowanie Stanisław Barańczak. Kraków: Wydawnictwo Znak, 2007.
- Liu, Cixin. *Ciemny las*. Translated by Andrzej Jankowski. Poznań: Rebis, 2017.
- . *Koniec śmierci*. Translated by Andrzej Jankowski. Poznań: Rebis, 2018.
- . *Problem trzech ciał*. Translated by Andrzej Jankowski. Poznań: Rebis, 2016.
- Majka, Paweł. *Niebiańskie pastwiska*. Poznań: Rebis, 2015.
- McDevitt, Jack. *The Engines of God*. New York: Ace Books, 1994.
- . *Cauldron*. New York: Ace Books, 2007.
- Podlewski, Marcin. *Głębia*. T. 1-4. Lublin: Fabryka Słów, 2015-2018.
- Simmons, Dann. *Endymion*. Translated by Wojciech Szypuła. Warszawa: Amber, 1998.
- . *Hyperion*. Vols. 1–2. Translated by Arkadiusz Nakoniecznik. Warszawa: Amber, 1994.
- . *Triumf Endymiona*. Translated by Wojciech Szypuła. Warszawa: Amber, 1999.
- . *Zagłada Hyperiona*. Translated by Radosław Januszewski and Marek Rudnik. Warszawa: Amber, 1997.
- Reynolds, Alastair. *Absolution Gap*. London: Gollancz, 2003.
- . *Inhibitor Phase*. London: Gollancz, 2021.
- . *Redemption Ark*. London: Gollancz, 2002.
- . *Revelation Space*. London: Gollancz, 2000.
- Peacocke, Arthur R. *Teologia i nauki przyrodnicze*. Translated by Leszek M. Sokółowski. Kraków: Znak, 1991.
- Teilhard de Chardin, Pierre. *Moja wizja świata*. Translated by Mieczysław Tazbir. Warszawa: Instytut Wydawniczy PAX, 1987.
- . *Zarys wszechświata personalistycznego i inne pisma*. Translated by M. Tazbir and Konrad Waloszczyk. Warszawa: Instytut Wydawniczy PAX, 1985.
- Trębicki, Grzegorz. “Cixin Liu’s The Remembrance of Earth’s Past Trilogy: In Search Of Interpretative Grounds.” In *W kręgu zagadnień literatury powszechnej*. Edited by Robert Kotowski and Iwona Mityk. Kielce: Muzeum Narodowe w Kielcach, 2019.
- . Kres wszystkich rzeczy: “Wspomnienie o przeszłości ziemi” Cixina Liu vs. „Cykl Hyperioński” Dana Simmonsa (unpublished).

- Vinge, Vernon. *Ogień nad otchłanią*. Translated by Rafał Wikoński. Warszawa: Prószyński i S-ka, 1998.
- Wright, John W. *The Golden Age*. New York: TOR Books, 2002
- . *The Golden Transcendence*. New York: TOR Books, 2003.
- . *The Phoenix Exultant*. New York: TOR Books, 2003.

## ABSTRAKT / ABSTRACT

Grzegorz TRĘBICKI – Mariaż dwóch światów. Quasi-chrześcijańska wizja ewolucji człowieka i kosmosu w „Pieśniach Hyperiona” Dana Simmonsa

DOI 10.12887/36-2023-4-144-09

„Pieśni Hyperiona” uznanego amerykańskiego pisarza Dana Simmonsa to jedno z najwybitniejszych osiągnięć „twardej” science fiction, opublikowanych w ostatnich dziesięcioleciach. Stanowią one zakrojoną na isticie epicką skalę opowieść o walce garstki bohaterów ze sztucznymi inteligencjami TechnoCentrum o przyszły kształt wszechświata. Zarazem jest to jedno ze stosunkowo nielicznych dzieł SF podejmujących próbę swoistej fuzji rzeczywistości fizycznej i metafizycznej, w znaczącym stopniu posługując się przy tym elementami filozofii chrześcijańskiej. Niniejszy artykuł stanowi krótką analizę powyższych motywów.

Słowa kluczowe: literatura fantastyczno-naukowa, sztuczne inteligencje, Dan Simmons, cykl hyperioński, eschatologia, ontologia, fizyka kwantowa, ewolucja wszechświata

Kontakt: Zakład Literatur Obcych, Instytut Literaturoznawstwa, Uniwersytet Jana Kochanowskiego, ul. Uniwersytecka 17, 25-405 Kielce  
E-mail: grzegorz.trebicki@ujk.edu.pl  
ORCID: 0000-0001-6740-8922

Grzegorz TRĘBICKI, The Marriage of Two Worlds: A Quasi-Christian Vision of the Evolution of the Universe and Humanity in Dan Simmons’ “Hyperion Cantos”

DOI 10.12887/36-2023-4-144-09

“The Hyperion Cantos” by a well-known American science fiction writer Dan Simmons is one of the most outstanding achievements of hard SF in the last decades. It is an epic story of the struggle of a handful of protagonists against artificial intelligences of the TechnoCore for the future shape of the universe. At the same time, it is one of the few SF works attempting to unite the physical and metaphysical realities with the help of the elements of Christian philosophy. The present article briefly analyzes the above motifs.

Keywords: science fiction, artificial intelligences, Dan Simmons, Hyperion cycle, eschatology, ontology, quantum physics, evolution of the universe

Contact: Zakład Literatur Obcych, Instytut Literaturoznawstwa, Uniwersytet Jana Kochanowskiego, ul. Uniwersytecka 17, 25-405 Kielce, Poland

E-mail: [grzegorz.trebicki@ujk.edu.pl](mailto:grzegorz.trebicki@ujk.edu.pl)

ORCID: 0000-0001-6740-8922



NADZIEJE I OBAWY





Bartłomiej KNOSALA

## ZARZĄDZANIE ŚRODOWISKIEM NATURALNYM PRZEZ SZTUCZNĄ INTELIGENCJĘ Ograniczenia i wyzwania narracji postnaturny

*Koncepcyjny projekt autonomicznego systemu infrastruktury krajobrazowej, czyli propozycja włączenia systemów sztucznej inteligencji w proces zarządzania środowiskiem naturalnym, przedstawiony przez Bradley'a Cantrella, Erle'a Ellisa i Laure Martin na łamach czasopisma „Trends in Ecology and Evolution” w roku 2017, stanowi dobry przykład technologii, która znajdując się w fazie projektowej, podlega próbom przypisania jej konkretnych znaczeń, które mogą decydować o jej dalszych losach. Odślonięcie znaczeń i koncepcji, w ramach których propozycja ta jest umiejscawiana, może przyczynić się do zwiększenia transparentności potencjalnej debaty publicznej zogniskowanej wokół tej technologii.*

Punktem wyjścia prezentowanych w niniejszym artykule rozważań jest przekonanie o istotnej roli, jaką pełnią narracje, idee i koncepcje w procesie kształtowania świata wokół nas. Podążając za intuicjami takich autorów, jak Ernst Cassirer, Charles Taylor czy Paul Ricoeur<sup>1</sup>, przyjmujemy, że konstrukcja świata symbolicznego oraz intersubiektywne sieci znaczeń określają formy naszego działania i wyznaczają kierunki przemian naszych społeczeństw. Z kolei za Arminem Grunwaldem rozszerzamy pole analiz na narracje związane z technologiami przyszłość. Według Grunwalda jednym z głównych czynników decydujących o rozwoju danej technologii są związane z nią narracje – od tego, jakie znaczenia zostaną przypisane do NEST (ang. New and Emerging Science and Technology), może ostatecznie zależeć rozpoczęcie prac badawczych nad daną technologią lub ich zatrzymanie.<sup>2</sup>

Koncepcyjny projekt autonomicznego systemu infrastruktury krajobrazowej, czyli propozycja włączenia systemów sztucznej inteligencji w proces zarządzania środowiskiem naturalnym, przedstawiony przez Bradley'a Cantrella, Erle'a Ellisa i Laure Martin na łamach czasopisma „Trends in Ecology and Evolution” w roku 2017<sup>3</sup>, stanowi dobry przykład technologii, która znajdując się w fazie projektowej, podlega próbom przypisania jej konkretnych znaczeń,

<sup>1</sup> Por. P. Brooks, *Seduced by Story: The Use and Abuse of Narrative*, The New York Review of Books, New York 2022, s. 10.

<sup>2</sup> Por. A. Grunwald, *The Hermeneutical Side of Responsible Research and Innovation*, Iste Ltd, London, 2016, s. 4.

<sup>3</sup> Por. B. Cantrell, L. Martin, E. Ellis, *Designing Autonomy: Opportunities for New Wildness in the Anthropocene*, „Trends in Ecology & Evolution” 32 (2017), nr 3, s. 156-166.

które mogą decydować o jej dalszych losach. Odsłonięcie znaczeń i koncepcji, w ramach których propozycja ta jest umiejscawiana, może przyczynić się do zwiększenia transparentności potencjalnej debaty publicznej zogniskowanej wokół tej technologii.

Warto zauważyć, że w omawianej w niniejszym artykule propozycji włączenia systemów sztucznej inteligencji w proces zarządzania środowiskiem naturalnym krzyżują się ze sobą dwie wpływowe w kulturze zachodniej narracje. Pierwsza dotyczy relacji, jaka zachodzi między naturą i kulturą, druga związana jest z szeroko pojętym techno- optymizmem – z wiarą w zbawczą moc technologii, z oświeceniową ideą postępu oraz z przekonaniem o możliwości nieograniczonej kontroli środowiska, którą zapewnia nauka i technologia. Jednak tym, co odróżnia koncepcyjny projekt autonomicznego systemu infrastruktury krajobrazowej od wielu innych propozycji wiążących optymistyczną perspektywę ekologiczną z rozwojem nauki i technologii<sup>4</sup>, jest uznanie koncepcji postnaturny jako elementu strategii perswazyjnej. Warto zauważyć, że w dotychczasowej refleksji środowiskowej obecny był pogląd, że za kryzys ekologiczny odpowiada koniunkcja racjonalizmu (rozumianego jako bezkrytyczny kult rozumu) oraz dualizmu natury i kultury<sup>5</sup>. Wprowadzenie w kontekst nowych technologii koncepcji postnaturny, która problematyzuje dualizm natury i kultury, stawia nowe wyzwania przed refleksją środowiskową. Jedno z tych wyzwań z pewnością będzie związane z krytycznym namysłem nad nowymi powiązaniem powstającymi na styku ingerencji technologicznych i sieci ekosystemów.

## OD ANTROPOCENU DO POSTNATURNY

Od roku 2000 dyskusja dotycząca relacji między człowiekiem a środowiskiem naturalnym związana jest z koncepcją antropocenu. Antropocen – termin zaproponowany przez holenderskiego chemika atmosfery Paula Crutzena w roku 2000 – jest propozycją wprowadzenia nowej nazwy na trwającą obecnie epokę geologiczną. Według Crutzena postępująca w szybkim tempie urbanizacja świata, wyczerpywanie się paliw kopalnych, zanieczyszczenia środowiska oraz emisja gazów cieplarnianych sprawiają, że możliwe staje się uaktualnienie periodyzacji na geologicznej osi czasu. Inni badacze. działający w paradygmacie nauki o systemach planetarnych. podkreślają bezpreceden-

<sup>4</sup> Koncepcją wiążącą postęp technologiczny z nadzieją na ekologiczną przyszłość są na przykład propozycje Amory Lovinsa (zob. A. L o v i n s, P. H a w k e n, *Natural Capitalism: The Next Industrial Revolution*, Earthscan, London 1999).

<sup>5</sup> Por. V. P l u m w o o d, *Environmental Culture. The Ecological Crisis of Reason*, Routledge, London 2002, s. 4.

sowy charakter oddziaływania gatunku *homo sapiens* na systemy planetarne. Ludzkość jednocześnie wpływa na hydrosferę poprzez zakwaszanie oceanów, litosferę poprzez wyjaławianie gleb, atmosferę poprzez emisję dwutlenku węgla, biosferę poprzez tak zwane wielkie szóste wymieranie<sup>6</sup>. W związku z tym, że charakter wpływu ludzkości na systemy planetarne jest w wyraźny sposób determinowany kształtem cywilizacji oraz że mamy tu do czynienia z systemami wzajemnie powiązаныmi, między którymi dochodzi do różnorodnych sprzężeń zwrotnych, natomiast sama natura w takim ujęciu uzyskuje jakiś rodzaj sprawczości, część autorów uznaje, że epoka antropocenu jest jednocześnie epoką postnaturalną.

#### NARRACJA POSTNATURY NADZIEJE I OGRANICZENIA

Warto zaznaczyć, że w obrębie dyskursu o antropocenie funkcjonuje kilka narracji. Christophe Bonneuil w artykule *The Geological Turn. Narratives of the Anthropocene* wymienia cztery wielkie narracje (ang. Grand Narratives): narracja naturalistyczna, narracja postnaturalna, narracja eko-katastroficzna i narracja eko-marksistowska. Podział ten utrzymują inni autorzy, między innymi Ewa Bińczyk w książce *Epoka człowieka. Retoryka i marazm antropocenu*.<sup>7</sup>

Źródła narracji postnatURY mają charakter heterogeniczny. Zarówno filozofki i filozofowie kojarzeni z postmodernizmem i eko-konstrukttywizmem – Donna Haraway<sup>8</sup>, Bruno Latour<sup>9</sup> i Isabelle Stengers<sup>10</sup>, jak i naturaliści (przedstawiciele przyrodoznawstwa) oraz techno-utopijne ośrodki analityczne (na przykład Breakthrough Institute) ogłaszają nadejście antropocenu jako epokę postnatURY<sup>11</sup>. Krytykę koncepcji dzikiej natury (krytyka taka stanowi istotny składnik postnatURY) możemy znaleźć również u aktywistów domagających się praw dla ludów rdzennych oraz zorientowanych postmodernistycznie przedstawi-

<sup>6</sup> Por. E. Bińczyk, *Epoka człowieka. Retoryka i marazm antropocenu*, Wydawnictwo Naukowe PWN, Warszawa 2018, s. 90.

<sup>7</sup> Por. tamże, s. 281.

<sup>8</sup> Zob. D. Haraway, *Staying with the Trouble: Making Kin in the Chthulucene*, Duke University Press, Durham–London, 2016.

<sup>9</sup> Zob. B. Latour, *Facing Gaia: Eight Lectures on the New Climatic Regime*, Polity, Cambridge 2017.

<sup>10</sup> Por. I. Stengers, *Accepting the Reality of Gaia. A Fundamental Shift?*, w: *The Anthropocene and the Global Environmental Crisis: Rethinking Modernity in a New Epoch*, red. C. Hamilton, C. Bonneuil, F. Gemenne, Routledge, London 2015, s.134-144.

<sup>11</sup> Por. C. Bonneuil, *The Geological Turn: Narratives of the Anthropocene*, w: *The Anthropocene and the Global Environmental Crisis: Rethinking Modernity in a New Epoch*, s. 17-30.

cieli nauk społecznych. Wśród nich warto wymienić Ramachandra Guha, Williama Cronona i Philipe Descola<sup>12</sup>. Jednak, jak pisze Christophe Bonneuil, narracja postnatury jest już obecna w refleksji środowiskowej głównego nurtu<sup>13</sup>.

Cechą charakterystyczną narracji postnatury jest negacja możliwości utrzymania charakterystycznego dla nowoczesności podziału na naturę i kulturę. „Podczas gdy nowoczesność obiecywała wyzwolenie społeczeństwa z determinizmu natury, antropocen ogłasza nieuchronny związek między losem ludzkości a wielkimi ziemskimi cyklami oraz złączenie dwóch odmiennych perspektyw czasowych, które przez ostatnie dwa stulecia były postrzegane jako oddzielne – krótkookresowej perspektywy ludzkiej historii oraz długo okresowej historii Ziemi”<sup>14</sup>. Takie ujęcie relacji między naturą i kulturą nieuchronnie wpływa na cały system wiedzy, który oparty jest na jakościowej odrębności nauk społecznych i humanistycznych z jednej strony i przyrodoznawstwa z drugiej. Można powiedzieć, że narracja postnatury przypomina koncepcję natury jako sieci, którą na początku dziewiętnastego wieku wypracował Aleksander von Humboldt. Dla Humboldta natura była niezwykle czułą siecią, w której zerwanie jednej nici mogło spowodować sprucie całej tkaniny. W narracji postnatury na tę sieć wpływają również wyznawane przez nas wartości i koncepcje, zgodnie z którymi porządkujemy rzeczywistość. Naukowcy z obszaru nauk o systemach planetarnych (ang. Earth System Science) jako źródło globalnego ocieplenia wskazują przede wszystkim nadmierną emisję dwutlenku węgla oraz masowe wylesianie, jednak szukając głębszych przyczyn, można również wskazać na koncepcję natury w naukach przyrodniczych<sup>15</sup>, system ekonomiczno-społeczny, dla którego głównym wskaźnikiem jest nieskończony wzrost PKB<sup>16</sup> czy charakterystyczne dla niektórych odłamów protestantyzmu przekonanie o konieczności panowania nad naturą<sup>17</sup>.

<sup>12</sup> Por. G. R a m a c h a n d r a, „Radical American Environmentalism and Wilderness Preservation: A Third World Critique” (1997), w: *The Future of Nature: Documents of Global Change*, red. L. Robin, S. Sörlin, P. Warde, Yale University Press, New Haven 2013, s. 409-432; W. C r o n o n, *The Trouble with Wilderness: Or Getting Back to the Wrong Nature*, w: *Uncommon Ground: Rethinking the Human Place in Nature*, Norton, New York 1996, s. 69-80; P. D e s c o l a, *Beyond Nature and Culture*, University of Chicago Press, Chicago 2013.

<sup>13</sup> Por. B o n n e u i l, *The Geological Turn: Narratives of the Anthropocene*, s. 24.

<sup>14</sup> Tamże.

<sup>15</sup> Por. Z. W r ó b l e w s k i, *Natura i cele. Dyskusja argumentu teleologicznego na rzecz ochrony przyrody*, Wydawnictwo KUL, Lublin 2010, s. 7.

<sup>16</sup> Por. T. J a c k s o n, *Dobrobyt bez wzrostu. Ekonomia dla planety o ograniczonych zasobach*, tłum. M. Polkowski, Wydawnictwo UMK, Toruń 2015.

<sup>17</sup> Warto zauważyć, że w refleksji środowiskowej od momentu opublikowania w 1967 roku słynnego artykułu Lynne’a W h i t e zatytułowanego *The Historical Roots of Our Ecologic Crisis* toczy się spór o rolę chrześcijaństwa w kwestii ekologicznej. Stanowiska w tym sporze są zróżnicowane, a sam spór jest niezwykle dynamiczny. Jak zauważa Ryszard F. Sadowski, istnieje ponad dwieście publikacji odnoszących się do tego tematu (por. R. F. S a d o w s k i, *Filozoficzny spór*

W tym kontekście można utrzymywać, że to, co obecnie nazywamy systemem przyrodniczym, jest ściśle związane z sferą wartości. Z perspektywy nauk humanistycznych koncepcja postprzyrody stanowi szansę na włączenie się przedstawicieli tych nauk w dyskusję o tym, w jaki sposób zapewnić stabilną przyszłość naszej planecie. Carolyn Merchant w pracy *The Anthropocene and the Humanities*<sup>18</sup> wskazuje na te obszary humanistyki, które mogą dostarczyć nowych i pluralistycznych sposobów rozumienia sieci zależności między naturą i kulturą – obecny w wielu wyznaniach nurt „religii i ekologii”<sup>19</sup>, etyka partnerstwa między ludźmi i nieludzką naturą czy filozoficzne dociekania dotyczące nieliniarnych i chaotycznych systemów oparte na pracach Edwarda Lorentza czy Ilya Prigogine’a<sup>20</sup> to tylko niektóre propozycje. Merchant postulując przejście od antropocenu do epoki zrównoważonej (ang. Age of Sustainability), przekonuje, że „przemiana ta będzie legitymizowana przez zmiany w wartościach oraz w sposobach postrzegania, poznawania i doświadczania rzeczywistości przez ludzi”<sup>21</sup>. Warto jednak zauważyć, że koncepcja postnatu-ry niesie ze sobą aspekty, które można uznać za problematyczne w kontekście strategii zabezpieczania stabilnej przyszłości naszej planety. Pierwszy związany jest z osłabianiem politycznego przesłania ruchów ekologicznych, które koncentrują się na ochronie nietkniętej ludzką ręką natury. Zgodnie z narracją postnatu-ry ludzkość z takim bytem jak „dzika natura” nigdy nie miała do czynienia. Jak pisze Ursula K. Heise, „antropologowie środowiska, geografowie i historycy wykazali, że w Australii, Ameryce Łacińskiej i Ameryce Północnej krajobrazy, które europejscy osadnicy postrzegali jako nietknięte przez ludzi «pustkowia», w rzeczywistości były zmieniane przez rdzenne społeczności przez tysiąclecia przed ich przybyciem”<sup>22</sup>. Oznacza to, że oddolne ruchy społeczne, które od lat sześćdziesiątych wpływały na polityczne i prawne odniesienia do świata natury oraz były odpowiedzialne za daleko idące przemiany kulturowe, mogą wytracić impet społecznego oddziaływania. Jak zauważa Heise, „zmiana głównych linii narracyjnych i podstawowych konceptów, które

---

*o rolę chrześcijaństwa w kwestii ekologicznej*, Towarzystwo Naukowe Franciszka Salezego, Warszawa 2015, s. 103). Z kolei Wojciech Bołoz uważa, że twierdzenie, iż chrześcijaństwo jest winne eksploatacji przyrody może ewentualnie dotyczyć „chrześcijaństwa w wydaniu protestanckim” (W. B o ł o z, *Kościół i ekologia. W obronie człowieka i środowiska naturalnego*, Wydawnictwo Homo Dei, Kraków 2010, s.12).

<sup>18</sup> Zob. C. M e r c h a n t, *The Anthropocene and the Humanities: From Climate Change to a New Age of Sustainability*, Yale University Press, New Heaven and London, 2020.

<sup>19</sup> Tamże, s. 90-196.

<sup>20</sup> Por. tamże, s. 123.

<sup>21</sup> Tamże, s. 153.

<sup>22</sup> U. H e i s e, *Imagining Extinction: The Cultural Meanings of Endangered Species*, University of Chicago Press, Chicago 2016, s. 9. Jeśli nie podano inaczej, tłumaczenie fragmentów obcojęzycznych – B.K.

przez tak długi okres czasu mobilizowały społeczeństwa do działania niemal na wszystkich kontynentach, niesie ryzyko osłabienia ekologicznej kultury i ograniczeniem politycznych wpływów ruchów ekologicznych<sup>23</sup>. Rezygnując z pojęcia natury, narażamy zdolność mobilizacji szerokich mas społecznych do działań na rzecz kwestii ekologicznej.

#### KONCEPCJA POSTNATURY JAKO STRATEGIA PERSWAZYJNA

Drugi aspekt związany jest z wykorzystywaniem koncepcji postnatyry jako strategii perswazyjnej usprawiedliwiającej coraz powszechniejsze stosowanie systemów technologicznych w kluczowych aspekty funkcjonowania naszej planety. Jeśli pojęcie natury nie stanowi już punktu odniesienia naszych działań, zwolennicy interwencji technologicznych mających na celu na przykład schładzanie klimatu na drodze geoinżynierii czy zabezpieczanie ekosystemów poprzez stosowanie sztucznej inteligencji zyskują nową linię argumentacyjną – skoro nie istnieje żaden konkretny stan sprzed interwencji człowieka, który stanowiłby punkt odniesienia, i do przywrócenia którego można by dążyć, otworzenie się na nowe technologie i gotowość do coraz głębszych interwencji technologicznych w środowisko naturalne pozwala na wytyczenie ścieżek prowadzących ku tak zwanemu dobremu antropocenowi<sup>24</sup>. W manifestie ekomodernistycznym wydanym przez Breakthrough Institute dobry antropocen to etap, który ludzkość może osiągnąć dzięki mądrym stosowaniu wiedzy i technologii. Jednak dobry antropocen nie oznacza osiągnięcia stanu harmonijnych relacji między ludźmi i naturą. Autorzy manifestu odrzucają przekonanie zgodnie z którym społeczności ludzkie muszą żyć w harmonii z naturą, aby uniknąć upadku ekonomicznego i katastrofy ekologicznej<sup>25</sup>. Wręcz przeciwnie, podkreślają oni konieczność wzmocnienia rozwoju technologicznego, aby uniezależnić ludzkość od dobrostanu przyrody. W tym celu zachęcają do podjęcia intensywnych wysiłków na rzecz rozwijania takich technologii jak: genetyczna modyfikacja roślin, technologie jądrowe czy modyfikowanie klimatu przy pomocy geoinżynierii<sup>26</sup>. Według autorów manifestu ekomodernistycznego wraz z oddzieleniem ludzkiego dobrobytu od wpływu na środowisko przyroda będzie kształtowana zgodnie z wytycznymi kulturowymi. „Wszędzie tam, gdzie

<sup>23</sup> Tamże, s. 11.

<sup>24</sup> Argumentację taką prezentuje manifest ekomodernistyczny (zob. *Manifest modernistyczny*, <http://www.ecomodernism.org/jzyk-polski>).

<sup>25</sup> Zob. tamże.

<sup>26</sup> Zob. M. S h e l l e n b e r g e r, T. N o r d h a u s, *On Becoming an Ecomodernist: A Positive Vision of Our Environmental Future*, <https://thebreakthrough.org/articles/on-becoming-an-ecomodernist>.



uniezależnienie rozwoju człowieka od przyrody zmniejsza presję wywieraną na krajobraz i ekosystemy w celu zaspokojenia podstawowych potrzeb ludzkich, rolnicy, społeczności i rządy w dalszym ciągu będą musiały zdecydować, jakie założenia estetyczne lub ekonomiczne mają ten krajobraz kształtować<sup>27</sup>. Warto jednak zauważyć, że podstawowe idee manifestu ekomodernistycznego spotkały się z wieloraką krytyką. Z perspektywy oceny technologii (ang. Technology Assessment) wskazuje się na ignorowanie ambiwalencji technologii oraz pomijanie problemu niezamierzonych konsekwencji wprowadzania innowacji technologicznych<sup>28</sup>. Z perspektywy krytyki nowoczesności, zwraca się z kolei uwagę, że prezentowana w manifestie czysto instrumentalna koncepcja natury w połączeniu z wiarą w zbawczą moc technologii jest podstawową przyczyną destrukcji środowiska. Trudno zatem spodziewać się, że taka postawa będzie jednocześnie rozwiązaniem problemów środowiskowych<sup>29</sup>. Również z perspektywy analizy dyskursów antropocenu podkreśla się, że niektóre rozwiązania promowane przez Breakthrough Institute, takie jak na przykład geoinżynieria, mogą działać demobilizacyjnie, gdyż samo snucie takich wizji odsuwa nas od wciąż jeszcze możliwych do zrealizowania, instytucjonalnych, transnarodowych rozwiązań wspierających polityką niskoemisyjną<sup>30</sup>. Badacze związani z ekonomią postwzrostu także przedstawiają cały szereg zarzutów – od pomijania niewygodnych danych i zjawisk (na przykład koszty środowiskowe konwencjonalnego, wielkoskalowego rolnictwa czy skala przełowienia oceanów), po ignorowanie podstawowej wiedzy z zakresu ekologii i termodynamiki mówiącej, że gatunki i społeczeństwa posiadają granice wzrostu<sup>31</sup>. Anne Fermaux zauważa, że prezentowana przez ekomodernistów koncepcja natury nawet w perspektywie filozofii postmodernistycznej też może budzić wątpliwości. Chociaż w manifestie ekomodernistów koncepcja nieskażonej ludzką ręką natury jest mocno krytykowana, to nie pojawia się w zamian ważna dla filozofii postmodernistycznej koncepcja „alternatywnych natur”, umożliwiająca na przykład demokratyczną debatę na temat środowiskowej sprawiedliwości. Przeciwnie, porządek społeczno-ekologiczny postulowany przez eko-modernistów miałby być zdaniem Fernaux antydemokratyczny i technokratyczny (to naukowcy mieliby decydować, które ustalenia dotyczą-

<sup>27</sup> *Manifest ekomodernistyczny*.

<sup>28</sup> Por. A. Grunwald, *Diverging Pathways to Overcoming the Environmental Crisis: A Critique of Ecomodernism from Technology Assessment Perspective*, „Journal of Cleaner Production” 197(2018) nr 2, s. 1854-1862.

<sup>29</sup> Por. Grunwald, dz. cyt., s. 164.

<sup>30</sup> Por. E. Bieńczyk, *Inżynieria klimatu a inżynieria człowieka. Dyskursy na temat środowiska w epoce antropocenu*, „Ethos” 28(2015) nr 3(111), s. 171.

<sup>31</sup> Zob. J. Caradonna in., *An Ecomodernist Manifesto: A Degrowth Critique*, [http://www.resilience.org/wp-content/uploads/articles/General/2015/05\\_May/A-Degrowth-Response-to-An-Ecomodernist-Manifesto.pdf](http://www.resilience.org/wp-content/uploads/articles/General/2015/05_May/A-Degrowth-Response-to-An-Ecomodernist-Manifesto.pdf).

ce natury będą realizowane), zarazem konserwatywny (uprzywilejowujący obecny, kapitalistyczny sposób zarządzania naturą) oraz uniwersalistyczny<sup>32</sup>.

## SZTUCZNA INTELIGENCJA I AUTOMATYCZNE ZARZĄDZANIE ŚRODOWISKIEM NATURALNYM

Mimo wzmożonej krytyki sygnatariusze manifestu ekomodernistycznego nie tylko nie zrewidowali swoich podstawowych tez promujących wzmocnienie interwencji naukowo-technicznych na globalną skalę, ale również przedstawili nową propozycję poszerzenia zakresu stosowania infrastruktury technologicznej na obszar bioróżnorodności. W artykule z 2017 roku zatytułowanym *Designing Autonomy: Opportunities for New Wildness in the Anthropocene* Erle Ellis (sygnatariusz manifestu ekomodernistycznego), Bradley Cantrell i Laura J. Martin proponują delegowanie zarządzania środowiskiem naturalnym systemom sztucznej inteligencji. Warto zwrócić uwagę, że główna oś argumentacyjna artykułu związana jest właśnie z koncepcją postnatury. Punktem wyjścia prowadzonych rozważań jest przekonanie, że na obecnym etapie rozwoju cywilizacyjnego wycofywanie interwencji człowieka z obszarów przyrodniczych, tak by odzyskać tak zwaną dziką naturę, byłoby praktyką nieodpowiedzialną. Ich zdaniem nawet w chronionych obszarach dzikiej przyrody wyzwaniem stało się utrzymanie wzorców i procesów ekologicznych bez coraz częstszych i intensywniejszych interwencji zarządczych, w tym kontroli gatunków inwazyjnych, zarządzania zagrożonymi populacjami i remediacji zanieczyszczeń<sup>33</sup>. Warto zauważyć, że w refleksji środowiskowej stanowisko mówiące o konieczności odejścia od modelu ochrony przyrody (ang. preservation) na rzecz restaurowania (ang. restoration) jest szeroko podzielane. Na przykład Bińczyk pisze o statusie natury w antropocenie: „Ochrona całkowicie dziewiczych, nieskażonych fragmentów środowiska naturalnego to obecnie postulat niemożliwy już do zrealizowania. [...] Z tego powodu krytykuje się paradygmat ochrony przyrody jako przestarzały, głosząc potrzebę restaurowania – rekompensowania planecie szkód wyrządzonych ręką ludzką. Chodzi tu o odnawianie ekosystemów rozumiane na różne sposoby: jako planetarne ogrodnictwo, rodzaj designu, inżynierii czy zarządzania”<sup>34</sup>.

<sup>32</sup> Zob. A. F e r m a u x, *The Return of Nature in the Capitococene*, w: *Rethinking the Environment for the Anthropocene: Political Theory and Sociocultural Relations in the New Geological Epoch*, red. M. Aries-Maldonado, Z. Trachtenberg, Routledge, New York 2019, s. 19-37.

<sup>33</sup> Zob. B. C a n t r e l l, L. M a r t i n, E. E l l i s, *Designing Autonomy: Opportunities for New Wildness in the Anthropocene*, „Trends in Ecology & Evolution” 32(2017) nr 3, s. 156-166.

<sup>34</sup> Por. B i Ń c z y k, dz. cyt., s. 121.

Skoro iluzją jest autonomiczne funkcjonowanie dzikiej przyrody, a interwencje ludzkie w podtrzymywanie „dzikości” są normą, to zdaniem autorów omawianego artykułu otwiera to drogę do postawienia pytania o zarządzanie środowiskiem naturalnym za pomocą sztucznej inteligencji. Zadaniem sztucznej inteligencji byłoby podtrzymywanie autonomii nie-ludzkich procesów ekologicznych, która w tym zakresie miałaby zastąpić bezpośrednią interwencję człowieka.

Nie tylko jednak charakterystyczne dla antropocenu pogłębiające się zależności między ludźmi i ekosystemami każą myśleć o projektowaniu środowiska, które miałyby funkcjonować w oparciu o sztuczną inteligencję. Drugim powodem, który podnoszą autorzy omawianego artykułu, jest szybki rozwój systemów opartych na tak zwanym głębokim uczeniu przez wzmacnianie (ang. Deep Reinforcement Learning, DRL). Uczenie przez wzmacnianie – obok uczenia nadzorowanego i uczenia nienadzorowanego – jest jednym z trzech nurtów uczenia maszynowego. Uczeniu przez wzmacnianie nie przygotowuje się zestawu danych, a jedynie środowisko, z którego model zbiera dane. W pętli procesu uczenia się algorytmy ulepszają swoje sposoby działania, ucząc się najkorzystniejszego oddziaływania na środowisko.

Jednocześnie Cantrell, Martin i Ellis wskazują na udane próby ingerencji człowieka w środowisko, których celem było zachowania dzikości. Część z tych ingerencji polega na wprowadzaniu nowych gatunków, których zadaniem jest przejęcie funkcji gatunków wymarłych (na przykład w Oostvaardersplassen), inne związane są z wprowadzaniem technologii opartych na sztucznej inteligencji. Autorzy omawiają przykładowo wprowadzenie autonomicznego pojazdu podwodnego (COTSbot), który identyfikuje rozgwiazdę koroniastą (COTS) i wstrzykuje jej śmiertelne sole żółciowe w celu ochrony rodzimych gatunków koralowców<sup>35</sup>.

Jak piszą autorzy artykułu, w odpowiedzi na te trendy ekolodzy i biolodzy zajmujący się ochroną przyrody zaczęli badać możliwości tworzenia nowych ekosystemów jako „wytworzoną dzikość” (ang. created wildness). W tym podejściu paradoksalnie wpływy ludzkie są maksymalizowane (tworzenie systemów sztucznej inteligencji) w celu wzmocnienia wpływów pozaludzkich – ekosystemy i aktorzy „nie-ludzcy” funkcjonują poza bezpośrednim nadzorem człowieka. Podstawowym narzędziem do osiągnięcia „wytworzonej dzikości” byłaby w pełni autonomiczna infrastruktura sztucznej inteligencji, która łączyłaby DRL z systemami czujników i mechanizmów w celu wytworzenia reagujących interakcji, które tworzą i podtrzymują „nie-ludzką” dzikość

---

<sup>35</sup> Zob. *New Robot Has Crown-of-Thorns Starfish in Its Sights*, <https://www.qut.edu.au/news?news-id=95438>.

bez potrzeby ciągłej interwencji człowieka. Cantrell, Martin i Ellis taką infrastrukturę nazywają kreatorem dzikości.

### KREATOR DZIKOŚCI

Kreator dzikości to koncepcyjny projekt autonomicznego systemu infrastruktury krajobrazowej, który tworzy i podtrzymuje dzikość poprzez wzmacnianie wpływów pozaludzkich przy jednoczesnym przeciwdziałaniu wszelkim formom wpływu człowieka. Jest to system obliczeniowy oparty na głębokim uczeniu, który kontroluje infrastrukturę fizyczną. Dzięki systemom czujników i innych urządzeń infrastruktura ta może wyczuwać i manipulować środowiskiem oraz wchodzić w interakcje z organizmami. „Kreator dzikości – piszą Cantrell, Martin i Ellis – reprezentuje nie-ludzkiego inteligentnego aktora, początkowo opracowanego przez ludzkich projektantów, ale będącego w stanie uczyć się własnych nowatorskich zachowań strategicznych dzięki trwałym interakcjom środowiskowym w tworzeniu i utrzymaniu dzikiego ekosystemu wolnego od wpływów człowieka. Algorytmy, które kontrolują zachowanie systemu są wyuczone z jego kontekstu i działań, i nie są programowane przez ludzi”<sup>36</sup>.

Oto przykładowe zasady działania kreatora dzikości: (1) Operacje i działania są niewidoczne i niezrozumiałe dla ludzkich obserwatorów: procesy rządzące systemem są ukryte przed ludźmi i funkcjonalnie niepoznawalne dla ludzi.

(2) Ludzie odwiedzający „wytworzoną dzikość” są w stanie doświadczyć bycia w ekosystemie, który działa bez jakiegokolwiek wpływu człowieka.

(3) Kreator dzikości nieustannie monitoruje wpływy ludzkie i je usuwa lub ingeruje w nie. Na przykład antropogeniczny hałas jest eliminowany, antropogeniczne światło jest blokowane, a ludzkie artefakty i zanieczyszczenia są usuwane i filtrowane.

Cantrell, Martin i Ellis hipotetyczne użycie „kreator dzikości” opisują w ten sposób: „Kreator dzikości może zostać rozmieszczony i zacząć działać na terenie nadbrzeżnego mokradła. System najpierw wykorzystałby systemy wykrywania i uczenia się, aby zidentyfikować żywe organizmy i nie-ludzkie wzorce środowiskowe na całym terenie. Następnie system szukałby dowodów na wpływ człowieka i rozpoczęłby operacje w celu pomocy nie-ludzkim aktorom, takim jak rośliny, w przekształceniu wzorców środowiskowych w kierunku warunków pozbawionych wpływu człowieka. Prowadząc te działania, system uczyłby się najbardziej efektywnych strategii. Wraz z upływem czasu i podejmowaniem coraz to nowych działań, system uczy się w miarę zachodze-

<sup>36</sup> Cantrell, Martin, Ellis, dz. cyt., s. 161.

nia procesów sukcesji ekologicznej. Ostatecznie system jest w stanie rozwijać i wprowadzać w życie zupełnie nowe algorytmy i zachowania, które byłyby unikalne dla „twórcy dzikości” i potencjalnie nieprzewidziane przez żadne wcześniejsze zachowanie systemu<sup>37</sup>.

Z czasem „twórca dzikości” aktywnie poszukiwałby antropogenicznych czynników, w tym zanieczyszczeń pochodzących z procesów przemysłowych, antropogenicznego hałasu oraz innych pozostałości po obecności człowieka, ucząc się i wdrażając nowe protokoły, aby przeciwdziałać tym efektom. Twórca dzikości będzie również maskował swoje działania i modyfikacje tak, aby ludzie obserwatorzy postrzegali florę, faunę i abiotyczne środowisko jako nietkniętą ludzką ręką. Z czasem działania „kreatora dzikości” stałyby się nierozpoznawalne i niezrozumiałe dla człowieka. Natomiast same wzory i procesy ekologiczne zostałyby przekształcone w taki sposób, aby zatrzeć ślad oddziaływania człowieka.

#### KRYTYKA

Omawiając propozycję delegowania zarządzania środowiskiem naturalnym przy pomocy sztucznej inteligencji, warto zobaczyć to zagadnienie w szerszej, wykraczającej poza rozstrzygnięcia typowo inżynierskiej, perspektywie. We współczesnej humanistyce oraz naukach społecznych istnieje wiele, często komplementarnych tradycji, rozpoznawania relacji jakie zachodzą między wprowadzeniem nowych technologii a sferą kultury – studia nad nauką i technologią (STS), ekologia medialna (ang. media ecology), filozofia techniki czy ocena technologii (ang. technology assessment) są obecnie głównymi źródłami humanistycznej refleksji nad rolą infrastruktury technologicznej we współczesnym świecie. Aby właściwie ocenić skutki wprowadzenia danego rozwiązania technicznego, dobrze jest przyjąć pluralistyczną perspektywę. W przypadku zarządzania środowiskiem naturalnym przez sztuczną inteligencję należałoby zastanowić się nad następującymi kwestiami:

Czy jesteśmy gotowi do przekazania systemom sztucznej inteligencji decyzji o charakterze normatywnym?

W jaki sposób delegowanie kompetencji systemom cyfrowym wpływa na nasze zdolności rozumienia świata oraz na naszą motywację walki o lepszy świat?

Jak oparcie się na wielkoskalowej infrastrukturze sztucznej inteligencji wpłynie na tak zwany problem metali rzadki, e-odpady oraz kryzysy społeczno-ekologiczne?

---

<sup>37</sup> Tamże, s. 163.

## KTÓRE GATUNKI CHRONIĆ, KTÓRE TĘPIĆ? NORMATYWNY WYMIAR NAPRAWY PRZYRODY

Warto zauważyć, że koncepcja naprawy przyrody (ang. *restauration*) posiada niezbywalny normatywny charakter. O ile analizy biologów i ekologów mogą dostarczać wiedzy o gatunkach obecnie tworzących biotopy, o ich wzajemnych powiązaniach przez sieci pokarmowe oraz sposobach wzajemnego oddziaływania, to decyzja które kombinacje gatunków powinny w przyszłości zamieszkiwać konkretny biotop ma charakter normatywny<sup>38</sup>. Czy jesteśmy gotowi przekazać systemom sztucznej inteligencji możliwość podejmowania takich decyzji? I czy w związku z tym, że są to systemy w pełni zautomatyzowane oparte na DRL (decyzje sztucznej inteligencji będą dla nas nieprzejryste), jesteśmy gotowi zrezygnować z możliwości rozumienia świata wokół nas?

## ZDOLNOŚCI, KOMPETENCJE I OGRANICZANIE ŚWIADOMOŚĆ ZMIAN ŚRODOWISKOWYCH

Wśród autorów zajmujących się badaniem wdrażania systemów zautomatyzowanych stale rośnie zainteresowanie wpływem, jaki wywiera wprowadzanie nowych systemów technologicznych na nasze umiejętności. Badacze podkreślają, że automatyzacja nie zastępuje w prosty sposób ludzkich aktywności, lecz zmienia je – często w sposób trudny do przewidzenia. Delegowanie zarządzania środowiskiem naturalnym systemom sztucznej inteligencji niewymagającym ludzkiej kontroli może doprowadzić do podobnych zaburzeń świadomości zmian w środowisku zewnętrznym, jakie zachodzą wśród użytkowników autonomicznych samochodów – zmniejszona świadomość sytuacji oraz samozadowolenie. Gdy człowiek oddaje kontrolę nad podejmowaniem decyzji, staje się mniej świadomy zmian w stanie środowiska. Jeśli zautomatyzowany system jest wysoce, chociaż nie do końca, niezawodny w podejmowaniu decyzji, to ludzie mogą nie zauważyć momentów, w których automatyka zawodzi<sup>39</sup>. Zważywszy, że biosfera jest jednym z systemów planetarnych o kluczowym znaczeniu dla funkcjonowania całego systemu, skutki takich zaniedbań mogą być katastrofalne w skali globalnej.

<sup>38</sup> Zob. A. D u n k e r, *Rediscovering Earth. Ten dialogues on the future of nature*, OR Books, London, New York 2021, s. 40. Szerzej na temat niezbywalnego normatywnego charakteru nauki zob. A. L e k k a - K o w a l i k, *Dlaczego nauka nie może być wolna od wartości*, „Roczniki Filozoficzne” 52(2004) nr 2, s. 275-293; t a ż, *Odkrywanie aksjologicznego wymiaru nauki*, Wydawnictwo KUL, Lublin 2008.

<sup>39</sup> Por. P a r a s u r m a n i n, *A Model for Types and Levels of Human Interaction with Automation*, „IEEE Transactions on Systems Man and Cybernetics – Part A: Systems and Humans” 30(2000) nr 3, s. 289.



Warto również zauważyć, że obecnie w refleksji środowiskowej promuje się postawy związane z obywatelskim, oddolnym zaangażowaniem w troskę o środowisko naturalnym. Ruchy takie jak nauka obywatelska, włączająca w badanie zagrożonych gatunków zwykłych obywateli, postrzegane są jako jedna z możliwych strategii korekty funkcjonowania nauki – zaangażowanie w badanie środowiska wymaga nie tylko rozumienia funkcjonowania ekosystemów, ale również może prowadzić do bardziej holistycznego ujęcia związków między światem przyrody i ludźmi, do uzyskania „całościowego obrazu” (ang. *toto picture*), w którym sztuka, nauka i doświadczenie są zintegrowane<sup>40</sup>. Wydaje się, że większym pożytkiem dla stanu bioróżnorodności, ale również dla kondycji społecznej jest inwestowanie energii społecznej w strategię typu nauka obywatelska niż w kontrolę środowiska naturalnego przy pomocy sztucznej inteligencji.

Delegowanie opieki nad środowiskiem naturalnym systemom sztucznej inteligencji może prowadzić do społecznej znieczulicy na kwestie związane z zanikiem bioróżnorodności zgodnie z mechanizmem wyuczonej bezradności. Systemy sztucznej inteligencji, wyręczając ludzi w aktywności związanej z dbaniem o ekosystemy, mogą – jako skutek uboczny – powodować spadek zainteresowania światem przyrody oraz zmniejszać motywację do angażowania się w utrzymywanie bioróżnorodności. Warto również zadać pytanie co jest celem delegowania opieki nad środowiskiem naturalnym systemom sztucznej inteligencji? Czy jest nim zabezpieczenie stabilnej przyszłości naszej planety czy jedynie stworzenie doświadczenia przebywania w ekosystemie, który działa bez jakiegokolwiek wpływu człowieka? Temu drugiemu celowi można zarzucić promowanie stanowiska skrajnego antropocentryzmu, którego krytyka posiada długą i złożoną historię w refleksji środowiskowej.

## BIORÓŻNORODNOŚĆ, RÓŻNORODNOŚĆ KULTUROWA I WYOBRAŻNIA RELIGIJNA

Delegowanie troski o przyrodę maszynom może mieć jeszcze inne negatywne konsekwencje. Według wielu autorów istnieje ściśle powiązanie między różnorodnością kulturową i bioróżnorodnością<sup>41</sup>. Wprowadzenie „kreatora dzikości” może to powiązanie przerwać lub oddziaływać na nie w negatywny spo-

<sup>40</sup> M. H a n n I b a l, *Citizen Scientist. Searching for Heroes and Hope in an Age of Extinction*, New York: The Experiment 2016, s. 7.

<sup>41</sup> Por. T. S c h a a f, C. L e e, *Conserving Cultural and Biological Diversity: The Role of Sacred Natural Sites and Cultural Landscapes*, Proceedings of UNESCO-IUCN International Conference, Tokyo, Japan, UNESCO, 2006; N. D u d l e y, L. H i g g i n s – Z o g i b, S. M a n s o u r i a n, *The Links between Protected Areas, Faiths and Sacred Natural Sites*, “Conservation Biology” 23 (2009), s. 568-577.



sób. Kultury, które opierają swoją tożsamość na związkach ze światem natury (chodzi tu głównie o społeczności rdzenne i lokalne, ale również o wyznawców takich religii jak: buddyzm, taoizm, hinduizm, dżinizm, sintoizm, sikhizm, zoroastryzm) mogą stracić podstawowy punkt zaczepienia, dzięki któremu rozumieją same siebie. Trudno przewidzieć skalę negatywnych konsekwencji niszczenia takich rdzennych i lokalnych tożsamości, zwłaszcza w sytuacji, kiedy aktualna refleksja środowiskowa (humanistyka ekologiczna) coraz śmielej sięga do tak zwanej tradycyjnej wiedzy ekologicznej<sup>42</sup>. Warto również zauważyć, że ingerencja technologiczna w sposób funkcjonowania środowiska może wpłynąć również na naszą wyobraźnię religijną, gdyż – jak zauważa amerykański ekoteolog Thomas Berry – istnieje ściśle powiązanie między tymi dwoma aspektami (mianowicie przyrodą i wyobraźnią religijną)<sup>43</sup>.

#### PLANOWANA PRZESTARZAŁOŚĆ, METALE RZADKIE ORAZ KRYZYS SPOŁECZNO-EKOLOGICZNY

Warto również zwrócić uwagę na koszty wytworzenia i utrzymania cyfrowych infrastruktur technologicznych. Mimo deklaracji, że przy zastosowaniu bardzo trwałych materiałów, infrastruktura będzie mogła służyć latami, należy pamiętać, że globalny kapitalizm działa zgodnie z logiką projektowanej przestarzałości<sup>44</sup>. To oznacza, że trwałość systemów technologicznych jest niezwykle trudna do osiągnięcia. Systemy produkowane przez prywatne firmy muszą zapewnić odpowiednią stopę zwrotu, więc albo będą niezwykle drogie, albo nietrwałe, co będzie oznaczało konieczność częstych wymian, a w ostatecznym efekcie wydobywanie metali rzadkich na szeroką skalę. Jeśli weźmiemy pod uwagę ograniczone zasoby naszej planety, to minimalizowanie naszej zależności od cyfrowych infrastruktur technologicznych zyskuje negatywny wymiar ekologiczny.

Dodatkowo część metali rzadkich wydobywana jest na terytoriach, na których (z powodu dużej dostępności tych metali) toczą się niekończące się wojny. Metale te określane są jako minerały konfliktowe (ang. conflicts mineral). Zwiększanie popytu na te metale może niekorzystnie wpływać na przebieg procesów pokojowych<sup>45</sup>. Warto również zaznaczyć, że zwiększający się popyt na metale rzadkie niekorzystnie wpływa na sytuacje pracowników kopalń

<sup>42</sup> Zob. E. D o m a n i s k a, *Humanistyka ekologiczna*, „Teksty Drugie” 2013, nr 1-2, s. 13-32.

<sup>43</sup> Por. R. S a d o w s k i, *Thomas Berry – prorok ery ekozoicznej*, „Studia Ecologiae et Bioethicae” 7(2009) nr 2, s. 209.

<sup>44</sup> Por. M. P e r e l m a n, *Manufacturing Discontent: The Trap of Individualism in Corporate Society*, Pluto Press, London 2005, s. 27-30.

<sup>45</sup> Taki wniosek można wyciągnąć z pracy *How Green is Your Smartphone*, w której autorzy wskazują na rolę zwiększającego się popytu na metale niezbędne do produkcji smartfonów w de-

w krajach trzeciego świata. Jak piszą autorzy opracowania *How Green is Your Smartphone*: „Są oni [pracownicy kopalń metali rzadkich – B.K.] narażeni na zagrożenia dla dróg oddechowych i oddziaływanie pierwiastków promieniotwórczych, często w krajach, w których przemysł nie jest uregulowany lub przepisy są słabo egzekwowane. Górnicy często pracują na niestabilnym terenie lub w miejscach zagrożonych, gdzie mają do czynienia z łatwopalnym pyłem, pożarami i ryzykiem zawału. Wdychany przez nich pył rudny może prowadzić do chorób płuc, w tym zapalenia oskrzeli, krzemicy i raka. Produkty uboczne wydobywania złota [jednego z kluczowych metali koniecznych dla funkcjonowania infrastruktury cyfrowej – B.K.] zawierają takie trujące neurotoksyny jak ołów, cyjanek i rtęć<sup>46</sup>. Wątpliwości również budzi problem składowania odpadów infrastruktury cyfrowej (ang. e-waste), które tylko w ograniczonym stopniu są podawane recyklingowi. Biorąc pod uwagę fakt, że „okres połowicznego rozpadu toksycznych odpadów emitowanych do gleby z zakładów produkujących półprzewodniki sprawia, że wody gruntowe i grunty nie nadają się do użytku lub są wysoce niebezpieczne dla zamieszkującej je ludności na długo po tym, jak winne firmy odeszły<sup>47</sup>, warto zadać ponownie zadać pytanie o ekologiczny charakter infrastruktur cyfrowych. Jednocześnie na przykładzie zagadnienia metali rzadkich, ich ograniczonego występowania oraz niekorzystnych skutków społecznych, których przyczyną jest wzrastający na nie popyt, widzimy, jak złożonym zagadnieniem jest problem związany z zarządzaniem środowiskiem naturalnym przez sztuczną inteligencję.

\*

Promowane przez Cantrella, Martina i Ellisa tworzenie dzikości za pomocą systemów sztucznej inteligencji należy starać się ująć w szerokim kontekście. Z całą pewnością branie pod uwagę jedynie aspektów inżynierskich, pytań technicznych o możliwości omawianej infrastruktury cyfrowej jest niewystarczające. Propozycja zautomatyzowania opieki nad dzikimi miejscami powinna uruchomić całą gamę pytań o charakterze ontologicznym, etycznym, kulturowym, psychologicznym i społecznym. W ramach dyskusji prowadzonych w obrębie dyscyplin humanistycznych szczególną uwagę powinno się poświęcić koncepcji postnatyry, która w aspekcie omawianego przez nas projektu, ukazuje swój ambiwalentny charakter. Z jednej strony pojęcie to

stabilizacji sytuacji w Demokratycznej Republice Kongo. Por. R. M a x w e l l, T. M I l l e r, *How Green is Your Smartphone?*, Polity Press, Cambridge 2020, s. 71.

<sup>46</sup> Tamże, s. 70.

<sup>47</sup> Tamże, s. 79.

posiada emancypacyjny potencjał, ponieważ wskazuje na rolę wartości i narracji w procesie przeobrażania naszego świata. Dzięki temu strategie mające na celu zabezpieczenie stabilnej przyszłości naszej planety mogą wyjść poza rozwiązania naukowo-techniczne i zacząć zadawać pytania o rolę przemian w obrębie szeroko rozumianej kulturze. Z drugiej strony koncepcja postnatury może być użyta w roli perswazyjnej, aby uzasadniać zintensyfikowane interwencje technologiczne w obrębie tak zwanego środowiska naturalnego.

#### BIBLIOGRAFIA / BIBLIOGRAPHY

- Bińczyk, Ewa. *Epoka człowieka. Retoryka i marazm antropocenu*. Warszawa: PWN, 2018.
- . “Inżynieria klimatu a inżynieria człowieka: Dyskursy na temat środowiska w epoce antropocenu.” *Ethos* 28, no. 3(111), (2015): 153–75.
- Bołoz, Wojciech. *Kościół i ekologia. W obronie człowieka i środowiska naturalnego*. Kraków: Wydawnictwo Homo Dei, 2010.
- Bonneuil, Christophe. “The Geological Turn: Narratives of the Anthropocene.” In *The Anthropocene and the Global Environmental Crisis. Rethinking modernity in a new epoch*. Edited by Clive Hamilton, Christophe Bonneuil, François Gemenne. London: Routledge, 2015.
- Brooks, Peter. *Seduced by Story: The Use and Abuse of Narrative*. New York: The New York Review of Books, 2022.
- Cantrell, Bradley, Laura Martin, and Erle Ellis. “Designing Autonomy: Opportunities for New Wildness in the Anthropocene.” *Trends in Ecology & Evolution* 32, no. 3 (2017): 156–66.
- Caradonna, Jeremy, et al. *An Ecomodernist Manifesto: A Degrowth Critique*. [http://www.resilience.org/wp-content/uploads/articles/General/2015/05\\_May/A-Degrowth-Response-to-An-Ecomodernist-Manifesto.pdf](http://www.resilience.org/wp-content/uploads/articles/General/2015/05_May/A-Degrowth-Response-to-An-Ecomodernist-Manifesto.pdf).
- Cronon, William. “The Trouble with Wilderness: Or Getting Back to the Wrong Nature.” In *Uncommon Ground: Rethinking the Human Place in Nature*. Edited by William Cronon. New York: Norton, 1996.
- Deepfield-robotics.com. [www.deepfield-robotics.com](http://www.deepfield-robotics.com).
- Descola, Philippe, *Beyond Nature and Culture*. Chicago: University of Chicago Press, 2013.
- Domańska, Ewa. “Humanistyka ekologiczna.” *Teksty Drugie* 1-2 (2013): 13–32.
- Dudley, Nigel, Liza Higgins-Zogib, and Stephanie Mansourian Stephanie. “The Links between Protected Areas, Faiths and Sacred Natural Sites.” *Conservation Biology* 23 (2009): 568–77.
- Dunker, Anders. *Rediscovering Earth: Ten Dialogues on the Future of Nature*. London, New York: OR Books, 2021.

- Fermaux, Ann. "The Return of Nature in the Capitolocene." In *Rethinking the Environment for the Anthropocene: Political Theory and Sociocultural Relations in the New Geological Epoch*. Edited by Manuel Aries-Maldonado, Zev Trachtenberg. London: Routledge, 2019.
- Grunwald, Armin. "Diverging Pathways to Overcoming the Environmental Crisis: A Critique of Ecomodernism from Technology Assessment Perspective." *Journal of Cleaner Production* 197, no. 2 (2018): 1854–62.
- . *The Hermeneutical Side of Responsible Research and Innovation*. ISTE Ltd: London 2016.
- Guha, Ramachandra. "Radical American Environmentalism and Wilderness Preservation: A Third World Critique." In *The Future of Nature: Documents of Global Change*. Edited by Libby Robin, Sverker Sörlin, and Paul Warde. New Haven: Yale University Press, 2013.
- Hannibal, Mary Ellen. *Citizen Scientist: Searching for Heroes and Hope in an Age of Extinction*. New York: The Experiment, 2016.
- Haraway, Donna. *Staying with the Trouble: Making Kin in the Chthulucene*. Durham and London: Duke University Press, 2016.
- Heise, Ursula. *Imagining Extinction: The Cultural Meanings of Endangered Species*. Chicago: The University of Chicago Press, 2016.
- Jachowski, David, Rob Slotow, and Joshua J. Millspaugh, "Good Virtual Fences Make Good Neighbors: Opportunities for Conservation." *Animal Conservation* 17, no. 3 (2014): 187–96.
- Jackson, Tim. *Dobrobyt bez wzrostu. Ekonomia dla planety o ograniczonych zasobach*. Translated by Marcin Polakowski. Toruń: Wydawnictwo UMK, 2015.
- Latour, Bruno. *Facing Gaia. Eight Lectures on the New Climatic Regime*. Cambridge: Polity 2017.
- Lekka-Kowalik, Agnieszka. "Dlaczego nauka nie może być wolna od wartości." *Roczniki Filozoficzne* 52, no. 2 (2004): 275–93.
- . *Odkrywanie aksjologicznego wymiaru nauki*. Lublin: Wydawnictwo KUL, 2008.
- Lovins, Amory, and Paul Hawken. *Natural Capitalism: The Next Industrial Revolution*. London: Earthscan, 1999.
- Manifest modernistyczny*. [www.ecomodernism.org/jzyk-polski](http://www.ecomodernism.org/jzyk-polski).
- Maxwell, Richard, and Toby Miller. *How Green is Your Smartphone?* Cambridge: Polity Press, 2020.
- Merchant, Carolyn. *The Anthropocene and the Humanities: From Climate Change to a New Age of Sustainability*. New Heaven, London: Yale University Press, 2020.
- New Robot Has Crown-of-Thorns Starfish in Its Sights*. <https://www.qut.edu.au/news?news-id=95438>.
- On Becoming an Ecomodernist: A Positive Vision of Our Environmental Future*. BreakThrough Institute. [www.thebreakthrough.org/articles/on-becoming-an-ecomodernist](http://www.thebreakthrough.org/articles/on-becoming-an-ecomodernist).

- Parasuraman, Raja, Thomas Sheridan, and Christopher Wickens. "A Model for Types and Levels of Human Interaction with Automation." *IEEE Transactions on Systems Man and Cybernetics: Part A; Systems and Humans* 30, no. 3 (2000): 286–97.
- Perelman, Michael. *Manufacturing Discontent: The Trap of Individualism in Corporate Society*. London. Ann Arbor. MI: Pluto Press, 2005.
- Plumwood, Val. *Environmental Culture: The Ecological Crisis of Reason*. London: Routledge, 2002.
- Sadowski, Ryszard. "Thomas Berry – prorok ery ekozoicznej." *Studia Ecologiae et Bioethicae* 7, no. 2 (2009): 205–16.
- . *Filozoficzny spór o rolę chrześcijaństwa w kwestii ekologicznej*. Warszawa: Towarzystwo Naukowe Franciszka Salezego, 2015.
- Schaaf, Thomas, and Cathy Lee. *Conserving Cultural and Biological Diversity: The Role of Sacred Natural Sites and Cultural Landscapes*. Proceedings of UNESCO-IUCN International Conference. Tokyo. Japan. UNESCO. Paris 2006.
- Stengers, Isabelle. "Accepting the Reality of Gaia: A Fundamental Shift?" In *The Anthropocene and the Global Environmental Crisis: Rethinking Modernity in a New Epoch*. Edited by Clive Hamilton, Christophe Bonneuil, and François Gemenne. London: Routledge, 2015.
- Wróblewski, Zbigniew. *Natura i cele. Dyskusja argumentu teleologicznego na rzecz ochrony przyrody*. Lublin: Wydawnictwo KUL, 2010.

## ABSTRAKT / ABSTRACT

Bartłomiej KNOSALA – Zarządzanie środowiskiem naturalnym przez sztuczną inteligencję. Ograniczenia i wyzwania narracji postnatURY

DOI 10.12887/36-2023-4-144-10

Artykuł stanowi próbę krytycznej analizy wielowymiarowych efektów propozycji delegowania możliwości zarządzania środowiskiem naturalnym systemom sztucznej inteligencji opartym na tak zwanym głębokim uczeniu przez wzmacnianie (DRL). Punktem wyjścia jest koncepcja zaproponowana przez Armina Grunwalda, która dla zwiększenia transparentności potencjalnej debaty publicznej wokół NEST (ang. New and Emerging Science and Technology), postuluje ocenę i analizę znaczeń, które są przypisywane do innowacji naukowych i technologicznych na bardzo wczesnym etapie ich opracowywania. Kluczowym pojęciem, w ramach którego usytuowane są propozycje delegowania zarządzania środowiskiem naturalnym przez sztuczną inteligencję, jest koncepcja postnatURY. W wyniku przeprowadzonych analiz pojęcie to ukazuje swój ambiwalentny charakter. Z jednej strony posiada ono emancypacyjny potencjał, wskazuje bowiem na rolę wartości i narracji w procesie przeobrażania naszego świata. Dzięki temu strategię mające na celu zabezpieczenie stabilnej przyszłości naszej planety mogą wyjść poza wąskie rozwiązania naukowo-techniczne i zacząć zadawać pytania o rolę przemian w obrębie szeroko rozumianej

kultury. Z drugiej strony koncepcja postnatury może być użyta w roli perswazyjnej jako sposób, dzięki któremu łatwiej jest przekonać do zintensyfikowanych interwencji technologicznych w obrębie tak zwanego środowiska naturalnego.

Słowa kluczowe: zarządzanie środowiskiem naturalnym, sztuczna inteligencja, antropocen, postnatura, narracje, techno-optimizm, hermeneutyczny model oceny technologii

Kontakt: Katedra Stosowanych Nauk Społecznych, Wydział Organizacji i Zarządzania, Politechnika Śląska, ul. Roosevelta 26-28, 41-800 Zabrze  
Tel. 32 27773

E-mail: bartlomiej.knosala@polsl.pl

Bartłomiej KNOSALA, Environmental Management by Artificial Intelligence: Challenges and Limitations of the Concept of Postnature

DOI 10.12887/36-2023-4-144-10

This article attempts to critically analyze multidimensional effects of the proposals to delegate environmental management capabilities to artificial intelligence systems based on the so-called deep learning through reinforcement (DRL). The starting point is the concept proposed by Armin Grunwald, which, in order to make the potential public debate around NEST (New and Emerging Science and Technology) more transparent, advocates an assessment and analysis of the meanings that are attached to scientific and technological innovations at a very early stage of their development. The key idea within which the proposals to delegate management of the environment by artificial intelligence are situated is the concept of post-nature. As a result of the analyses carried out in this article, the concept in question shows its ambivalent character. On the one hand, it has an emancipatory potential, as it points to the role of values and narratives in the process of transforming our world. In this way, strategies aimed at securing a stable future for our planet can go beyond narrow scientific and technological solutions and we can begin to ask questions about the role of transformation within the broader culture. On the other hand, the concept of postnature can be used in a persuasive manner as a way through which it is easier to win acceptance of intensified technological interventions within the so-called natural environment.

Keywords: environmental management, artificial intelligence, Anthropocene, postnature, narratives, techno-optimism, hermeneutic model of technology assessment

Contact: Department of Applied Social Science. Faculty of Organization and Management, Silesian University of Technology, ul. Roosevelta 26-28, 41-800 Zabrze, Poland

Phone: +48 32 27773

E-mail: bartlomiej.knosala@polsl.pl



Izabela KACZMARZYK

## RZECZ „GODNA WIDZENIA” CZY „APOKALIPTYCZNY POTWÓR”?

Konsekwencje maszynizacji górnośląskiej rzeczywistości kulturowej

*Ogólne ramy zmiany kulturowej związanej z maszynizacją, naszkicowane przez naszych osiemnastowiecznych i dziewiętnastowiecznych przodków, wprowadziły nas w świat technologicznego przyspieszenia, które zadecydowało o zmechanizowaniu i umaszynowaniu (również cyfrowym) ludzkiej egzystencji. Wydaje się, że mimo swojej „apokaliptyczności” wcześniejszy, przedcybernetyczny etap maszynizacji (zarówno w perspektywie górnośląskiej, jak i w szerszym, może nawet globalnym wymiarze), mimo obaw, które budził, z perspektywy czasu aż tak niebezpieczny dla istoty bycia człowiekiem się nie okazał.*

Mieczysław Orłowicz tak opisywał przed pierwszą wojną światową Górną Śląsk: „Pod względem wyglądu da się Górną Śląsk podzielić na trzy części: *Biały Śląsk* od Opola po Gliwice, tak nazwany od licznych fabryk i kopalń wapna i cementu, *Zielony Śląsk* w okolicach Raciborza, kraj pagórków, lasów i parków oraz *Czarny Śląsk*, czyli Górnośląskie Zagłębie Węglowe, zaciemnione dymem hut, fabryk i kopalni”<sup>1</sup>. Istotą funkcjonowania tych „hut, fabryk i kopalni” były różnego typu maszyny zarówno stosowane od średniowiecza w górnictwie kruszcowym maszyny proste (jak chociażby różnego typu kołowroty), jak i maszyny bardziej skomplikowane, wykorzystujące siłę pary, a później także elektryczności (od maszyny parowej przez ładowarki, wrębiarki aż po pierwsze kombajny górnicze, które zaczęto stosować w połowie dwudziestego wieku). Wraz z coraz liczniejszym i powszechniejszym zastosowaniem takich maszyn oraz kolejnymi kilometrami torów, którymi poruszały się pociągi, kolejki wąskotorowe, a później również tramwaje, rzeczywistość kulturową Czarnego Śląska, przenikało i naznaczało doświadczenie maszynizacji, rozumianej jako jeden z elementów procesu kulturowego, który i zaważył na kulturowej specyfice tego regionu<sup>2</sup>. W tej perspektywie celem artykułu jest próba rekonstrukcji roli i znaczenia maszyn w kształtowaniu się górnośląskiej rzeczywistości kulturowej od końca osiemnastego do przełomu dziewiętnastego i dwudziestego wieku oraz wpływu, jaki funkcjonowanie w umaszynowanym świecie miało na wykształcenie się

<sup>1</sup> M. O r ł o w i c z, *Przewodnik po ziemiach dawnej Polski, Litwy i Rusi*, nakł. Eugeniusza Starczewskiego, Kraków 1914, s. 172.

<sup>2</sup> Był to zarazem jeden z istotnych czynników stereotypizacji wizerunku Górnego Śląska, ale ten aspekt kwestii uprzemysłowienia przekracza ramy niniejszego artykułu.



specyficznych cech i umiejętności charakteryzujących postawę górnośląskich robotników wobec pracy<sup>3</sup>.

Bazę źródłową stanowią będzie polskojęzyczna i niemieckojęzyczna epistolografia, relacje etnograficzne, podróżnicze, dziennikarskie oraz dokumenty memoratywne, ze szczególnym uwzględnieniem tak zwanej *Złotej księgi Tarnowskich Gór*<sup>4</sup>, zawierające obserwacje i refleksje związane z obecnością maszyn w górnośląskim przemyśle<sup>5</sup>. Uwaga badawcza skoncentrowana zostanie na analizie wątków i motywów pozostających w kręgu dyskursu o maszynie<sup>6</sup> na podstawie niefikcyjnych tekstów kultury, związanych tematycznie z przestrzenią między Mysłowicami a Gliwicami oraz Bytomiem a Tarnowskimi Górami<sup>7</sup>, gdzie nastąpiła największa koncentracja przemysłu, ponieważ na tym obszarze znajdowały się bogate złoża naturalne i który cechowały znaczne, jak na koniec dziewiętnastego wieku, udogodnienia komunikacyjne. Zwracał na to uwagę już ks. Adolf Hytrek, pisząc „Koleje żelazne przeryniają kraj cały wzdłuż i wszerz tak, że prawie wszystkie miasta niemi są połączone. Nadto przez cały kraj przechodzi szeroka, bita droga prowadząca z Wrocławia do Krakowa [...], którą w dawnych czasach krakowscy kupcy przez Szlązk na jarmarki niemieckie jeździli”<sup>8</sup>.

Elementem łączącym analizowane teksty jest przede wszystkim wzmiankowanie obecności i wskazywanie roli maszyn jako dowodu zmian cywilizacyjnych (w rozumieniu poziomu rozwoju technologicznego). Teksty te mają współcześnie charakter rozproszony. Nie ma bowiem antologii, która gro-

<sup>3</sup> Tak rozumiany proces maszynizacji stał się ważnym elementem kształtowania się górnośląskiego imaginarijumu przemysłowego, o czym pisałam szerzej, zwłaszcza w kontekście przemian górnośląskiego krajobrazu kulturowego (zob. I. K a c z m a r z y k, *Krajobraz, portret, maska. Studia o górnośląskim imaginarijumu przemysłowym*, Wydawnictwo Naukowe Akademii Ignatianum, Kraków 2019).

<sup>4</sup> Zob. *Das Goldene Buch von Tarnowitz (Fremdenbuch der Frieigrichsgrube von 1788 ab.). Ehrentafel. Verzeihnis der freiwilligen Beiträge der Berg- und Hüttenleute 1813/14 nebst Begleitbriefe*, Breslau 1913; *Złota księga Tarnowskich Gór*, oprac. M. Wroński, tłum. z jęz. niem. E. Szyga, tłum. z jęz. łac. Z. Jedynak, tłum. z jęz. fr. A. Górka, tłum. z jęz. szw. B. Kokot, tłum. z jęz. ang. N. Jarzyńska, Instytut Tarnogórski, Muzeum Instytutu Tarnogórskiego, Tarnowskie Góry 2011.

<sup>5</sup> Osobną grupę tekstów stanowią typowo techniczne opisy urządzeń przygotowywanych na potrzeby dokumentacji poszczególnych zakładów przemysłowych, zwłaszcza kopalń. Należy też zaznaczyć, że maszyny, ich techniczna specyfika i formy eksploatacji były również tematem prac dyplomowych przygotowywanych przez adeptów górnictwa. Zob. np. J. J a r o s, *Historia kopalni Król w Chorzowie (1791-1945)*, Wydawnictwo Górniczo-Hutnicze, Katowice 1962.

<sup>6</sup> Jest to określenie Macieja Wróblewskiego (zob. M. W r ó b l e w s k i, *Literatura i maszyna*, Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, Toruń 2015).

<sup>7</sup> W odniesieniu do tego obszaru zaczęto w dziewiętnastym wieku używać określenia „Oberschlesische Industriegebiet”.

<sup>8</sup> A. H y t r e k, *Górny Szlązk pod względem obyczajów, języka i usposobienia ludności*, oprac. B. Kalinowski, „Przegląd Polski” 14 (1879) nr 3, s. 297.

madziłaby ich najbardziej reprezentatywne egzemplifikacje<sup>9</sup>. Należy jednak podkreślić, że nie są to wyłącznie pokryte patyną czasu źródła do badania dziejów górnośląskiej kultury technicznej czy historii gospodarczej, ale – jeżeli spojrzeć na nie w kontekście analityczno-interpretacyjnych działań podejmowanych z perspektywy „impulsu antropologicznego”<sup>10</sup> – są to teksty, których konkretyzacja czytelnicza (by odwołać się do klasycznego Ingardenowskiego ujęcia<sup>11</sup>) pozwala na diachroniczne spojrzenie na kształtowanie się ambiwalentnych relacji pomiędzy człowiekiem a maszyną.

#### U POCZĄTKÓW DYCHOTOMII W PERCEPCJI MASZYNOWEGO ŚWIATA

Piotr Łaszczycza w odniesieniu do współczesnych pytań o relacje między człowiekiem (czy ściślej: ludzkim umysłem) a maszyną wskazywał, że rozwój cywilizacyjny, tempo pracy oraz przekształcanie środowiska sprawiają, że maszyny występują w rolach operatorów, protez i konkurentów człowieka. Pełnienie tych ról przez maszyny nie jest stałe, odgórnie przyporządkowane, wręcz przeciwnie, podlega nieustannym zmianom, „rodząc problemy dostosowania funkcji człowieka i maszyn, zastępowania pracy człowieka przez maszyny, rosnącej roli przetwarzania informacji i przekraczania biologicznych ograniczeń ciała ludzkiego”<sup>12</sup>. Pojawia się wobec tego naturalne pytanie, czy postęp narastający w ciągu geometrycznym to przede wszystkim rzecz „widerzenia godna”<sup>13</sup>, nad którą człowiek – pewien swojej siły – może spoglądać z podziwem dla samego siebie, a zarazem z uczuciem wyższości, wynikającej

<sup>9</sup> Niezwykle interesującą antologię tekstów związanych z kolejnictwem opracował Wojciech Tomasiak (zob. W. T o m a s i a k, *Romantycy i technika*, t. 1-2, Instytut Badań Literackich PAN, Warszawa 2017). Wybory tekstów źródłowych o śląskiej proweniencji, w których obecne są wątki związane z przemysłem, to na przykład: *Górny Śląsk i Zagłębie w dawnych opisach. Wiek XIX*, wybór, wstęp i oprac. A. Zieliński, Wydawnictwo Śląsk, Katowice 1984; *Von der Dampfmaschine zur Eisenbahn. Bildquellen und Dokumenten zur Frühindustrialisierung Oberschlesiens 1780-1860/ Od maszyny parowej do kolei żelaznej. Źródła ikonograficzne i dokumenty do wczesnego uprzemysłowienia Górnego Śląska 1780-1860*, red. P. Chmiel i in., tłum. P. Chmiel i in., Oberschlesisches Landesmuseum-Muzeum Górnictwa Węglowego, Ratingen-Zabrze 2004. W publikacjach tych nie został jednak w sposób autonomiczny potraktowany wątek maszyn i umaszynowienia.

<sup>10</sup> Zob. D. Ś n i e ż k o, *Antropologia w badaniach nad literaturą dawną*, w: *Jaka antropologia literatury jest dzisiaj możliwa?*, red. P. Czaplinski, A. Legeżyńska, M. Telicki, Wydawnictwo „Poznańskie Studia Polonistyczne”, Poznań 2010.

<sup>11</sup> Zob. R. I n g a r d e n, *Szkice z filozofii literatury*, Spółdzielnia Wydawnicza Polonista, Łódź 1947.

<sup>12</sup> P. Ł a s z c z y c z a, *Człowiek i jego maszyny. Operatorzy i protezy*, „Filo-Sofija” 2017, nr 39, s. 50.

<sup>13</sup> J. L o m p a, *Krótki rys geografii Szląska dla nauki początkowej*, nakł. Henryka Handla, Głogówek 1847, s. 31.

z samego faktu bycia istotą ludzką, czy też spuszczone została z łańcucha maszynowa bestia, której apokaliptyczności bynajmniej lekceważyć nie wolno, bo dla człowieka pozostawia ona coraz mniej miejsca?

Machiny różnego typu towarzyszą człowiekowi od zawsze, a proste zasady mechaniki, pozwalające na przykład uzyskać efekt dźwigni (dwustronnej albo przynajmniej jednostronnej), czyli tak zwana swoista „skrzydlata umiejętność”, są wykorzystywane przez człowieka odkąd zorientował się, że jest zdolny zwiększyć swoją naturalną siłę i możliwość oddziaływania na otoczenie. Legendarne są w tym kontekście przypisywane Archimedesowi słowa: „Dajcie mi punkt podparcia, a poruszę Ziemię”. Istotą tego typu prostych mechanizmów było jednak przede wszystkim to, że nie wywoływały one uczucia niepokoju, bo zasada oraz efekty działania były widoczne i oczywiste. I chociaż takich prostych maszyn przez długi czas było zdecydowanie najwięcej, to trzeba też pamiętać, że nigdy nie miały one technologicznego monopolu, konstruowano również bardziej skomplikowane urządzenia, niepozwalające gołym okiem dostrzec sposobu, w jaki funkcjonują. Były to urządzenia, w których widoczny był efekt końcowy ich działania. Takie urządzenia sprawiały wrażenie, że „działają same”, a już na pewno człowiek nie wchodził z nimi w interakcję tak bezpośrednią jak w przypadku, kiedy chciał posłużyć się łomem. I bez wątpienia to właśnie te pierwowzory maszynowego rozwoju stanowiły źródło dychotomicznego stosunku człowieka do maszyny, z jednej strony jako artefaktu budzącego podziw, a z drugiej wywołującego trwogę.

Dzisiaj, chociaż groźne ślepieia lokomotywy w ciemności już nie budzą takiego lęku, to jednak niejednokrotnie niepokoi nas, a czasem nawet przeraża, nie tyle fizyczne niebezpieczeństwo, na jakie w kontakcie z maszynami jesteśmy narażeni, ale możliwość zawładnięcia najważniejszym z naszych organów – mózgiem. W konsekwencji rodzi się pytanie, czy „bunt maszyn” jest możliwy i co począć z naszym strachem przed technologiczną władzą algorytmów, przed tym, że maszyna niekoniecznie jest rodzajem posłusznego, udomowionego zwierzątka, że może raczej zdecydowanie bliżej jest jej jednak do groźnej bestii. I pewnie, dlatego wcale nie jesteśmy pewni, że ów umaszynowany potwór zawsze będzie się poddawał rozkazom naszego intelektu. Najnowsze generacje maszyn mocno się emancypują, a technologia – jak pisał już dekadę temu Zbyszko Melosik – staje się dla człowieka jednym z decydujących układów odniesienia w konstruowaniu własnej tożsamości i życia: „Stanowi część matrycy stworzonej przez interakcje wolności i zniewolenia, podmiotowości i podporządkowania; matrycy, z której wyjść nie można, ponieważ konstytuuje tożsamość [człowieka – I.K.]”<sup>14</sup>. Melosik zwracał też uwagę, odwołując się do

<sup>14</sup> Z. M e l o s i k, *Technologizacja życia i tożsamości w kulturze współczesnej*, „Studia Edukacyjne” 2016, nr 38, s. 57 (<https://repozytorium.amu.edu.pl/handle/10593/14837>).

myśli francuskiego filozofa techniki Jacques'a Ellula, że rozwoju technologii zatrzymać nie można, gdyż jej immanentną cechą jest to, że nieodwołalnie postępuje naprzód, przy czym, „rozwojowość ta ma także swój wymiar przestrzenny; technologia nie może uzyskać odpowiedniego rozmachu, dopóki nie jest zastosowana wszędzie. Nie może pozostawić żadnej dziedziny nienaruszonej”<sup>15</sup>. „Tak więc, o ile «w przeszłości technologia była pochłaniana przez cywilizację, której była częścią», to obecnie «stała się [...] pochłaniającym elementem, w ramach którego wszystko jest usytuowane»”<sup>16</sup>.

Współczesna technologizacja, zwłaszcza po doświadczeniach funkcjonowania w rzeczywistości pandemii, oderwała człowieka w dużym stopniu przede wszystkim od przywiązania do określonego terytorium i sprawiła, że mocno względny stał się czas. Te dwa wyznaczniki zdecydowanie różnią nasze doświadczenie kontaktu z maszynami od wcześniejszych. Jednak – poza tymi elementami – ogólne ramy zmiany kulturowej związanej z maszynizacją zostały naszkicowane przez naszych osiemnastowiecznych i dziewiętnastowiecznych przodków. Nie były one co prawda pierwszym znaczącym osiągnięciem ludzkiego intelektu w dziejach świata, jednak wprowadziły nas w świat technologicznego przyspieszenia, które zadecydowało o zmechanizowaniu i umaszynowaniu (również cyfrowym) ludzkiej egzystencji, w tym w coraz bardziej zaawansowane procesy uczenia maszynowego (samouczenia się maszyn, systemów uczących się), czyli systemów potrafiących się doskonalić na bazie doświadczenia. I chociaż wykorzystywane w tych procesach algorytmy bazują na danych statystycznych i prawdopodobieństwie, to jednak w relacji z takimi systemami mamy poczucie co najmniej emocjonalnej ambiwalencji w zakresie naszej sprawczości i kontroli nad światem. Tego typu „proteżowanie ludzkiej zdolności do przetwarzania informacji przez maszyny stwarza [bowiem] – na razie fantastyczną – perspektywę rywalizacji inteligencji maszyn z inteligencją człowieka”<sup>17</sup>.

#### POCZĄTEK „WIELKICH DZIEŁ TAK POŻĄDANEGO POSTĘPU”<sup>18</sup>

Erik Brynjolfsson i Andrew McAfee wskazywali, że przez wieki „[...] postęp przebiegał boleśnie powoli, niemal niepostrzeżenie. [...] Dopiero nieco ponad dwieście lat temu nagle pojawiło się coś, co wygięło krzywą ludzkiej historii pod kątem niemal 90 stopni. [...] Chodzi mianowicie o rewolucję prze-

<sup>15</sup> J. E l l u l, *The Technological System*, tłum. J. Neugroschel, Continuum, New York 1980, s. 193; cyt. za: M e l o s i k, dz. cyt., s. 44.

<sup>16</sup> Tamże.

<sup>17</sup> Ł a s z c z y c a, dz. cyt., s. 51.

<sup>18</sup> *Złota księga Tarnowskich Gór*, s. 6.

mysłową<sup>19</sup>. Proces ten zdynamizował się zwłaszcza w drugiej połowie dziewiętnastego wieku. To wówczas maszyny zaczęły być coraz silniej związane z życiem przeciętnego człowieka, który stykał się z nimi chociażby, korzystając z przedmiotów codziennego użytku, wytwarzanych już nie w rzemieślniczych warsztatach, ale w fabrykach produkujących na skalę masową. Wiek pary i elektryczności stanowi tym samym rodzaj kanwy, na której pojawiają się wątki i motywy związane z rolą maszyn. Jest przy tym charakterystyczne, że chociaż niejednokrotnie dochodziło do głosu owo wspomniane uczucie trwogi, oparte na przekonaniu, że maszyny są „prawie” żywe<sup>20</sup>, to jednak partykuła ta oznaczała, że „całkiem żywe” nie są, a tym samym nie podawano w wątpliwość ogólnej zasady nadrzędności człowieka. Ten etap kontaktu z maszynowym światem cechował bowiem jeszcze wyraźny instrumentalizm. Każdy techniczny wynalazek i jego zastosowanie, zwłaszcza na szerszą skalę, był wartościowany pozytywnie. Analizując różne podejścia do maszyny i techniki, zwracała na to uwagę Grażyna Osika, podkreślając, że z tego punktu widzenia maszynowe artefakty rozpoznaje się co prawda jako obce byty, ale takie, które człowiek sam powołał do istnienia, kierując się racjonalnymi przesłankami usprawniania działania. Tym samym wspomniana wyżej nadrzędność nad maszynowym światem pozostaje poza wszelką dyskusją, a tym samym „w tym prostoliniowym spojrzeniu oczywiste jest, że mamy pełne panowanie nad sytuacją, bo to człowiek wyznacza cele i wartości, jakim służy dane rozwiązanie techniczne”<sup>21</sup>.

Maszynowe ekfrazy (również te górnośląskie) nie budziły więc niepokoju (a tym bardziej lęku) przed przejęciem władzy nad ludzkim umysłem, a jedynie miały co najwyżej animizacyjny wydźwięk, idący w kierunku paraleli do mechanicznego bestiariusz, najczęściej z wyraźną infernalną nutą: „Nie przesadzam, jeśli tu z 50 machin parowych nie sapie i dyszy jak potwory apokaliptyczne”<sup>22</sup> – opisywał swoją wizytę w górnośląskim zakładzie przemysłowym jeden z dziewiętnastowiecznych inżynierów, a korespondent „Kraju” po odwiedzeniu Katowic pisał z kolei: „Miejscowość otoczona lasem kominów sterczących przy kopalniach i fabrykach wysoko w górę. Olbrzymie zakłady górnicze i fabryczne. Wieczorem, gdy nad ziemią rozpostrze się ciemność nocy, zieją piece płomieniem [...]. Niekształtne fabryczne aneksy w postaci olbrzymich kotłów, powywracanych do góry nogami, gdzie ogień nie wygasa,

<sup>19</sup> E. Brynjolfsson, A. McAfee, *Drugi wiek maszyny. Praca, postęp i dobrobyt w czasach genialnych technologii*, przeł. B. Sałbut, Warszawa 2015, s. 15. Cyt. za: G. Osika, *Czekając na osobliwość – o modelach interpretacji techniki*, „Filo-Sofija” 2017, nr 39, s. 65.

<sup>20</sup> Por. *Historia piękna*, red. U. Eco, tłum. A. Kuciak, Dom Wydawniczy Rebis, Poznań 2005, s. 382.

<sup>21</sup> Osika, dz. cyt., s. 71.

<sup>22</sup> *Na Szląsku i w Styrii. Z notat inżyniera Cz [spisał K. Saryusz]*, „Wędrowiec” 3(1878) seria 3, nr 53, s. 346 (<https://bcu.lib.uni.lodz.pl/dlibra/publication/378>).

przybierają w nocnej ciemni kształty strasznych smoków, ziejących palącą się siarką<sup>23</sup>. Identyczne infernalne obrazy pojawiają się również w tekstach o lokalnej proveniencji, w tym w pieśniach śpiewanych przez górnośląskich robotników, które zostały zebrane przez Łukasza i Stanisława Wallisów: „Gdzie z komina ogień bucha/ młot dzielnie grucha/ Niby w piekle, w ogniu, w dymie/ pot kipi ze mnie/ wszystko z żelaza i stali/ gorąco pali”<sup>24</sup>. Wyraźnie bowiem to doświadczenie kontaktu z umaszynowanym, technicyzowanym światem nie skłania w planie wyrażania do poszukiwań nowych środków wyrazu, ale odwołuje się do głęboko osadzonych kulturowych wzorców definiowania grozy: potworów, smoków, płomieni oraz frazeologizmów budowanych wokół pojęcia „piekło”, a jednocześnie zachwyty nad wzniosłością owego „piękna maszynowego dzieła”<sup>25</sup>.

#### TARNOGÓRSKA „MASZYNA OGNIOWA” JAKO PARS PRO TOTO GÓRNOŚLĄSKIEGO UMASZYNOWIENIA

Powszechna obecność maszyn na obszarze „czarnej” części Górnego Śląska wynikała z dynamicznego rozwoju przemysłowego. Rozwój ten miał istotne znaczenie dla pruskiej gospodarki, ale też stał się, jak słusznie konstataowała Irena Bukowska-Floreńska, impulsem zmian społeczno-kulturowych i nową siłą kulturotwórczą, co sprawiło, że na Górnym Śląsku bardzo wyraźnie zaznaczyła się zmiana tradycyjnego układu między człowiekiem a przyrodą na modernizacyjną relację człowieka i maszyny. Dynamika tego procesu została zapoczątkowana w ostatnich dekadach osiemnastego wieku, jednak czas jej szczególnego natężenia przypadł na drugą połowę wieku dziewiętnastego i początek dwudziestego wieku<sup>26</sup>. Dominacja relacji człowiek-maszyna zredefiniowała wówczas niemal każdy wymiar codzienności, wpływając zwłaszcza na strukturę przestrzeni, znaczenie czasu i wzory kulturowe. Podobne procesy – co trzeba wyraźnie zaznaczyć – nie stanowiły szczególnego wyróżnika górnośląskiej rzeczywistości kulturowej, dotyczyły również Zagłębia Ruhry, Nord-Pas-de-Calais, północnych hrabstw Anglii i innych, europejskich regionów górniczo-przemysłowych. Można też podkreślić, że obecność i użytkowanie maszyn ani na Śląsku, ani w innych częściach Europy nie są charakterystycznym wyznacznikiem wyłącznie wieku pary i elektryczności,

<sup>23</sup> *Na Szlązku. Listy korespondenta „Kraju”, „Kraj” 35(1899), s. 102.*

<sup>24</sup> *Pieśni związane z pracą* (zebrane przez Łukasza i Stanisława Wallisów), Biblioteka Śląska w Katowicach, rps, sygn. R 2058 III.

<sup>25</sup> *Złota księga Tarnowskich Gór*, s. 33.

<sup>26</sup> Por. I. B u k o w s k a - F l o r e Ń s k a, *Uwarunkowania społeczne i kulturowe folkloru śląskiego (do 1939 r.)*, w: *Z dziejów i dorobku folklorystyki śląskiej (do 1939 r.)*, red. J. Pośpiech, T. Smolińska, Wydawnictwo Uniwersytetu Opolskiego, Opole 2002, s. 15.



ale to para i elektryczność po raz pierwszy zintensyfikowały na taką skalę zależność człowieka od maszyny, chociaż wciąż jeszcze wyłącznie w kontekście relacji operacyjnych i protezujących, ponieważ maszyny podporządkowane były wówczas przede wszystkim prawom mechaniki, a nie elektroniki czy informatyki, co ułatwiało ich instrumentalne traktowanie.

Dynamiczna industrializacja była przyczyną wzmożonego zainteresowania Górnym Śląskiem, który już od końca osiemnastego wieku zaczął wzbudzać ciekawość przede wszystkim ze względu na możliwość przyjrzenia się „czynnym zakładom pracy” (jak tego rodzaju aktywność nazywa się współcześnie), a także z uwagi na różnego typu ideologiczne pobudki, które wpływały na koncentrację uwagi na skutkach rozwoju przemysłu na Górnym Śląsku<sup>27</sup>. To doświadczenie spotkania z okolicami „ożywionymi przemysłem”, w których „wznoszą się po polach samotne domy fabryczne, kuźnie i warsztaty, i wysokie kominy parowych maszyn, i długie koszary robotników, a obok nich sterczą wszędzie mogiły rudy, żużla i węgla kamiennych”<sup>28</sup>, znalazło swoje odzwierciedlenie we wskazanych na początku artykułu rozmaitych tekstach epistolograficznych, chorograficznych, podróżniczych, sprawozdawczych czy memoratywnych. Nie zawsze kwestie związane z przemysłem, techniką czy umaszynowaniem stanowiły ich motyw przewodni, ponieważ podejmowały one zazwyczaj szerszą problematykę, dlatego na szczególną uwagę zasługuje unikalny dokument: księga gości (czy precyzyjniej księga meldunkowa) kopalni Friedrichsgrube w Tarnowitz (Tarnowskich Górach) zatytułowana *Fremdenbuch der Friedrichsgrube bei Tarnowitz*. Niestety oryginał dokumentu zaginął w trakcie działań wojennych na przełomie 1944 i 1945 roku. Istnieje co prawda duże prawdopodobieństwo, że jak wiele innych cennych archiwaliów został wywieziony bądź ukryty z obawy przed nadciągającą Armią Czerwoną, a nie, że został zniszczony, jednak jego losów do tej pory nie udało się ustalić. Na szczęście w 1913 roku przygotowano z okazji ogólnoniemieckiego zjazdu górniczego, który odbył się we Wrocławiu, okolicznościowe wydawnictwo *Festschrift zum XII. Allgemeinen Deutschen Bergmannstage in Breslau 1913*, do którego jako aneks dołączono odpis tarnogórskiej księgi pamiątkowej. Nadano jej wówczas tytuł *Das Goldene Buch von Tarnowitz (Złota księga Tarnowskich Gór)*. Reprint tego wydania wraz z polskojęzycznym tłumaczeniem i opracowaniem Marka Wrońskiego został opublikowany w 1994 roku (kolejne wydania ukazały się w 2009, 2011 i 2019 roku).

<sup>27</sup> Szerzej na ten temat zob. Kaczmarzyk, *Krajobraz, portret, maska*, s. 37-61 (rozdz. „U progu industrialnej inności. Determinanty kształtowania się wizerunku Górnego Śląska na przełomie XVIII i XIX w.”).

<sup>28</sup> W. P o l, *Z wycieczki*, „Biblioteka Naukowego Zakładu im. Ossolińskich” 2(847) nr 5, s. 545.



Księga ta zawiera ponad tysiąc wpisów (najwięcej z lat 1788-1835), poświadczających nie tylko kto odwiedził wówczas tarnogórską kopalnię i „podziwiał maszyny”, ale też, jakie refleksje, uczucia i emocje wzbudził ów kontakt ze sprowadzoną z Anglii maszyną parową, nazywaną wówczas „Feuremaschine” (maszyną ogniową). Pod koniec osiemnastego wieku była to poważna i odważna inwestycja, która wyznaczyła zupełnie nowy etap w rozwoju górnośląskiego przemysłu. Zainteresowanie wdrożeniem tej maszyny do tarnogórskiego procesu technologicznego sprawiło, że kopalnia stała się wizytówką pruskiego przemysłu państwowego. Z tego względu odwiedzali ją ówczesni wielcy tego świata, z Fryderykiem Wilhelmem II na czele, podziwiając ten – jak się wówczas wydawało – dowód absolutnego geniuszu człowieka. Kopalnia stała się dzięki temu celem pobytów studyjnych i kurtuazyjno-turystycznych wycieczek. Takie wycieczki organizowano, co ciekawe, nie tylko w „męskim gronie”, ale również dla pań lub w towarzystwie pań. Były wśród nich arystokratki z odległych miejsc w Europie, jak tajemnicza „Madame Barbarigo<sup>29</sup>, Noble Veniciene”<sup>30</sup>, najczęściej jednak kobiety zwiedzające kopalnię związane były z rodzącą się grupą śląskiej arystokracji „węgla i stali”. W 1790 roku w towarzystwie księżnej i niedawnej regentki Charlotte Amalie von Saxe-Meiningen, pracy maszyny parowej przyglądała się na przykład młodzianka wówczas naręczona księcia Hansa Heinricha VI von Hochberga – Anne Emilie księżna Anhalt Coethen. Potomkowie tego małżeństwa współtworzyli później przez kolejne sto pięćdziesiąt lat historię przemysłowego rodu książąt pszczyńskich, który miał ogromny wpływ na śląski przemysł aż do końca drugiej wojny światowej.

W pierwszym okresie maszynizacji Górnego Śląska wyraźnie dominowała maszynowa apologetyka. To, co można było zobaczyć w Tarnowskich Górach, było dla wielu odwiedzających (zwłaszcza ludzi niezwiązanych wprost z jakąkolwiek eksploatacją złóż naturalnych) doświadczeniem przede wszystkim nowości. Ta grupa ludzi, wkraczając w przestrzeń Friedrichsgrube, przyglądała się maszynie z jednej strony pewnemu curiosum, a z drugiej jako czemuś wyjątkowemu pod względem estetycznym, przykładając do swoich ocen kryteria podobne do oceny starożytności czy dzieł sztuki, z którymi niejednen z owych pierwszych tarnogórskich turystów spotkał się w czasie własnych Grand Tour

<sup>29</sup> Ród Barbarigo był możliwym rodem patrycjuszowskim, ważnym w dziejach Republiki Weneckiej, przedstawicielką tego rodu – Contarinę poznała w czasie swojego pobytu w Wenecji Katarzyna Płaterowa. Jej zdaniem Contarina Barbarigo wyróżniała się na tle nudnych i zajmujących się tylko plotkami weneccjanek (jednak sprawdzenie, czy to właśnie ona odwiedziła Tarnowskie Góry, wymagałoby przeprowadzenia badań archiwalnych wykraczających poza ramy tego artykułu). Por. K. Płaterowa, *Moja podróż do Włoch. Dziennik z lat 1785-1786*, oprac. M.E. Kowalczyk, tłum. A. Pikor-Półtorak, Wydawnictwo LTW, Łomianki – Dziekanów Leśny 2013, s. 76-78.

<sup>30</sup> *Złota księga Tarnowskich Gór*, s. 9.

po kontynencie europejskim. Jednak ani wówczas, ani współcześnie maszyna i dzieło sztuki nie są w perspektywie ocen estetycznych tożsame.

Maszyny nie są bowiem – jak dzieła sztuki – niepowtarzalne. Ich istotą i siłą jest seryjność i dlatego warto zwrócić uwagę również na wypowiedź anonimowego kosmopolity z końca osiemnastego wieku. Autor *Schilderungen oder Reisen eines Kosmopoliten. Reise von...n durch Polen und Schlesien in den Jahren 1792 und 93* (*Opowiadania albo podróży jednego Kosmopolity. Podróże po Polsce i Śląsku w latach 1792 i 1793*), aczkolwiek z dystansem podchodził do obiegowych opinii na jakiś temat, uznał, że musi zobaczyć Tarnowskie Góry, ponieważ chciał się przekonać, ile prawdy jest w słowach „nie widział Górnego Śląska, kto nie był w zakładach w Tarnowskich Górach”<sup>31</sup>. To, co dostrzegł, to przede wszystkim niezwykle zawężający poznawczo sens tak sformułowanego zdania, wynikający jednak z typowości i powtarzalności przemysłowego świata, które sprawiały, że w różnych miejscach w Europie zobaczyć można było wówczas właściwie to samo, dlatego jego zdaniem „Górny Śląsk można tak samo dobrze poznać w Londynie, lub, jeśli kto woli bliżej w Altensalz bądź w Rothenburgu, tzn. w każdym miejscu, gdzie, jak w tarnogórskich zakładach, pracuje jedna lub kilka maszyn parowych”<sup>32</sup>. Kosmopolita nie należał do ludzi zafascynowanych technicznym postępem, raczej skomplikowane i niezrozumiałe dla laika działanie maszyny parowej skłaniało go do refleksji nad istotą ludzkiego rozumu, który skoncentrowany na „błahostkach hiperfizycznych” nie jest już zdolny do zajmowania się rzeczami prawdziwie „wielkimi i wzniosłymi”. I na pewno wzniosłości tej nie ujmował w kategoriach, które pozwalałaby wykorzystać jego relację w poszukiwaniach egzemplifikacji wzniosłości technologicznej<sup>33</sup>.

Nie wiadomo, kto jest autorem *Schilderungen oder Reisen eines Kosmopoliten*, jednak – zakładając, że jest to ktoś, kto rzeczywiście odwiedził Tarnowskie Góry – to być może również jego podpis widnieje w księdze pamiątkowej, są bowiem również takie wpisy, które ograniczają się wyłącznie do podpisu. Nie jest więc wykluczone, że w tej grupie należałoby szukać odpowiedzi na pytanie o tożsamość tajemniczego Kosmopolity<sup>34</sup>. Aczkolwiek nie ulega wątpliwości, że wśród wielu z tych, którzy z różnych powodów przybywali do

---

<sup>31</sup> *Opowiadania albo podróże jednego Kosmopolity. Podróż po Polsce i Śląsku w latach 1792 i 1793*, w: *Trzy relacje z podróży po Śląsku w XVIII i XIX wieku*, oprac. J. Szymański, tłum. L. Szybkowski, Dom Współpracy Polsko-Niemieckiej; Fundacja Zamek Chudów, Opole–Katowice 2006, s. 80.

<sup>32</sup> Tamże, s. 80.

<sup>33</sup> Zob. L. M a r x, *The Machine in the Garden. Technology and the Pastoral Ideal in America*, Oxford University Press, New York 1964 (<http://wtf.tw/ref/marx.pdf>).

<sup>34</sup> Próba potwierdzenia bądź obalenia tej hipotezy wymagałaby szczegółowych badań archiwalnych, które nie mają dla poruszanej w artykule problematyki istotnego znaczenia.

Friedrichsgrube, zainstalowana w kopalni maszyna parowa nie wzbudzała żadnych ambiwalentnych uczuć, a jedynie wywoływała podziw.

Oczywiście pewien wpływ na apologetyczny ton wpisów miało zapewne to, że była to księga gości (księga meldunkowa). Złamaniem konwencji byłoby więc umieszczanie uwag krytycznych, jednak mimo to, zwłaszcza bardziej rozbudowane wpisy odzwierciedlają fakt, że podważanie instrumentalnej roli „maszyny ogniowej” najwyraźniej nie przychodziło nikomu na myśl. Hierarchia człowiek–maszyna nie wydawała się tym samym w żaden sposób zaburzona. To człowiek w przekonaniu ówczesnych przybywających do Friedrichsgrube (zarówno w celach złożenia oficjalnej wizyty, jak i ze względów zawodowych czy też z powodów wyłącznie turystycznych) sprawował władzę i to człowiek, dzięki swojej sile oraz przemyślności był w stanie zapanować przede wszystkim nad wodą – żywiołem od wieków skutecznie uniemożliwiającym eksploatację podziemnych złóż. Władza ta nie miała jednak charakteru absolutnego. W tarnogórskiej księdze pamiątkowej znajdują się również wpisy jasno określające providencjalistyczny charakter funkcjonowania owej umaszynowanej rzeczywistości. Tarnogórski pastor, Emanuel Pohle, wprost pisał o Opatrzności, która w śląską ziemię włożyła swoje błogosławieństwo<sup>35</sup>, a bliżej nieznany inspektor Naglo z huty Malapane (w dzisiejszym Ozimku) podkreślał wcześniej, że „Wszystko od błogosławieństwa Bożego zależy”<sup>36</sup>. W podobnym tonie wpisał się późniejszy słynny polski kaznodzieja ewangelicki w Cieszynie Jan Kłapsia (Johann Klapsia), który wracając z Dolnego Śląska do Jaworza, gdzie pełnił posługę duszpasterską, odwiedził Friedrichsgrube. Ujrzone tam „maszyny ogniowe” skłoniły go to do refleksji, „jak duże jest bogactwo uszczęśliwiającej ludzi mądrości i dobroć Wszechojca oraz jak prawdziwe jest to, że człowiek, arcydzieło stworzenia, jest boskiego pochodzenia, potrafi bogactwo tego błogosławieństwa Bożego nie tylko odnaleźć, lecz także do doskonalenia swojego pokolenia spożytkować”<sup>37</sup>.

Z tego punktu widzenia zamontowanie maszyny, a później maszyn<sup>38</sup>, które chroniły kopalnię przed zalaniem, a także stosowanie coraz bardziej skompli-

<sup>35</sup> Por. *Złota księga Tarnowskich Gór*, s. 7.

<sup>36</sup> Tamże, s. 5.

<sup>37</sup> Tamże, s. 33.

<sup>38</sup> Wydajność tej maszyny po dwóch latach musiała zostać zwiększona poprzez zakup kolejnej, również angielskiej, później maszyny produkowane były w Gliwicach i Ozimku. Pierwsza maszyna parowa z Tarnowskich Gór pracowała przez siedemdziesiąt lat, jej ostatnim zadaniem było gaszenie największego na Śląsku w dziewiętnastym stuleciu pożaru podziemnych złóż węgla w kopalni Fanny w Michałkowicach. Obecnie prowadzone są prace eksploracyjne, których zadaniem jest odnalezienie śladów związanych z funkcjonowaniem tej pierwszej maszyny w Tarnowskich Górach (zob. np. A. F r u ż y ń s k i, *Fryderyk August Holtzhausen – pierwszy budowniczy maszyn parowych na Śląsku*, „Hereditas Minaniorum” 2014, nr 1, <http://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-8b9064a2-3453-4ccc-8120-7b2e2c32beeb>).

kowanych urządzeń wspomagających pracę górników, rodziło poczucie czy może nawet przekonanie, że maszyny w relacji człowiek–natura stoją zawsze po stronie człowieka, nie naruszają jego podmiotowości, ale są dowodem siły jego rozumu, dzięki któremu potrafi on nad siłami przyrody zawsze zapanować. Chęć pokonania żywiołów była odwiecznym pragnieniem człowieka, stąd działanie maszyn, które w górnictwie (i innych przemysłowych zakładach) zaczęto stosować, należało poznać, a później spróbować zrozumieć, na czym polega ich znaczenie nie tylko dla tarnogórskiej kopalni, ale Śląska w ogóle. Osoby z maszynami najbliższe związane, przede wszystkim inwestorzy (mówiąc językiem współczesnej ekonomii), którzy zaryzykowali spore sumy pieniędzy na potrzeby procesu technologicznego, ale też wykwalifikowani robotnicy obsługujący poszczególne urządzenia, zapewne byli bardzo dumni ze swojej odwagi i zręczności. I może dlatego tak wiele kontrowersji wzbudził wpis Johanna Wolfganga Goethego, który przybył w 1790 roku do Tarnowskich Gór z racji swoich ministerialnych obowiązków i zainteresowań mineralogią. Goethe napisał, że oto znalazł się: „z dala od wykształconych ludzi, na krańcach rzeszy”<sup>39</sup>, chociaż najprawdopodobniej chciał w ten sposób wyrazić uznanie dla tego, co na owych „krańcach państwa” (niem. am Ende des Reiches) zobaczył, nawet jeżeli było to podszyte mocno protekcjonistyczną nutą<sup>40</sup>.

Nie każdy z odwiedzających miał jednak – jak Goethe – na celu poszerzenie swojej wiedzy technologicznej ze względu na jakieś pragmatyczne okoliczności. Wiele osób przybywało do zakładu, ponieważ stanowił on rodzaj turystycznej atrakcji. „Byłem tu, by obejrzeć maszynę i swoje nazwisko Gottl. Woida zanotowałem” – napisał w marcu 1792 roku lakonicznie, ale dobitnie jeden z gości<sup>41</sup>, a inny – Ignatz Scholz, Królewski Komisarz Sądowy z Brzegu, przypomniał sobie „w czasie oglądania maszyn ogniowych urywek z Szekspira (!), że istnieją na niebie, i na ziemi rzeczy, o których się naszym filozofom nie śniło”<sup>42</sup>.

## ROBOTNICZA SYMBIOZA Z MASZYNĄ

Analiza wpisów w *Złotej księdze Tarnowskich Gór* oraz innych opisów czy relacji poświęconych górnośląskiemu przemysłowi dowodzi, że negatywnych bądź malkontenckich opinii na temat postępującej w tym regionie

<sup>39</sup> *Złota księga Tarnowskich Gór*, s. 23.

<sup>40</sup> Tamże, s. 132.

<sup>41</sup> Tamże, s. 33.

<sup>42</sup> Tamże.

maszynizacji było zdecydowanie mniej, zwłaszcza, że rola maszyn w śląskiej rzeczywistości kulturowej wzrastała i nie chodziło o to, że Friedrichsgrube była warta odwiedzenia przestrzenia, którą pruska władza mogła się szczyścić, czy też, która mogła kogoś zaintrygować. Świadomość znaczenia, jakie przemysł odegrał w rozwoju Śląska, sprawiała, że wracano do pamiętnego roku 1788, uznając, że jest to wydarzenie warte przypomnienia. Z tej okazji opracowano w 1913 roku *Złotą księgę Tarnowskich Gór*, dzięki czemu cenne zapisy tarnogórskiej księgi gości w znakomitej większości zostały ocalone<sup>43</sup>, a Józef Piernikarczyk, śląski historyk i działacz kulturalno-oświatowy przypomniał o roli tej pierwszej maszyny parowej w okolicznościowym artykule opublikowanym w roku 1938<sup>44</sup>.

Jednak znaczenie uprzemysłowienia Górnego Śląska, którego maszynizacja jest integralnym elementem, nie wiązało się jedynie z przekształceniem peryferyjnego w gruncie rzeczy obszaru w region o istotnym znaczeniu ekonomicznym, ale polegało również na szerokim wpływie kulturowym i społecznym. Wokół tarnogórskich maszyn parowych i innych umaszynowanych zakładów pracy, w tym kopalń węgla kamiennego, skupiała się bowiem grupa wykwalifikowanych robotników. Potrzeba istnienia takiej grupy i jej poszerzenia była na tyle znacząca, że Friedrich Anton von Heynitz w swoim memoriale o stanie przemysłu na Śląsku z 1801 roku dobitnie wskazywał, że brak wykwalifikowanych robotników w górnictwie i hutnictwie stanowi jeden z głównych czynników utrudniających dalszy rozwój gospodarczy. Heynitz podkreślał, że nawet dziedziczenie zawodu i zatrudnianie młodych ludzi, których ojcowie już w przemyśle pracowali, jest wciąż zbyt wolne<sup>45</sup>.

Jednak w miarę powstawania kolejnych zakładów na Górnym Śląsku następowały zmiany demograficzne, w przemysłowych osadach osiedlało się coraz więcej ludzi, nieobawiających się pracy obok maszyn, z maszynami i przy maszynach. Pozwoliło to kilkadziesiąt lat później reporterowi niezwykle poczytnego ówczesnie w krajach niemieckojęzycznych pisma „Die Gartenlaube” (będącego protoplastą licznych odmian współczesnej prasy popularnej), w swojej relacji z Górnego Śląska entuzjastycznie pisać, o tym, jak znakomicie

<sup>43</sup> Należy pamiętać, że *Złota księga Tarnowskich Gór* nie była wydaniem krytycznym, a publikacją okolicznościową, poza tym zostały w niej zamieszczone informacje, że niektóre wpisy zostały pominięte.

<sup>44</sup> Zob. J. Piernikarczyk, *W 150 lat po uruchomieniu pierwszej maszyny parowej na Górnym Śląsku*, [http://sbc.org.pl/Content/158298/ii147714-1937\\_1939-12-0001.pdf](http://sbc.org.pl/Content/158298/ii147714-1937_1939-12-0001.pdf). Również współcześnie pamięć ta jest kultywowana: z inicjatywy Stowarzyszenia Miłośników Ziemi Tarnogórskiej od roku 2019 organizowane jest cykliczne wydarzenie kulturalne na terenie Skansenu Maszyn Parowych przy Zabytkowej Kopalni Srebra w Tarnowskich Górach – „Święto Pary” (zob. *Święto pary już po raz trzeci*, Tarnowskie Góry, <https://tarnowskiegory.pl/2023/10/swieto-pary-juz-po-raz-trzeci/>).

<sup>45</sup> Zob. A.J. Wójcik, *Fossores ex Polonia. Rekrutacja polskich górników do Prus*, „Hereditas Minariorum” 2017, nr 4, <https://www.dbc.wroc.pl/dlibra/publication/140293/edition/72368/content>.

górnoląscy robotnicy radzą sobie w tym coraz bardziej zmechanizowanym świecie. Relacjonował on: „Większość owych robotników to urodzeni Górnoślązacy, i mój przyjaciel zapewnił mnie, że ów tak osławiony lud posiadał szczególny spryt i zręczność. Weź górnośląskiego chłopca, pokaż mu maszynę parową, wytłumacz, na jakiej zasadzie działa i jak ją należy obsługiwać, on to od razu pojmie, zbada w mig mechanizm i już po dwudziestu czterech godzinach będzie potrafił ją sprawnie obsługiwać. Tak wielki talent ma on także do każdej innej technicznej pracy. Nie świadczy to wcale o jego ograniczonych zdolnościach do działalności umysłowej, brakuje mu tylko okazji, aby wykształcić i rozwijać takowe zdolności”<sup>46</sup>. Tarnogórskie „maszyny ogniowe” stały się tym samym kamieniem węgielnym procesu, który wyznaczył dalszy kierunek rozwoju Górnego Śląska. Wincenty Pol, który w połowie dziewiętnastego wieku podróżował przez śląskie ziemie, zachwycony kolejnym etapem industrialnego rozwoju – koleją, pisał, używając nieprzypadkowo maszynowej metafory, że „Tu wypada być tym tłokiem, co w ruch wprawia wóz parowy, a nie tym drągiem, co w szprychach płynącego taboru zawadza”<sup>47</sup>. Ta konkluzja wydaje się być kwintesencją pierwszego okresu w kształtowaniu na Górnym Śląsku relacji między człowiekiem a maszyną.

Pierwsze maszyny parowe nie były jeszcze, co należy podkreślić, odhumanizowane, ich opisy wskazują przede wszystkim na ową operacyjną i protezową funkcję w stosunku do człowieka, a także w stosunku do konia, którego siła była dotąd największym człowiekiem sprzymierzeńcem (jednostkowa moc konia stała się przelicznikiem mocy maszynowej – konia mechanicznego). W Tarnowskich Górach przelicznik opłacalności wskazywał, że co prawda urządzenie kosztowało więcej niż całoroczny zarobek wszystkich pracowników kopalni, jednak wydatki ponoszone później na jego utrzymanie (trzy tysiące siedemset talarów rocznie) były i tak trzykrotnie niższe, niż pieniądze przeznaczone do tej pory na funkcjonowanie odwadniarek konnych, czyli utrzymanie sto dwadzieścia koni, dzięki którym kieraty funkcjonowały dwadzieścia cztery godziny na dobę<sup>48</sup>. Zastosowanie maszyn sytuowało człowieka – co z punktu widzenia rozwoju ekonomii jest oczywiste – jako czynnik bilansu zysków i strat przedsiębiorstwa.

Konsekwencją pracy w zmechanizowanym środowisku było wykształcenie wśród pracowników górnośląskich zakładów przemysłowych umiejętności, które dopełniały siłę fizyczną, a polegały na sprawności wykonywania określonych czynności w określonym czasie i określonym sposobie współ-

<sup>46</sup> Cyt. za: A. H a l o r, *Przewodnik siemianowicki. Wędrówki po mieście i okolicy*, nakł. Urzędu Miejskiego, Siemianowice Śląskie 2000, s. 28.

<sup>47</sup> W. P o l, *Z wycieczki*, „Biblioteka Naukowego Zakładu im. Ossolińskich” 2 (1847), nr 5, s. 551.

<sup>48</sup> Zob. *Reaktywacja górnictwa*, Zabytkowa kopalnia srebra i sztolnia Czarnego Pstrąga, <https://kopalniasrebra.pl/historia/>.



działania z innymi, a czasem z racjonalizatorską inwencją. Z tego względu tylko do pierwszych maszyn sprowadzono części zapasowe z Anglii, później z tych części w Gliwicach i Ozimku składano maszyny już na Górnym Śląsku. Po dekadzie nie było to już potrzebne, ponieważ od 1798 roku maszyny konstruowano w całości na miejscu. Józef Piernikarczyk przytacza historię (oczywiście trudno określić, na ile anegdotyczną) chłopca, dzięki któremu udało się działanie maszyny zrationalizować. Pisał on: „Należy dodać, że pierwsze maszyny nie były jeszcze zaopatrzone w tak zwane kulisy. Musiano dla ciągłego ruchu maszyn utrzymywać do ich obsługi chłopców z okolicy. Zauważono jednak kilkakrotnie, że pewnego chłopca przy maszynie nie było, a maszyna mimo to szła. Dziwiono się temu i postanowiono to zbadać. Obserwowano chłopca z ukrycia i przekonano się, że gonił za ptakami i aby zaś maszyna nie stanęła, wynalazł sobie sam jakiś prosty, automatyczny sposób jej regulacji. Chłopak miał się w ten sposób przyczynić do wynalezienia tak zwanych kulisów przy maszynie”<sup>49</sup>. Ta historia (Piernikarczyk nie wskazuje jej źródła) wpisuje się w szerszą narrację o wynalazczości jako *enfant prodige*, odkryciu stanowiącym konsekwencję przypadkowego w gruncie rzeczy działania, którego w odniesieniu do maszyny parowej miał według biograficzno-technologicznej legendy dokonać sam James Watt, obserwując, jak para z czajnika podnosi pokrywkę<sup>50</sup>.

W początkach kształtowania się relacji człowiek-maszyna takie racjonalizatorskie udoskonalenia tylko potwierdzały wiodącą rolę człowieka w tym procesie i to człowieka trochę w rozumieniu Everymana, bo przecież każdy (jak ów anegdotyczny-legendarny chłopiec z Tarnowskich Gór) może usprawnić działanie maszyny. Oczywiście takie możliwości były udziałem nielicznych, natomiast znacznie częściej kontakt z maszyną prowadził do wykształcenia się określonych cech czy rozwoju predyspozycji, które pozwalały sprawnie funkcjonować na danym stanowisku pracy. Wiązały się one z tym, że niezbędna była sprawność obsługiwanie różnych mechanizmów tak, aby nie zakłócić ich pracy. Maszyna jest bowiem wymagająca, żąda uwagi i nie toleruje błędów. Błąd pociąga za sobą konsekwencje. Mechanizm nadaje rytm i wymaga określonego działania, jakakolwiek pomyłka, niewłaściwa kolejność działań czy nieuprawniona ingerencja prowadzi w najlepszym razie do awarii, a w najgorszym do wypadku, którego ofiarą jest człowiek. Tym samym maszyny ujawniać zaczęły może jeszcze nie władzę, ale na swój sposób kontrolę

<sup>49</sup> Piernikarczyk, *W 150 lat po uruchomieniu pierwszej maszyny parowej na Górnym Śląsku*, s. 5.

<sup>50</sup> Por. M. Litwinowicz - Drożdziel, *Wiek XIX. Wynalazek. Wstępne rozpoznanie*, w: *Praktyka, utopia, metafora. Wynalazek w XIX wieku*, red. J. Kubicka, M. Litwinowicz-Drożdziel, Uniwersytet Warszawski – Convivo Anna Matusiak, Warszawa 2016, s. 7-21.



nad rzeczywistością, mogły bowiem stać się siłą sprawczą, odmieniającą życie człowieka i wyznaczającą niejednokrotnie tragiczną cezurę „przed i po”.

### MASZYNA JAKO ŹRÓDŁO ZAGROŻENIA

Patrząc z perspektywy zakładu przemysłowego, który stał się symbolem industrializacji Górnego Śląska – kopalni – rola, jaką w możliwościach eksploataowania coraz głębszych pokładów węgla kamiennego odgrywały maszyny, jest ogromna. Bez ich zastosowania rozwój tej gałęzi przemysłu w ogóle nie byłby możliwy. Jednak ich zastosowanie na szerszą skalę miało (i ma) również swoją ciemną stronę związaną z warunkami i konsekwencjami ich użytkowania. Najpoważniejszą z tych konsekwencji jest wypadkowość. Paradoksalnie bowiem podatność na wypadki (również współcześnie) wynika z braku maszynizacji w ogóle albo tego, że jakiś umaszynowany element zawiódł.

Kopalnia zawsze była miejscem niebezpiecznym. W kopalniach zdarzały się wypadki, chociaż ich liczba nie była wbrew pozorom duża. W połowie dziewiętnastego wieku średnio na tysiąc zatrudnionych w górnictwie węglowym śmierć na stanowisku pracy ponosiło dwóch, trzech górników<sup>51</sup> i nawet, jeżeli przyjąć, że dane te są niepełne, to jednak nie sposób nie zauważyć, że to postęp technologiczny, udoskonalanie sposobów wydobywania pociągnęło za sobą wzrost zagrożenia wypadkiem. W 1913 roku w broszurze: *O czym powinien pamiętać każdy górnik* zamieszczono dramatyczne ostrzeżenie: „każdy górnik może się spodziewać, że raz na trzy lata spotka go wypadek”<sup>52</sup>. Przyczyną wielu takich dramatycznych zdarzeń były tak zwane maszyny pomocnicze<sup>53</sup>, których człowiek był operatorem i które protezowały siłę jego mięśni. Rozwój technologiczny, stosowanie coraz bardziej skomplikowanych urządzeń ułatwiało bowiem pracę, ale jednocześnie rodziło nowe zagrożenia.

Niebezpieczeństwo pracy w kopalniach zawsze związane było (i jest) z działaniem sił natury. W górnictwie kruszcowym wiązało się ono z siłą wód podziemnych, w górnictwie węglowym pojawiło się ponadto zagrożenie metanowe oraz wybuchem pyłu węglowego. Wypadki powodowane były wyrzutami gazów, tąpnięciami i zawałami, będącymi następstwem procesów zachodzących w górotworze. Równie niebezpieczne były pożary (choć najczęściej spowodowane przez człowieka). Górnicy mieli świadomość tych

<sup>51</sup> Por. J. J a r o s, *Zarys dziejów górnictwa węglowego*, Państwowe Wydawnictwo Naukowe, Warszawa–Kraków 1975, s. 82.

<sup>52</sup> D. Ś w i t a ł a - T r y b e k, *Tragedia w kopalni. Kulturowe konteksty katastrof i wypadków górniczych*, Wydawnictwo Uniwersytetu Opolskiego, Opole 2014, s. 117.

<sup>53</sup> Por. J. P i e r n i k a r c z y k, *Historia górnictwa i hutnictwa na Górnym Śląsku*, t. 2, nakł. Śląskiego Związku Akademickiego, Katowice 1936, s. 41.

zagrożeń, przy czym zerwanie liny czy osunięcie się na drabinach w szybie zwykle oznaczało jednostkową śmierć, dopiero zastosowanie różnorodnych mechanicznych urządzeń na większą skalę paradoksalnie zwiększyło wypadkowość i liczebność ofiar. Wiele urządzeń pomocniczych zastosowano w celu podniesienia efektywności wydobywania (bezpieczeństwo było zdecydowanie, mówiąc oględnie, na drugim planie). W takiej sytuacji, jeżeli zawiodły maszyny odpowiedzialne za systemy wentylacyjne czy wyciągowe, katastrofa była nieunikniona, jak na przykład w 1898 roku w ówczesnej kopalni Paulus w osadzie Morgenroth (dzisiaj Chebziu – dzielnicy Rudy Śląskiej) z powodu wyslizgnięcia się kopalnianej szoli i awarii łapaczów, klatka z górnikami runęła do szybu, grzebiąc dwadzieścia pięć osób<sup>54</sup>. Nie należały do rzadkości również wybuchy kotłów parowych.

Takie doświadczenia i świadomość, że mogą się one wydarzyć, miały wpływ na kształtowanie się wzoru kulturowego górnika (który na zasadzie *pars pro toto* stał się zarazem mocno zmityzowanym, zwłaszcza w czasach współczesnych, wzorem Ślązaka w ogóle). Z pewnym patosem, ale też przebijającym z tej wypowiedzi emocjonalnym i znanym z autopsji związkiem z industrialną rzeczywistością, mówił o tym w swoim odczycie Gustaw Morcinek, podkreślając, że górnicy i hutnicy to ludzie, którzy w swojej pracy, na co dzień „kumają się ze śmiercią” i dlatego „jeżeli co dnia patrzeć w jej oczy ślepiące z głębi zapadłego wyrobiska, czy spod spękanego stropu, czy też wyczuwać jej tchnienie w obrębie ryczącej strugi płynnego żelaza czy śmigających po zadytmionej hali czerwonym węży stalowych – to nie dziwota, że człowiek nie może być innym. Hardziej w jej towarzystwie, wyzbywa się wszelkiej ckliwości, zamienia się w srogięgo i zawziętego zapaśnika, co musi umieć przechytrzyć podłe zamiary czającej się śmierci, a potem wykpić ją serdecznie”<sup>55</sup>.

Tym samym w kontakcie, czy może nawet starciu z jednej strony z siłami natury, a z drugiej z maszynami, człowiek nie zawsze miał poczucie sprawczości i dominacji, zwłaszcza że maszyny z pozycji pomocnika, dzięki któremu można było zyskać przewagę nad siłami natury, mogły zmienić się niepostrzeżenie w nieprzewidywalnego przeciwnika, swoisty żywy organizm. Tak właśnie określił kopalnię współcześnie jeden z górników: „Kopalnia to był żywy organizm. U góry te budynki stały jak stały. Ale ten dół się zmieniał. Dzisiaj jestem w pracy, jutro przychodzę i już nie ma tego miejsca. Już jest parę metrów dalej to stanowisko pracy. To się zmieniało, to chodziło. Cały czas coś nowego. Coś się dzieje: tam się woda pokazała, tam się metan..., wentylacja znowu. I to cały czas trzeba było dopracowywać. Prąd cały czas musiał być doprowadzony, powietrze musiało być doprowadzane, tory doprowadzane [...]”.

<sup>54</sup> Por. tamże, s. 110-113.

<sup>55</sup> Tamże, s. 105.

Żywy organizm, żywy, to żyło [...]. Nie wiedziało się, czy się zjedzie, czy się nie zjedzie”<sup>56</sup>.

Beata Piecha van Schagen, realizując projekt dotyczący współczesnych narracji górniczych, pisała, że „[...] górnicy patrzą na kopalnię przez pryzmat bycia, a więc ma ona wartość ontyczną. Doświadczenie owej wartości jest przez górników budowane emocjonalnie, rozumowo i chyba nawet duchowo, co powoduje, że są oni w stanie w oparciu o nią skonstruować swoją tożsamość. A więc bycie-w-kopalni ustanawia ich bycie w ogóle i pozwala im na ukonstytuowanie tożsamości górniczej”<sup>57</sup>. Ukształtowanie się na Górnym Śląsku takiej tożsamości (mimo że z oczywistych względów nie jest to już współcześnie tożsamość dominująca) nie byłoby możliwe, gdyby w owym pojedynku z naturą człowiek nie został wsparty przez maszyny właśnie. Mimo bestiaryjnych konotacji i piekielnych analogii oraz rzeczywistych zagrożeń, intensyfikacja tych kontaktów sprawiła, że przemysł (i związana nim maszynizacja) stały się jednym z ważniejszych czynników górniczego (śląskiego) etosu. Jego ukształtowanie się przypadło jednak na czas, kiedy maszyny wciąż jeszcze były na miarę człowieka, chociażby dlatego, że jak w ramach tego samego projektu powiedział inny, emerytowany górnik, dzisiaj już nie ma takiej współpracy i zaangażowania, skończyło się. „Kiedyś były obudowy indywidualne, zmechanizowane, ale troszeczkę inna była procedura wydobywania. Inne maszyny. W tej chwili na dole nie można nic zrobić bez maszyn, bez urządzeń. Kiedyś [...] na przenośniku szkatowym spalił się silnik, silnik nieduży, ważył gdzieś osiemdziesiąt, sto kilo. – No sztygarze, to jo polece i przyniese. Na plecach go przyniósł. W tej chwili taki silnik waży półtorej, dwie tony [...]. Kiedyś na indywidualnej obudowie, na takiej ścianie pracowało dwudziestu, dwudziestu pięciu górników. W tej chwili pracuje czterech. I jeden drugiego przez całą dniówkę praktycznie nie widzi. Nie ma tej współpracy. A przedtem jeden drugiemu pomógł, usiedli, porozmawiali, i tam różne kawały, nie kawały. I tym była właśnie ta więź i ta współpraca”<sup>58</sup>.

\*

W górnictwie, podobnie jak w wielu innych miejscach, etyka pracy opartej na współpracy ustąpiła współcześnie bez wątpienia myśleniu korporacyj-

<sup>56</sup> Cyt. za: B. Piecha van Schagen, „Kopalnia to miłość”. *Szkic o tożsamości*, w: *Narracje górnicze z terenu Zabrze. Kopalnia to je do mnie wszystko*, red. B. Linek, Muzeum Górnictwa Węglowego w Zabrzu, Zabrze 2016, s. 62.

<sup>57</sup> Tamże, s. 62.

<sup>58</sup> Tamże, s.74.

nemu<sup>59</sup>. Sprzyja temu eliminowanie konieczności wzajemnych społecznych interakcji, będące jedną z konsekwencji kolejnych etapów umaszynowania (a zwłaszcza cyfryzacji) kulturowej rzeczywistości. Wydaje się, że mimo swojej „apokaliptyczności” wcześniejszy, przedcybernetyczny etap maszynizacji (zarówno w perspektywie górnośląskiej, jak i w szerszym, może nawet globalnym wymiarze), mimo obaw, które budził, z perspektywy czasu aż tak niebezpieczny dla istoty bycia człowiekiem się nie okazał. Trzeba też pamiętać, że paradoksalnie każdy wynalazek i każda zmiana zawsze ma swoich entuzjastów, ale też budzi reakcje zupełnie nieprzystające do skali zagrożenia. Wincenty Pol, pisząc o kolei żelaznej przytoczył anegdotę o Kozaku, który znajdował się w służbie i kiedy w ramach tej służby zmuszony był wsiąść do pociągu, uznał, że po powrocie do domu będzie się musiał spowiadać, że „nas ono wozilo po świecie, bo dla chrześcijańskiej duszy nie godzi się tem jeździć, pannuńciu”<sup>60</sup>. Czy nostalgia za przedcybernetyczną hierarchią w relacji człowiek–maszyna, kiedy nie była ona jeszcze tak silnym konkurentem człowieka, nie naruszała jego podmiotowości, okaże się równie anachroniczna jak strach owego Kozaka jadącego pociągiem do Belgii? Czy wręcz przeciwnie dopiero współczesne (i przyszłe) maszyny okażą się prawdziwie apokaliptyczną bestią i zmienią rzeczywistość nie tylko Górnego Śląska? Nie sposób oczywiście na te pytania odpowiedzieć inaczej niż w duchu profetycznych wizji, chociaż nie sposób też nie zauważyć, że rozmaite przesłanki coraz dalej idących konsekwencji maszynizacji świata, a co za tym idzie władzy technologicznych systemów nie wydają się być całkiem optymistyczne.

#### BIBLIOGRAFIA / BIBLIOGRAPHY

- Bukowska-Floreńska, Irena. “Uwarunkowania społeczne i kulturowe folkloru śląskiego (do 1939 r.)” In *Z dziejów i dorobku folklorystyki śląskiej (do 1939 r.)*. Edited by Jerzy Pośpiech, Teresa Smolińska. Opole: Wydawnictwo Uniwersytetu Opolskiego, 2002.
- Chojcka, Ewa, et al. *Sztuka Górnego Śląska od średniowiecza do końca XX wieku*. Katowice: Muzeum Śląskie, 2004.
- Ellul, Jacques. *The Technological System*. Translated by Joachim Neugroschel. New York: Continuum, 1980.
- Frużyński, Adam. “Fryderyk August Holtzhausen – pierwszy budowniczy maszyn parowych na Śląsku.” *Hereditas Minaniorum*, no. 1 (2014). <http://yadda.icm.edu.pl/baztech/element/bwmeta1.element.baztech-8b9064a2-3453-4ccc-8120-7b2e2c32beeb>.

<sup>59</sup> Por. tamże, s.74.

<sup>60</sup> P o l, *Z wycieczki*, s. 557.

- Das Goldene Buch von Tarnowitz (Fremdenbuch der Friedrichsgrube von 1788 ab.): Ehrentafel; Verzeichnis der freiwilligen Beiträge der Berg- und Hüttenleute 1813/14 nebst Begleitbriefen.* Breslau 1913.
- Eco, Umberto, ed. *Historia piękna*. Translated by Agnieszka Kuciak. Poznań: Dom Wydawniczy Rebis, 2005.
- Górny Śląsk i Zagłębie w dawnych opisach. Wiek XIX.* Edited by Andrzej Zieliński. Katowice: Wydawnictwo Śląsk, 1984.
- Halor, Antoni. *Przewodnik siemianowicki: Wędrówki po mieście i okolicy*. Siemianowice Śląskie: Urząd Miejski, 2000.
- Hytrek, Adolf. “Górny Szląsk pod względem obyczajów, języka i usposobienia ludności.” Edited by Bernard Kalicki. *Przegląd Polski* 14, no. 3 (1879): 291–319.
- Ingarden, Roman. *Szkice z filozofii literatury*. Łódź: Spółdzielnia Wydawnicza Polonista, 1947.
- Jaros, Jerzy. *Historia kopalni Król w Chorzowie (1791-1945)*. Katowice: Wydawnictwo Górnictwo-Hutnicze, 1962.
- . *Zarys dziejów górnictwa węglowego*. Warszawa-Kraków: Państwowe Wydawnictwo Naukowe, 1975.
- Kaczmarzyk, Izabela. *Krajobraz, portret, maska. Studia o górnośląskim imaginariu przemysłowym*. Kraków: Wydawnictwo Naukowe Akademii Ignatianum w Krakowie, 2019.
- Litwinowicz-Drożdziel, Małgorzata. “Wiek XIX. Wynalazek. Wstępne rozpoznanie.” In *Praktyka, utopia, metafora. Wynalazek w XIX wieku*. Edited by Joanna Kubicka, Małgorzata Litwinowicz-Drożdziel. Warszawa: Uniwersytet Warszawski and Convivo Anna Matusiak, 2016.
- Lompa, Józef. *Krótki rys jeografii Szląska dla nauki początkowej*. Głogówek: Henryk Handl, 1847.
- Łaszczycza, Piotr. “Człowiek i jego maszyny: Operatorzy i protezy.” *Filo-Sofija*, no. 39 (2017): 49–64.
- Marx, Leo. *The Machine in the Garden. Technology and the Pastoral Ideal in America*. New York: Oxford University Press, 1964. <http://wtf.tw/ref/marx.pdf>.
- Melosik, Zbyszko. “Technologizacja życia i tożsamości w kulturze współczesnej.” *Studia Edukacyjne* 38 (2016): 43-59. <https://repozytorium.amu.edu.pl/handle/10593/14837>.
- “Na Szląsku i w Styrii. Z notat inżyniera Cz. [spisał K. Saryusz].” *Wędrowiec* 3, series 3, no. 53 (1878), <https://bcu1.lib.uni.lodz.pl/dlibra/publication/378>.
- “Na Szlązku. Listy korespondenta ‘Kraju’.” *Kraj* 35 (1899): 101–3.
- Opowiadania albo podróże jednego Kosmopolity. Podróż po Polsce i Śląsku w latach 1792 i 1793, w: Trzy relacje z podróży po Śląsku w XVIII i XIX wieku.* Edited by Jarosław Szymański. Translated by Lesław Szybkowski. Opole, Katowice: Dom Współpracy Polsko-Niemieckiej, Fundacja Zamek Chudów, 2006.
- Orłowicz, Mieczysław. *Przewodnik po ziemiach dawnej Polski, Litwy i Rusi*. Kraków: Eugeniusz Starczewski, 1914.

- Osika, Grażyna. "Czekając na osobliwość – o modelach interpretacji techniki." *Filosophia*, no. 39 (2017): 65–78.
- Piecha van Schagen, Barbara, "'Kopalnia to miłość'. Szkic o tożsamości." In *Narracje górnicze z terenu Zabrze. Kopalnia to je do mnie wszystko*. Edited by Bernard Linek. Zabrze: Muzeum Górnictwa Węglowego w Zabrzu, 2016.
- Piernikarczyk, Józef. *Historia górnictwa i hutnictwa na Górnym Śląsku*. Vol. 2. Katowice: Śląski Związek Akademicki, 1936.
- . *W 150 lat po uruchomieniu pierwszej maszyny parowej na Górnym Śląsku*, [http://sbc.org.pl/Content/158298/ii147714-1937\\_1939-12-0001.pdf](http://sbc.org.pl/Content/158298/ii147714-1937_1939-12-0001.pdf).
- Pieśni związane z pracą (zebrane przez Łukasza i Stanisława Wallisów)*. Manuscript. Biblioteka Śląska w Katowicach, R 2058 III.
- Platerowa, Katarzyna. *Moja podróż do Włoch. Dziennik z lat 1785-1786*. Edited by Małgorzata Ewa Kowalczyk. Translated by Anna Pikor-Półtorak. Łomianki, Dziekanów Leśny: Wydawnictwo LTW, 2013.
- Pol, Wincenty. "Z wycieczki." *Biblioteka Naukowego Zakładu im. Ossolińskich* 2, no. 5 (1847): 540–57.
- Reaktywacja górnictwa*. Zabytkowa kopalnia srebra i sztolnia Czarnego Pstrąga. <https://kopalniasrebra.pl/historia/>.
- Śnieżko, Daniel. "Antropologia w badaniach nad literaturą dawną." In *Jaka antropologia literatury jest dzisiaj możliwa?* Edited by Przemysław Czapliński, Anna Legeżyńska, and Marcin Telicki. Poznań: Poznańskie Studia Polonistyczne, 2010.
- Święto pary już po raz trzeci*. Tarnowskie Góry. <https://tarnowskiegory.pl/2023/10/swieto-pary-juz-po-raz-trzeci/>.
- Światała-Trybek, Dorota. *Tragedia w kopalni: Kulturowe konteksty katastrof i wypadków górniczych*. Opole: Wydawnictwo Uniwersytetu Opolskiego 2014.
- Tomasik, Wojciech. *Romantycy i technika*. Vols. 1-2. Warszawa: Instytut Badań Literackich PAN, 2017.
- Von der Dampfmaschine zur Eisenbahn. Bildquellen und Dokumenten zur Frühindustrialisierung Oberschlesiens 1780-1860 / Od maszyny parowej do kolei żelaznej. Źródła ikonograficzne i dokumenty do wczesnego uprzemysłowienia Górnego Śląska 1780-1860*. Edited by Peter Chmiel, et al. Translated by Peter Chmiel, et al. Ratingen-Zabrze: Oberschlesisches Landesmuseum-Muzeum Górnictwa Węglowego, 2004.
- Wójcik, Andrzej J. "Fossores ex Polonia. Rekrutacja polskich górników do Prus." *Hereditas Minariorum* no. 4 (2017), <https://www.dbc.wroc.pl/dlibra/publication/140293/edition/72368/content>.
- Wróblewski, Maciej. *Literatura i maszyna*. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika w Toruniu, 2015.
- Złota księga Tarnowskich Gór*. Edited by Marek Wroński. Translated from the German by Erwin Szyga, translated from the Latin by Zdzisław Jedynek, translated from the French by Anna Górka, translated from the Swiss by Beata Kokot, translated from the English by Nina Jarzyńska. Tarnowskie Góry: Instytut Tarnogórski, Muzeum Instytutu Tarnogórskiego, 2011.



## ABSTRAKT / ABSTRACT

Izabela Kaczmarzyk – Rzecz „godna widzenia” czy „apokaliptyczny potwór”? Konsekwencje maszynizacji górnośląskiej rzeczywistości kulturowej

DOI 10.12887/36-2023-4-144-11

Od końca osiemnastego wieku do przełomu wieku dziewiętnastego i dwudziestego na Górnym Śląsku następowały dynamiczne przemiany industrializacyjne i urbanizacyjne. Powszechna obecność maszyn miała silne oddziaływanie kulturotwórcze, dlatego wyraźnie zaznaczyła się w tym regionie zmiana wzorca kulturowego z tradycyjnego układu między człowiekiem a przyrodą na modernizacyjną relację człowieka i maszyny. Maszyny postrzegane były dwojako. Po pierwsze jako dowód ludzkiego geniuszu, dzięki któremu człowiek zyskiwał swoistą protezę powiększającą siłę jego mięśni i dzięki której mógł panować nad przyrodą (przykładem takiego rozumienia maszyny były opisy pierwszej maszyny parowej zainstalowanej w Friedrichsgrube w Tarnowskich Górach), a po drugie jako budzące niepokój, a czasem nawet trwogę mechaniczne bestiarium (takie ujęcie pojawia się częściej w tekstach późniejszych, z drugiej połowy dziewiętnastego i początku dwudziestego wieku). Ważną konsekwencją pracy w zmechanizowanym środowisku było też wykształcenie wśród mieszkańców uprzemysłowionej części Górnego Śląska umiejętności, które dopełniały siłę fizyczną, a polegały na sprawności wykonywania określonych czynności w określonym czasie, współdziałając z innymi w określony sposób, niekiedy z racjonalizatorską inwencją. Przemysł (i związana nim maszynizacja) stały się tym samym jednym z ważniejszych czynników górnośląskiego etosu. Jednak mimo zagrożeń, jakie wiązały się z obecnością maszyn i pracą przy maszynach, ten etap umaszynowania nie naruszał ludzkiej podmiotowości, ponieważ maszyna nie była jeszcze wówczas tak silnym konkurentem człowieka, jakim wydają się współczesne systemy technologiczne.

Słowa kluczowe: Górny Śląsk, maszyna, industrializacja

Kontakt: Instytut Kulturoznawstwa i Dziennikarstwa, Wydział Filozofii, Uniwersytet Ignatianum, ul. Mikołaja Kopernika 26, 31-501 Kraków

E-mail: [izabela.kaczmarzyk@ignatianum.edu.pl](mailto:izabela.kaczmarzyk@ignatianum.edu.pl)

ORCID: 0000-0001-7310-2209

Izabela Kaczmarzyk, A Thing “Worth Seeing” or an “Apocalyptic Monster”? Consequences of the Mechanisation of the Upper Silesian Cultural Environment

DOI 10.12887/36-2023-4-144-11

From the end of the eighteenth century to the turn of the nineteenth and twentieth centuries Upper Silesia saw dynamic industrialisation and urbanisation. The ubiquitous presence of machines had a strong culture-forming impact and therefore a change in the cultural pattern from the traditional system between



man and nature to the modernising relationship between man and machine was clearly marked in this region. Machines were perceived in two different ways. On the one hand, as the evidence of human genius owing to which human beings gained a sort of prosthesis increasing the strength of their muscles and enabling them to control nature (an example of such an understanding of the machine were the descriptions of the first steam engine installed in Friedrichsgrube in Tarnowskie Góry) and, on the other hand, as a mechanical bestiary, inducing anxiety, or sometimes even fright (such an approach appears more frequently in later texts, from the second half of the nineteenth and early twentieth centuries). An important consequence of working in a mechanised environment was the development of skills among the inhabitants of the industrialised part of Upper Silesia, which complemented their physical strength and consisted in the ability to perform certain activities in a certain time, cooperating with others in a certain way, sometimes with improvement-oriented inventiveness. Industry (and the related mechanisation) thus became one of the most important factors in the Upper Silesian ethos. Nevertheless, despite the perils associated with the presence of machines and work on machines, that stage of mechanisation did not violate human subjectivity since the machine was not then as strong a competitor as modern technological systems appear to be.

Keywords: Upper Silesia, machine, industrialisation

Contact: Contact: Institute of Cultural Studies, Faculty of Philosophy, Jesuit University Ignatianum, ul. Mikołaja Kopernika 26, 31-501 Cracow, Poland  
E-mail: [izabela.kaczmarzyk@ignatianum.edu.pl](mailto:izabela.kaczmarzyk@ignatianum.edu.pl)  
ORCID: 0000-0001-7310-2209

Zbigniew NAWRAT

## MEDYCYNĄ I ETHOS ROBOTÓW

*Wszystkie stosowane kliniczne systemy informatyczne powinny osiągać odpowiedni poziom integracji: niezawodność, użyteczność, interoperacyjność. Tylko wtedy będą mogły być skutecznie stosowane, aby odciążać lekarzy w trybie on-line, dostarczając wiarygodnych informacji i kontrolując działania w czasie wykonywanej procedury, wykorzystując protokoły, które mogą być wymieniane „ze zrozumieniem” przez ważne w tym procesie urządzenia. Pomimo postępów we wnioskowaniu przyczynowym AI nie „wyczuwa” jednak skutków podjętych decyzji dla pacjenta.*

O tym, że świat się składa z atomów, dowiedzieliśmy się od filozofów starożytnych, roboty stworzyli artyści, a pomysłodawcami Internetu byli twórcy powieści fantastycznych. Trudno urzeczywistnić to, co powstało w wyobraźni, ale właśnie dzięki innowacyjności, wyobraźni i pracy odnieśliśmy sukces jako gatunek ludzki. Postęp wiedzy związany z techniką prowadzi do zmian cywilizacyjnych. Sposób realizacji celów pracy zmienił się od czasu, kiedy udało nam się zdigitalizować treści zapisane wcześniej w sposób analogowy, a obliczenia sprowadzić do operacji zerojedynkowych, czyli skomputeryzować. Obecnie praca lekarza zmienia się, przechodzi od naturalnej ludzkiej inteligencji i pracy manualnej do wykorzystania sztucznej inteligencji i pracy robota. Robot medyczny staje się coraz częściej partnerem realizacji usług medycznych. To sprawia, że rodzi się potrzeba określenia etosu robota medycznego. To kolejny etap ewolucji człowieka i jego narzędzi.

Definiowanie pracy jako procesu złożonej aktywności fizyczno-umysłowej, której celem jest przekształcenie szeroko rozumianego środowiska w ten sposób, by zwiększyć szanse przeżycia gatunku ludzkiego<sup>1</sup>, jest dobrym jak opisem działalności lekarza, która dotyczy środowiska – człowieka (pacjenta).

Praca lekarza polega na udzielaniu świadczeń zdrowotnych, w szczególności: badaniu stanu zdrowia, rozpoznawaniu chorób i zapobieganiu im, leczeniu i rehabilitacji chorych, udzielaniu porad, a także wydawaniu opinii i orzeczeń lekarskich<sup>2</sup>.

<sup>1</sup> Zob. hasło „Praca”, w: *Encyklopedia Polskiego Wydawnictwa Naukowego*, <https://encyklopedia.pwn.pl/haslo/praca;3961600.html>.

<sup>2</sup> Zob. *[Wykonywanie zawodu lekarza] – Zawody lekarza i lekarza dentystry*, Dz.U.2023.1516, art. 2. <https://sip.lex.pl/akty-prawne/dzu-dziennik-ustaw/zawody-lekarza-i-lekarza-dentystry-16798282/art-2>.

Człowiek może wykonywać usługę medyczną w postaci pracy intelektualnej i działań mechanicznych za pomocą narzędzi oraz aparatury medycznej, z wykorzystaniem różnych zjawisk fizycznych. Jednak najczęściej praca lekarza to proces analizowania danych i podejmowania decyzji. Praca chirurga polega na inwazyjnym wtargnięciu w ludzki organizm, zadawaniu ran i leczeniu. To sztuka tworzenia chwilowych, kontrolowanych w danym obszarze organizmu sytuacji.

Roboty z powodzeniem są obecnie wykorzystywane jako: (1) roboty diagnostyczne, czyli roboty służące do nowoczesnej, cyfrowej, na przykład wykonanej w 3D diagnostyki obrazowej; (2) roboty chirurgiczne (narzędzia zwiększające jakość, precyzję interwencji chirurgicznej i często zmniejszające inwazyjność operacji); (3) roboty opiekuńcze, maszyny socjalne, które zwiększają jakość życia ludzi starszych, zniedołężniałych, z niewydolnymi narządami ruchu, ich wykorzystanie zwiększa samodzielność podopiecznych; (4) roboty rehabilitacyjne, czyli roboty służące do terapii, treningu, rehabilitacji przez kontrolowany ruch rehabilitowanych narządów ruchu; (5) roboty ratunkowe (roboty wykorzystywane zdalnie lub autonomicznie do akcji ratunkowych w różnym środowisku i różnych warunkach zagrożenia życia); (6) sztuczne narządy, zrobotyzowane elementy zastępcze uzupełniające działanie niektórych narządów organizmu człowieka; (7) bioroboty (roboty naukowe naśladujące ludzi lub zwierzęta, roboty wykorzystywane do celów poznawczych – neurofizjologii, patologii mózgu czy samoorganizacji społecznej); (8) edukacyjne roboty medyczne, roboty wykorzystywane do nauki zawodu lekarza, pielęgniarza czy ratownika, symulatory pacjenta.

Największe sukcesy robotyka medyczna odniosła do tej pory w zakresie robotów chirurgicznych i rehabilitacyjnych. W maju 2018 roku robot kardiologii inwazyjnej o nazwie Corindus Vascular Robotics otrzymał zgodę wydaną przez amerykańską Agencję ds. Żywności i Leków (U.S. Food and Drug Administration, FDA) na pierwszy zautomatyzowany ruch robotyczny zaprojektowany dla platformy CorPath GRX. Zastrzeżona funkcja oprogramowania, „Rotate on Retract” (RoR), jest pierwszym zautomatyzowanym mechanizmem robota, który umożliwia operatorowi szybką nawigację do docelowej zmiany anatomicznej przez automatyczne obracanie prowadnika podczas cofania joysticka<sup>3</sup>.

To dzięki wdrożeniu technologii produkcji seryjnej, automatyzacji i robotyzacji możliwy jest obecny wysoki poziom jakości ludzkiego życia. Medycyna, borykająca się z problemami kadrowymi, rosnącymi oczekiwaniami

<sup>3</sup> Por. Z. N a w r a t, *Roboty medyczne w systemach teleinformatycznych*, w: *Informatyka w medycynie*, red. M. Kurzyński i in., Akademicka Oficyna Wydawnicza Exit, Warszawa 2019, s. 727-760.

społecznymi i kosztami przekraczającymi możliwości ekonomiczne, potrzebuje robotyzacji. Ponad sto lat temu wprowadzone maszyny, automaty, roboty zmieniły manufaktury w zakłady przemysłowe, wpływając na politykę i relacje społeczne przełomu dziewiętnastego i dwudziestego wieku. Podobnie wprowadzane innowacje w medycynie poprawiają efektywność i zmieniają formy wykonywania usług medycznych oraz relacje pomiędzy wszystkimi uczestnikami procesu leczenia.

Przekonanie autora artykułu o znaczącej roli robotów dla dobra człowieka zostało ukształtowane przez lata pracy w zespole znakomitego, innowacyjnego kardiochirurga – prof. Zbigniewa Religi. Autor uczestniczył w opracowaniu i wdrożeniu systemu mechanicznego wspomaganie serca oraz prowadził prace nad robotem chirurgicznym Robin Heart.

Sztuczne narządy i AI są elementami robotyki medycznej. Obecnie stosowane urządzenia nie są samodzielne, ale prowadzone są prace, które mają na celu przekazanie części obowiązków decyzyjnych tym maszynom. To jest konieczne (ze względu na bezpieczeństwo pacjenta), jeśli chcemy rozwijać teleoperacje (zdalnie sterowane operacje za pomocą robotów na odległość zagrożone jest utratą komunikacji pomiędzy operatorem robotem) lub pozwolić pacjentom ze sztucznymi narządami przebywać w warunkach domowych. Tysiące pacjentów na świecie, żyjących dzięki sztuczным narządom, rehabilitowanych i poruszających się dzięki robotom, półtora miliona operacji rocznie wykonywanych z pomocą robotów – trudno o bardziej przekonujące dowody, że już dzisiaj roboty ratują życie, dbają o jego jakość i zmieniają standardy wykonywania pracy.

#### AI JAKO ALGORYTM

AI zwykle przetwarza informacje, stosując pewien zestaw reguł, instrukcji, czyli algorytmy. Na wejściu wprowadzana informacja „zamienia” się w wyniku procesu opartego na algorytmach na informację wyjściową – wynik. System może być nadzorowany przez programistę, można też stworzyć proces, zestaw algorytmów, w którym nie będzie potrzebna interwencja programisty do aktualizacji oprogramowania. Taki system określa umiejętność rozwiązywania problemów przez „uczenie się” na podstawie poprzednich wyników. Działając na dostarczonych danych i oceniając efekty (czy są one zbliżone do spodziewanej odpowiedzi), program może sam poszukiwać wzorców i doskonalić się. Tym właśnie zajmuje się uczenie maszynowe (ang. machine learning, ML). Algorytm to aspekt techniki matematycznej, a stosując algorytm uczenia maszynowego pobierane są dane wejściowe i wyjściowe, aby w trakcie działania (proces treningowy) przygotować się lepiej do pracy nad nowymi

danymi wejściowymi. Podobnie fizyka stosuje matematykę jako narzędzie opisu rzeczywistości. Nie każde rozwiązanie matematyczne ma jednak sens fizyczny. Ucząca się AI próbuje naśladować za pomocą narzędzi matematycznych działanie neuronów naszego mózgu, wytwarza rozwiązania. W procesie treningowym, na każdym poziomie dokonywania wyborów, w trakcie analizowania danych wejściowych nadawane są wagi, wyróżniające dobre rozwiązania, czyli takie, które zmierzają do wyjściowej informacji, zgodnej z obserwacją rzeczywistością.

Systemy AI są sprawne w szukaniu korelacji, powiązań faktów czy powtarzających się wzorców (pikseli na obrazach), znajdują ich więcej i szybciej niż ludzie, ale większość z nich nie jest zrozumiała dla człowieka. Działa to w pewnym stopniu jak wprowadzanie uważnego hobbysty do zespołu profesjonalistów, który przez naśladowanie i praktykowanie może osiągnąć wysoki poziom skuteczności działania. Inteligencja jest jak „spryt” – powoduje, że we właściwy (bezpieczny, określony przez kod kulturowy sposób) reagujemy na bodźce zmysłowe. Rozumność to umiejętność wykorzystania abstrakcyjnych modeli do rozwiązywania rzeczywistych zadań. Wystarczy inteligencja, by widząc fotografię sporządzoną przez mikroskop sił atomowych, uzyskać przekonanie o tym, że świat się składa z atomów. Dwa i pół tysiąca lat wcześniej rozumowanie doprowadziło też do takiego przekonania Demokryta z Abdery.

Można podać wiele przykładów, pokazujących, jak sztuczna inteligencja „sprytnie” omija „rozwiązywanie” problemu, optymalizując ekonomicznie obciążenie i czas pracy komputera. Na przykład poszukując przypadków gruźlicy na podstawie analizowanych zdjęć RTG, szybko odnajdzie na wybranych zdjęciach notatkę, że pochodzi ono ze szpitala gruźliczego, więc nada mu rangę zwiększonego ryzyka, skracając czas, omijając „faktyczne” analizowanie zdjęcia pod kątem wystąpienia zmian właściwych gruźlicy. Podobnie jak u ludzi, tak i tu wszystko zależy od metod kształcenia, szkolenia. Obecnie prowadzone są intensywne badania umożliwiające kontrolowanie decyzji na każdym etapie „pracy AI”. Potrzebujemy nowego typu uczenia. AI uczy się tworząc wzorce przez analizowanie przypadków, podobnie jak znachor, który zwykle nie ma pogłębionej wiedzy teoretycznej. Wybiera podobne przypadki, opisane na podstawie określonych badań, które powodują podobne skutki chorobowe. Stąd wnosi o podobnych przyczynach, wskazuje na podobne przewidywania i sugeruje podobny sposób leczenia. Przez analogię, taki system uczenia sprawdza w prawie amerykańskim opartym na *casusach* lub w projektowaniu akcji reklamowych opartych na wzorcach socjologicznych. Ma to jednak mniejsze zastosowanie w medycynie. Kilka tysięcy lat zajęło nam doskonalenie systemu edukacji, przez szkoły, uniwersytety, podręczniki i wynalezienie metod naukowych, abyśmy mogli w pewnym przybliżeniu poznać proces podejmowania decyzji przez innych ludzi. Metody AI rozwijają się od

prostego Perceptronu, przez skomplikowane sieci neuronowe i kolejne inspiracje biologiczne (działanie mózgu). Szybki rozwój możliwości obliczeniowych procesorów i komputerów spowodował, że możemy sobie pozwolić na wielokrotne powtarzanie procesu uczenia, opierając się na dużych zbiorach danych, by w ten sposób uzyskać, pomimo prostoty algorytmu, oczekiwane rezultaty. Jeśli chcemy wykorzystywać sztuczną inteligencję (AI) w różnych obszarach medycyny, które wiążą się z wysokim ryzykiem, konieczne jest opracowanie specjalnych technik uczenia sieci neuronowych. Inspiracją w tym zakresie może być sukces, jaki odniesiono poprzez zwiększenie bezpieczeństwa i efektywności świadczenia usług medycznych, dzięki standaryzacji procedur opartych na konkretnych wytycznych. Wytyczne te, stosowane w określonych dziedzinach i procedurach medycznych, są aktualizowane w oparciu o wnioski płynące z publikowanych badań naukowych. Transparentność metodologii ich tworzenia pozwala na identyfikację ograniczeń w ich stosowaniu oraz stwarza możliwość udoskonalania leczenia poprzez wdrożenie procesu ciągłego doskonalenia opartego na faktach.

### ŚWIAT WARTOŚCI W PRACY AI

Wszystkie stosowane kliniczne systemy informatyczne, szczególnie w infrastrukturze szpitala wyposażonego w roboty i urządzenia automatyczne, powinny osiągać odpowiedni poziom integracji: niezawodność, użyteczność, interoperacyjność (ang. reliability, usability, interoperability). Tylko wtedy będą mogły być skutecznie stosowane, aby odciążyć lekarzy w trybie on-line, dostarczając wiarygodne informacje i kontrolując działania w czasie wykonywanej procedury, wykorzystując protokoły, które mogą być wymieniane „ze zrozumieniem” przez ważne w tym procesie urządzenia. Pomimo postępów we wnioskowaniu przyczynowym AI nie „wyczuwa” skutków decyzji podjętych dla pacjenta. Wielopłaszczyznowa, wieloparametrowa analiza dokonana przez człowieka zawsze będzie konieczna do adekwatnej oceny przypadku chorobowego. Dla pacjenta czy lekarza decyzja robota może mieć wymiar nie tylko prawdy czy fałszu, ale też wyroku bądź wygranej w walce z chorobą.

Wykorzystanie technik ML wiąże się też z wątpliwościami. Ludzie pragną zrozumieć sposób podejmowania decyzji i udoskonalać procesy. To jest zgodne z zasadą bezpieczeństwa – nasze osiągnięcia w zakresie bezpieczeństwa są w znacznej mierze skutkiem jawności i wprowadzanych standardów – algorytmów sukcesu. Wszystkie systemy jakościowe ISO działają na zasadzie otwartości, dokumentowania, przejrzystości. Postęp nie byłby możliwy, gdyby nie możliwość nadzorowania (na każdym etapie) realizacji celów kolejnych zadań. Intuicyjnie czujemy potrzebę uwidocznienia każdego elementu łańcucha, aby uwierzyć, że spełni nasze oczekiwania.



Potrzeba zwrócenia uwagi na problemy etyczne, regulacji i standaryzacji wynika z oceny skutków, czyli ryzyka wprowadzania AI i robotów w medycynie. W tym obszarze organizacje pozarządowe czy redakcje czasopism specjalistycznych zwykle wyprzedzają prawo. Autor artykułu jest zaangażowany w te działania w ramach European Society for Artificial Organs, Międzynarodowego Stowarzyszenia na rzecz Robotyki Medycznej, Polskiego Stowarzyszenia „Sztuczna inteligencja w Medycynie”, Fundacji AI LAW TECH, Digital Innovation Hub Healthcare Robotics HERO i czasopism „AI Surgery” oraz „Medical Robotics Reports”. Na potrzeby tego opracowania warto zacytować rekomendacje opublikowane 22 maja 2019 roku przez Organizację Współpracy Gospodarczej i Rozwoju, (Organisation for Economic Cooperation and Development, OECD): *Recommendation of the Council on Artificial Intelligence*<sup>4</sup>. Celem ogłaszania rekomendacji jest promowanie odpowiedzialnego zarządzania oraz poszanowanie praw człowieka i wartości demokratycznych, wspieranie innowacji i budowanie zaufania do sztucznej inteligencji, czyli dbałość o jej twórców i użytkowników. Szczegółowe zalecenia obejmują:

(1) **R o z w ó j A I d l a d o b r o b y t u i z r ó w n o w a ż o n e g o r o z w o j u**: sztuczna inteligencja powinna zapewniać dobrobyt ludzkości i przynosić korzyści planecie; zwiększać ludzkie możliwości, rozwijać kreatywność, umożliwiać integrację grup wykluczonych, minimalizować nierówności gospodarcze, społeczne, płciowe i inne.

(2) **K o n c e n t r a c j a n a w a r t o ś c i a c h i c z ł o w i e k u**: oznacza respektowanie praworządności, praw człowieka i wartości demokratycznych (poszanowanie wolności, prywatności, godności człowieka, z uwzględnieniem ochrony danych, sprawiedliwości i równości społecznej).

(3) **T r a n s p a r e n t n o ś ć**: umożliwia uzyskanie przez użytkownika informacji o sposobach działania systemu, logice podejmowania decyzji przez algorytm i czynnikach wpływających na tę decyzję.

(4) **B e z p i e c z e ń s t w o**: AI nie tylko ma zapewnić bezpieczeństwo użytkownikowi, ale także przeciwdziałać wykorzystywaniu jej do nieodpowiednich celów.

(5) **O d p o w i e d z i a l n o ś ć**: twórcy sztucznej inteligencji są odpowiedzialni za prawidłowe funkcjonowanie systemów AI i przestrzeganie zasad etyki<sup>5</sup>.

OECD podkreśla rolę AI dla rozwoju cywilizacji, ale też i ograniczenia wynikające z potrzeby bezpieczeństwa wszystkich uczestników procesu,

<sup>4</sup> Zob. Organisation for Economic Cooperation and Development, *Recommendation of the Council on Artificial Intelligence*, OECD Legal Instruments, <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449>.

<sup>5</sup> Zob. tamże.



w których jest ona stosowana. Eksperti zalecają, aby systemy miały zabezpieczenia umożliwiające ludziom przejęcie kontroli nad maszyną, gdy zajdzie taka potrzeba. Ważne, aby użytkownikowi przysługiwało prawo wglądu w historię podejmowanych przez algorytmy decyzji oraz zakwestionowania decyzji podjętej przez maszynę.

AI pozwala na poszerzenie wiedzy pacjentów na temat stanu ich zdrowia. Demokratyzacja ochrony zdrowia dzięki technologii jest możliwa, ale związana jest ze zmianami systemowymi i utratą monopolu własności danych przez lekarzy. Rozwój AI wymaga dostępu do wyczerpujących danych, które dzisiaj są rozproszone. Scentralizowany polski system medyczny posiada duży potencjał w tym zakresie. Paradoksalnie jednak zwiększona liczba danych z urzędów domowych kontrolujących pacjentów może również zwiększać potrzeby realnych konsultacji (o tym informowali lekarze z MayoClinic, oceniając wzrost „fałszywie?” pozytywnych wykryć migotania przedsionków przez zegarki Apple).

Urządzenia wyposażone w AI są po to, by wspierać personel medyczny. Rosnąca różnica między podażą usług zdrowia a potrzebami tworzy permanentny stan kryzysowy, który objawia się zarówno zmniejszeniem satysfakcji z wykonywanego zawodu przez lekarzy, jak i niepokojącą liczbą błędów medycznych. Ociążenie ich przez outsourcing zadań administracyjnych i analizę danych przez maszyny daje lekarzom szansę na przywrócenie medycynie empatii. Relacje między wykonywaniem usług przez człowieka a rozwojem metod sztucznej inteligencji ujął Eric Topol, wprowadzając pojęcia głębokiej i płytkiej medycyny. „Płytką medycynę” charakteryzuje pośpiech i błędy spowodowane powierzchowną analizą informacji. Właśnie w zakresie analizy informacji AI może znacząco pomóc lekarzom. „Głęboka medycyna”, według Topola, to medycyna skupiona na pełnym, dogłębnym zrozumieniu każdego pacjenta, które trudno osiągnąć bez wsparcia AI<sup>6</sup>.

W modelu medycyny pogłębionej najważniejsza jest empatia, polegająca na utrzymaniu więzi między pacjentem i lekarzami. Wbrew powszechnym opiniom AI może przywrócić czynnik humanistyczny w praktyce medycznej. Właśnie to zrozumienie, wynikające ze wsparcia AI i lekarskiego wywiadu, możliwość nawiązania odpowiednio bliskich i rozłożonych w czasie relacji między człowiekiem-specjalistą a człowiekiem-pacjentem, może zwiększyć empatię lekarza. Choć jednak swoim studentom autor, dla ich bezpieczeństwa i zdrowia zaleca postawę współczucia (zamiast empatii) w relacjach z pacjentem. Zrozumienie drugiego człowieka to coś więcej niż wynik analizy danych pomiarowych.

---

<sup>6</sup> Zob. E. Topol, *Medycyna głęboka. Jak sztuczna inteligencja może ponownie uczynić opiekę zdrowotną ludzką*, red. A. Horzyk, T. Stompór, tłum. A. Boniszewska, ITEM Publishing, Warszawa 2020.

„Eric Topol przypomina, że gdy tworzone przez nas maszyny stają się coraz mądrzejsze i niebawem przejmą od nas większą liczbę zadań, my musimy jeszcze bardziej rozwijać w sobie człowieczeństwo i humanitarność, aby to zrównoważyć. Najgenialniejsze narzędzia AI pomogą nam dowiedzieć się o sobie – o naszych ciałach i umysłach – więcej, niż nam się kiedykolwiek wydawało, ale nie będą empatyczne w stosunku do pacjentów”<sup>7</sup> – podsumował Garry Kasparov (autor książki *Deep Thinking*, mistrz szachowy, który grał z komputerem Deep Blue), komentując książkę Topola.

### AI JAKO INFORMACJA

Żyjemy w przestrzeni informacji, czyli określonym punktem w czasie i przestrzeni możemy przypisać pewne dane, historię. Powracając do inteligencji – jest to nasze osobiste urządzenie do optymalizacji podejmowanych decyzji w wielowymiarowej przestrzeni informacji. Człowieka można również ująć jako zbiór informacji (około 1 terabajt danych w ciągu roku). Jeśli ktoś choruje na cukrzycę czy serce, informacji jest więcej. Analiza danych (ang. big data) odnosi spektakularne sukcesy (i porażki), analizując wpływ leków, diety czy zachowań na skuteczność procesu leczenia. Tam, gdzie kończy się nasza możliwość analizy przyczynowo-skutkowej, stosowania i weryfikowania teorii, ścisłych algorytmów, otwiera się pole dla zastosowania wielowarstwowych, komputerowych sieci neuronowych, samouczenia, czyli sztucznej inteligencji. Jeśli pacjent jest scharakteryzowany przez wyniki badań diagnostycznych oraz opis przebiegu i sposobu leczenia w sposób cyfrowy, to można wykorzystać metody sztucznej inteligencji dla optymalizacji procesu leczenia. Wprowadza się wiele programów doradczych analizujących dane i wspomagających podejmowanie decyzji przez lekarzy. Dzisiaj standardem jest takie wyposażenie programowe do każdej aparatury diagnostycznej. To już stanowi sukces. Jednak nawet superkomputer Watson firmy IBM, który z powodzeniem wygrał kilka popularnych gier i teleturniejów (czyli jest inteligentny), nie wykazał, w sposób jednoznaczny, swojej niezwyklej użyteczności jako diagnostyczne narzędzie doradcze dla celów medycznych. Oznacza to, że ciągle mamy do czynienia ze wspomaganiami przez AI procesu decyzyjnego, a nie zastąpieniem diagnostyki przez oprogramowanie.

W przypadku wątpliwości polegamy nadal na wiedzy i doświadczeniu lekarzy diagnostów. Sztuczna inteligencja powstała po to, aby komunikować maszyny, komputery z ludzką inteligencją. Skoro do tej pory nie mamy

<sup>7</sup> Wypowiedź Garry’ego Kasparova, Księgarnia Internetowa PWN, <https://ksiegarnia.pwn.pl/Medycyna-gleboka,871732290,p.html>.

możliwości użycia sztucznej inteligencji zamiast ludzkiej, to nie możemy usamodzielniać robotów i są one stosowane przez chirurga jako narzędzia chirurgiczne, czyli telemanipulatory. AI „poszerza” jego zmysły, co pozwala mu podejmować trafniejsze decyzje. Chirurg, lekarz, operator jest jednak niezbędnym elementem pętli procesu decyzyjnego.

Sztuczna inteligencja jest testowana w warunkach komunikacji i rywalizacji z człowiekiem. Gra według określonych reguł i kreatywność, związana ze stosowaniem reguł, stanowią idealne środowisko badania stopnia zaawansowania AI. Ada Lovelace, którą uznaje się za pierwszą programistkę i wizjonerkę AI, zwróciła uwagę, że „maszyna analityczna nie ma pretensji do dawania początku czemukolwiek, bo wykonuje wszystko, co jej polecimy”, a „kreatywność to dążenie do czegoś nowego, zaskakującego i wartościowego”, i „łatwo jest stworzyć coś nowego [...] trudniej stworzyć zaskoczenie i wartość”<sup>8</sup>. Lovelace znana jest przede wszystkim z publikacji na temat mechanicznego komputera Charlesa Babbage’a. Pierwsze programy matematyczne służyły do programowania maszyn włókienniczych. Anglia, uznawana za kolebkę przemysłu, była miejscem powstania pierwszych programów kontrolujących maszyny.

Kreatywność wymaga ryzyka i poniesienia skutków błędów. Dlatego gra między maszyną a człowiekiem wydaje się obiektywnym ringiem, miejscem starcia sztucznej i naturalnej inteligencji. W 1997 roku program Deep Blue firmy IBM pokonał mistrza szachowego. W 2016 roku AlphaGo firmy DeepMind pokonał mistrza gry w go Lee Sedola. Nowy program AlphaZero to wszechstronny algorytm uczący się od zera. Z kolei zespół Google’a (firma zakupiła w 2014 roku DeepMind) skupił się na projekcie MIZAR, którego celem jest zbudowanie biblioteki dowodów zapisanych w formalnym języku, który komputer mógłby zrozumieć i sprawdzić. Pomysłodawcą tego projektu jest polski matematyk Andrzej Trybulec<sup>9</sup>.

## AI I BADANIA MÓZGU

Komputery lepiej radzą sobie z dowodzeniem twierdzeń matematycznych niż rozpoznawaniem obrazu; z logiką niż kojarzeniem obrazów, logika bowiem oparta jest na prostym zbiorze reguł. Człowiek znajduje wyjście z zaskakującego problemu na podstawie generalizacji wcześniejszych doświadczeń. Pierwsze programy ekspertowe (na przykład MYCIN rozpoznający bakterie odpowiedzialne za choroby zakaźne czy CYC, którego celem było skodyfi-

<sup>8</sup> Cyt. za: M. du Sautoy, *Kod kreatywności. Sztuka i innowacje w epoce sztucznej inteligencji*, tłum. T. Chawziuk, Copernicus Center Press Kraków 2020, s. 8, 13.

<sup>9</sup> Por. tamże, s. 46, 258-261, 276.

kowanie ludzkiego rozsądku, opierały się na zbiorze reguł. Ekspert musieli dostarczyć informacji na temat faktów i reguł. System musiał być uaktualniany na podstawie nowych faktów, był nieskalowalny. W latach pięćdziesiątych dwudziestego wieku rozpoczęto badania mózgu. Pionierskie prace Alana Hodgkina, Andrew Huxleya (przepływ impulsów „albo nic”) oraz Bernarda Katza (działanie synaps w komunikacji neuronów oparte na przetwarzaniu sygnałów elektrycznych na sygnały chemiczne) tworzyły bazę wiedzy o mózgu. W latach osiemdziesiątych dwudziestego wieku obok grupy zajmującej się sztuczną inteligencją opartą na logice (funkcjonalizm) pojawiła się grupa matematyków i inżynierów inspirowana biologiczną budową mózgu oraz koncepcją „sieci neuronowych”, równoległego przetwarzania rozproszonego, logiki rozmytej. Do tej ostatniej należał Terrence Sejnowski<sup>10</sup>.

Sejnowski jest w gronie fizyków usiłujących wykorzystać prawa teorii kwantowej do badania tajemnic mózgu. Nasza inteligencja tworzyła się w procesie ewolucji. Każda istotna reakcja w komórce wykształciła enzym, którego celem jest katalizowanie tej reakcji. W naturze występuje pewien zakres inteligentnych zachowań, na których przykładzie mogą uczyć się sztuczne systemy – tym zajmuje się biologia algorytmiczna. Głębokie uczenie polega na optymalizowaniu funkcji strat. Odwrotność funkcji strat w ewolucji nazywa się dostosowaniem. W mózgu to potrzeba pożywienia, ciepła, bezpieczeństwa, tlenu czy rozmnażania pełni właśnie taką rolę przyrodzonej funkcji strat. W uczeniu ze wzmocnieniem podejmuje się działania w celu zoptymalizowania przyszłych nagród. Dopiero wprowadzenie tych rozwiązań przyczyniło się do rozwoju rozpoznawania obrazów (wzorców) przez automaty<sup>11</sup>.

## AI I MODELOWANIE

Kosmos i medycyna są obszarami badań, w których modelowanie, symulacje odgrywają ważną rolę, a często są jedyną możliwością. Próbuje się odtworzyć najistotniejsze cechy układu rzeczywistego w postaci modelu matematycznego lub fizycznego (modelowanie). Na tej podstawie możemy też tworzyć symulatory służące do badań. Obecnie najpopularniejsze są symulacje komputerowe (symulatory w postaci programów komputerowych). Z drugiej strony ze względu na oczywistą trudność w eksperymentowaniu na żywym organizmie opieramy dzisiaj standardy i wytyczne medyczne w dużej mierze na badaniach statystycznych.

<sup>10</sup> Por. T.J. Sejnowski, *Deep Learning. Głęboka rewolucja*, tłum. P. Cypryański, Wydawnictwo Poltext, Warszawa 2019, s. 47-49.

<sup>11</sup> Por. tamże, s. 43-102.

Metody sztucznej inteligencji przetwarzające informacje, sygnały, obrazy są w jakimś sensie modelem funkcjonalnym mózgu, w większym lub mniejszym uproszczeniu. Znane są ograniczenia stosowania modeli, również sieci neuronowe i AI mają swoje ograniczenia. Idea modelowania zyskała określenie „testu na bliźniaku pacjenta”, ponieważ dostępne dane diagnostyczne pozwalają na budowanie, w znacznej mierze przybliżonych do oryginału, modeli spersonalizowanych, konkretnego pacjenta. „Działając na nim”, możemy otrzymać nie ogólne, ale szczegółowe rozwiązanie, dopasowane do danych pozyskanych od konkretnej osoby.

W przyszłości jako potencjalni pacjenci będziemy mieli swój model, symulator, czyli swojego „bliźniaka” obejmującego genotyp, fenotyp, historię i skutki chorób czy urazów. Lekarze będą mogli testować na tym modelu – bliźniaku – potencjalne korzyści i niebezpieczeństwa wynikające ze stosowania różnych metod leczenia, stosowania leków, a przeprowadzenie operacji na bliźniaku ułatwi odpowiednie zaplanowanie operacji właściwej. Będzie można dokładnie rozplanować wybór materiałów i urządzeń, a nawet zespołu medycznego oraz przećwiczyć konkretne ruchy i czynności chirurgiczne podczas zabiegu.

Modelowanie fizyczne i symulacje komputerowe były podstawowymi narzędziami weryfikacji poprawności realizacji założeń dotyczących tworzenia sztucznego serca i robota chirurgicznego w ramach prowadzonych projektów przez autora. Projektowana konsola sterowania telemanipulatorem chirurgicznym zawierała monitor z funkcją bierną i aktywną doradztwa. Bierna rada to oczywiście dostęp do danych diagnostycznych, odpowiednio mapowanych na obraz pola operacyjnego, a aktywna – to próba oceny skutku, prognozy, odpowiedniego ruchu oraz czynności narzędzia chirurgicznego. Wirtualna sala operacyjna służy do planowania operacji. Chirurg przyszłości będzie mógł pracować, obserwując nie naturalny widok pola operacji, ale przetworzony, w którym każdy punkt jest aktualizowanym zbiorem informacji diagnostycznych, fizycznych i geometrycznych. Chirurg robotyczny wyposażony w AI będzie pracował jak pilot, który nie potrzebuje bezpośredniego dostępu do naturalnego widoku lotniska, żeby bezpiecznie wylądować.

Sztuczna inteligencja jest elementem robotyki. Pojęcie robota wywodzi się z pragnienia upodobnienia maszyny do człowieka. Warto tu wspomnieć o nowej dziedzinie o nazwie robonomika (połączenie dwóch wyrazów: robotyka i ekonomika), która zajmuje się robotami rozumianymi jako oprogramowanie (ang. software). Chodzi o automatyzację określonych czynności wykonywanych przez mózg człowieka, niejako na wzór aktywności mechanicznej (na przykład zdolności związanych ze sterowaniem pojazdem przez człowieka).

Człowiek nie tylko wykorzystuje technologię, ale jest też uwikłany w technologiczne związki. Telemedycyna, czyli technika przesyłania informacji na

odległość o treści medycznej staje się niezbędnym elementem świadczenia usług ochrony zdrowia. Nie dyskutując z wagą informacji we właściwym miejscu i czasie, trzeba jednak dodać, że informacja nie leczy. Jednak zła informacja, błędna, jest niebezpieczna i w konsekwencji przyczynia się do szkody pacjenta. Terapia wymaga zastosowania leków lub czynności, zabiegów medycznych, aby można było wykonywać zabiegi medyczne, potrzebne są narzędzia i maszyny.

Jednym z ciekawych aspektów w zakresie relacji człowiek–maszyna jest robot chirurgiczny. To nowe narzędzie w ręku chirurga. Zanim zacznie być samodzielny, autonomiczny, pozostaje telemanipulatorem w ludzkich rękach. Ingerując w ciało pacjenta chirurg próbuje usunąć skutki choroby albo urazu, naprawić lub usprawnić działanie narządów pacjenta. Chirurg posiada i stosuje narzędzia powodujące zwykle skutki nieodwracalne. Tego rodzaju działanie to uboczna konsekwencja działalności lekarza. Jednak w przypadku porażki chirurga, na etapie decyzyjnym lub na etapie wykonania, lekarz wpływa negatywnie na los pacjenta. Każda ingerencja w standardy chirurgii, w tym wprowadzanie autonomiczności narzędzi decyzyjnych (AI), a tym bardziej wykonawczych (MR AI), jest działaniem o najwyższym stopniu ryzyka. Im mniejsza inwazyjność operacji, tym częściej decydujemy się na operowanie z mniejszym dostępem do informacji i mniejszym zakresem możliwości bezpośredniego oddziaływania na całe pole operacji. Dlatego w chirurgii mini-inwazyjnej rola stosowanych narzędzi jest kluczowa, zarówno narzędzi do operowania, jak i narzędzi umożliwiających decydowanie o podjęciu lub zaniechaniu określonych działań chirurgicznych.

Człowiek, operator robota chirurgicznego, może być traktowany jako element systemu sterowania powiązany przez system informatyczny i działanie układu elektromechanicznego robota z narzędziem wykonawczym. Bierzymy tu pod uwagę zarówno przetwarzanie informacji w mózgu, jak i zależną od niego koordynację ruchową precyzyjnego sterowania poprzez interfejsy robota<sup>12</sup>.

Według Richarda Satavy robot chirurgiczny to dziś bardziej narzędzie informatyczne, aniżeli mechaniczne. Robot nie jest maszyną, lecz urządzeniem informatycznym i stwarza duże możliwości integracji całego systemu diagnostycznego z wykonawcą<sup>13</sup>. „Przyszłość technologii i medycyny nie leży w krwi i jelitach, ale w bitach i bajtach”<sup>14</sup>. W ten sposób Richard Satava, chirurg

<sup>12</sup> Por. Nawrat, dz. cyt., s. 727-760.

<sup>13</sup> Por. R.M. Satava, *Future Directions in Robotic Surgery*, w: *Surgical Robotics: Systems Applications and Visions*, red. J. Rosen, B. Hannaford, R.M. Satava, Springer, New York – Dordrecht – Heidelberg – London 2011, s. 3-11.

<sup>14</sup> R.T. Sant'Anna i in., *Robotic Systems in Cardiovascular Surgery*, „Revista Brasileira de Cirurgia Cardiovascular” 19(2004) nr 2, s. 171 (<https://www.scielo.br/j/rbcecv/a/GDNJJkcTGxDZVc-fj8cxhxgR/?lang=en>). Tłum. fragm. – Z.N.



z Uniwersytetu Waszyngtońskiego, który kierował pierwszym projektem robota chirurgicznego finansowanym przez Agencję Zaawansowanych Projektów Badawczych Obrony USA (The Defense Advanced Research Projects Agency, DARPA), podsumował swoje wieloletnie doświadczenie i wizje przyszłości chirurgii.

Kevin Warwick, nazywany przez dziennikarzy „pierwszym cyborgiem”, jako pierwszy człowiek wszczepił sobie pod skórę mikronadajnik, prowadzi również badania nad bezpośrednim połączeniem człowieka z maszyną, czyli robotami z biologicznymi mózgzami. Ludzkie ciało może przy wsparciu takiej technologii pokonać kolejne ograniczenia i zrealizować bardziej zaawansowane cele. Jednak Warwick zwraca uwagę, że zwykli ludzie „zostaną w tyle” w konkurencji z cyborgami<sup>15</sup>.

AI jest elementem lub zarządcą systemu sterowania, oprogramowania kontrolującego robota. Przy tworzeniu oprogramowania powinny być zastosowane zasady ergonomii, których celem jest zapewnienie skuteczności, wydajności i zadowolenia użytkownika.

Omawiając relacje człowiek–robot, należy wspomnieć o robotach ubieralnych, egzoszkieletach, stosowanych w rehabilitacji. Kolejną grupą robotów budowanych z myślą o uczestniczeniu w powszednim dniu człowieka są roboty socjalne, opiekuńcze. Dobry robot „rozumie” intencje użytkownika, uczy się, wykorzystując różne drogi komunikacji (sensory jak zmysły). Sztuczna inteligencja pełni niezbędną rolę protezy cybernetycznej. W tym ujęciu cybernetyka to nauka o sterowaniu, pozyskiwaniu i przekształcaniu informacji w człowieku i w urządzeniu. Tym bardziej istotna, gdy człowiek zespolony jest fizycznie z urządzeniem.

## ETYKA ROBOTÓW

Sto lat temu czeski pisarz Karel Čapek w sztuce *R.U.R. Rossum's Universal Robots* po raz pierwszy użył słowa „robot”<sup>16</sup>. Roboty pojawiły się na literackiej scenie, aby prowokować pytania dotyczące przyszłości świata i odpowiedzialności człowieka (za ten świat oraz jego mieszkańców). Sformułowane siedemdziesiąt lat temu przez Isaaca Asimova prawa robotyki wskazywały jako naczelną zasadę nadrzędności człowieka nad robotem, a w przypadku konfliktu nadrzędność wykonania zadania przez robota mniej „inteligentnego”

<sup>15</sup> Por. K. Warwick, *The Diminishing Human-Machine Interface*, „Medical Robotics Reports” 2(2013), s. 4-11.

<sup>16</sup> Zob. K. Čapek, *R.U.R. Rossum's Universal Robots*, tłum. M.M. Lemańczyk, Fundacja Nowoczesna Polska, Warszawa 2021.



w stosunku do robota wyposażonego w pełniejszy system oceny środowiska pracy. Moglibyśmy przyjąć za regułę główną zasadę nadrzędności człowieka nad AI. Niepokoi jednak – znana w relacjach międzyludzkich (nawet politycznych) – optymalizacja ryzyka, kwestia zachowania bezpieczeństwa, „obrona pokoju” przez danie pierwszeństwa mniej inteligentnym systemom<sup>17</sup>.

Etyka milcząco zakłada bezpośredni kontakt człowieka z człowiekiem. Nasze sumienie, empatia inaczej w sytuacji braku bezpośredniego kontaktu z bliźnim. Posiadane środki techniczne będą pozwalały na coraz większy wpływ na otoczenie, mnożąc możliwości, ale również ryzyko. Dzięki robotyzacji procesu leczenia możliwe jest leczenie na odległość. Na otrzymany efekt ma wpływ cały wielu ludzi (na przykład inżynierów, administratorów, ekonomistów) i zdarzeń losowych. Jakie będą skutki powiązania długości i jakości ludzkiego życia ze środkami technicznymi? Niegdyś oceniano jakość życia miarą dostępności do dobrej wody i pożywienia, dzisiaj dodajemy do tego dostęp do energii elektrycznej i informacji. Wiemy, że jednostka jest silnie powiązana z otoczeniem społecznym (globalnym) i technicznym. Obiektywne, proste modele symulacyjne podpowiadają, jak istotne w skali istnienia cywilizacji są obecnie podejmowane decyzje na poziomie domu, firmy, państwa. Poszukując mistrzów i prawd ułatwiających orientację w tej sytuacji, warto wrócić do Immanuela Kanta. Zaproponowany przez niego imperatyw kategoryczny nie stracił na aktualności: „Postępuj wedle takiej tylko maksymy, dzięki której możesz zarazem chcieć, żeby stała się powszechnym prawem”<sup>18</sup>. Podobnie jak praktyczne rozwinięcie imperatywu: „Postępuj tak, byś człowieczeństwa tak w twej osobie, jako też w osobie każdego innego używał zawsze zarazem jako celu, nigdy tylko jako środka”<sup>19</sup>. Poszanowanie godności człowieka, człowieczeństwo powinny być uwzględniane zarówno w rachunku ekonomicznym, jak i w praktyce społecznej, nawet wtedy lub szczególnie wtedy, gdy mamy do czynienia z informatyzacją i robotyzacją praktycznej działalności ludzkiej.

Dzięki harmonijnemu rozwojowi nauk technicznych i biologicznych oraz ich rozsądnemu wdrażaniu możemy wpływać na ewolucję gatunków i jakość życia na Ziemi. Sztuczna inteligencja jest częścią tej ewolucji. Skoro jest to element zarządzania na poziomie organów, organizmu, procedur medycznych na ciele, zarządzania szpitalem czy państwem – należy wprowadzić kryteria etyczne uzgodnione ze wszystkimi ludźmi<sup>20</sup>. Roboty muszą szanować prawa,

<sup>17</sup> Z. N a w r a t, *Etyka ery sztucznych narządów*, w: *Etyka wobec współczesnych wyzwań*, red. A. Bobko, Uniwersytet Rzeszowski, Rzeszów 2013, s. 39-50.

<sup>18</sup> I. K a n t, *Uzasadnienie metafizyki moralności*, tłum. M. Wartenberg, PWN, Warszawa 1981, s. 50.

<sup>19</sup> Tamże, s. 62.

<sup>20</sup> Por. Z. N a w r a t, *The Ethics of Artificial Organs*, w: *Implant Expert*, red. Z. Nawrat, M-Studio, Zabrze 2011, s. 127-135.

które zostaną im wskazane. Robot „żyje” w ignorancji (o co podejrzewał też ludzi już Sokrates i Platon), ale w szczególnej sytuacji możemy sobie wyobrazić, że „jego” „decyzje” mogą nam pomóc. W końcu wielu z nas będzie mogło liczyć za jakiś czas tylko na pomoc opiekuńczego robota. Czy robot może wyróżniać się niektórymi cnotami? Czy powinniśmy budować świadomość robota, aby kontrolowała proces decyzji? Kant uważał, że życzliwość człowieka nie zależy od efektów jego działań, ponieważ istnieje zbyt wiele czynników, które na te działania wpływają, a na które nie mamy wpływu. Rozwój dobrej woli jest najważniejszym celem człowieka. Natomiast ocena robota zależy wyłącznie od efektów jego działań. Celem działania maszyny jest najlepsze dopasowanie do potrzeb człowieka, dostosowane i zoptymalizowane do konkretnych zadań. Działanie zgodne z etyką utylitaryzmu oznacza, że tylko wyniki lub konsekwencje danego czynu decydują o jego wartości moralnej. Roboty nie są ani bardziej, ani mniej moralne niż ich twórcy. Wydaje się, że etyka sytuacyjna podpowiada najprostsze rozwiązanie – podejmowanie decyzji powinno opierać się na przewidywaniu konsekwencji danego czynu<sup>21</sup>. W końcu predykcyjność jest jednym z podstawowych zadań dla AI.

Czy jesteśmy gotowi mieć wokół siebie roboty? Czy możemy im ufać bardziej niż naszym ludzkim sąsiadom? Obecnie testowane są na drogach autonomiczne pojazdy w USA. Testowane są też kryteria odpowiedzialności. W robotyce medycznej ważny jest fakt bezpośredniego kontaktu z ciałem pacjenta, bezpośredni wpływ na życie człowieka, którym opiekuje się robot. Daje to zestaw wyzwań projektowych i wyzwań komunikacyjnych. Komunikacja i możliwość popełniania błędów prowokują do stawiania pytań. Odpowiedzi wymagają analizy pod względem etycznym. Wizja, w której roboty mogłyby być partnerami starszych i niedołączonych pacjentów, bardzo nas fascynuje. Robot osobisty powinien zapewniać bezpieczeństwo i poziom wolności człowieka<sup>22</sup>. Populacja świata sięga już ośmiu miliardów. Towarzyszy nam niezliczona liczba żywych organizmów i kilka milionów coraz bardziej inteligentnych robotów. Robotyka medyczna będzie odgrywać kluczową rolę w przyszłości społeczeństw. Jednak rola ta nie jest pozbawiona wielu wątpliwości i kontrowersji. Inteligentne i samoświadome roboty niedługo staną się członkami (elementami) ludzkiego społeczeństwa<sup>23</sup>.

<sup>21</sup> Z. Nawrat, M. Kawalski, *Questions about a New Ethics – Ethics of Medical Robots Era*, „Medical Robotics Reports” 2(2013), s. 22-25.

<sup>22</sup> Zob. Z. Nawrat, M. Kawalski, *Ethics – Challenges and Questions in Personal and Medical Robotics*, w: *Etyka wobec współczesnych wyzwań*, red. A. Bobko, Uniwersytet Rzeszowski, Rzeszów 2013, s. 61-70.

<sup>23</sup> Por. Nawrat, *The Ethics of Artificial Organs*, s. 127-135.

## ROBOTY MEDYCZNE

Można z powodzeniem opisać historię człowieka przez historię jego narzędzi. Wyobraźmy sobie, że obiektem analizowanym jest człowiek z narzędziem w dłoni. Taka hybryda biologiczno-techniczna. Roboty i AI to po prostu kolejny krok w ewolucji. Czy człowiek musi się oswajać ze swoimi wytworami? AI powstało z inspiracji naturą i obecnie rozwija się, kontynuując naśladowanie pracy mózgu. Algorytmy genetyczne i sztuczna ewolucja próbują zastosować teorię ewolucji Karola Darwina do AI. Sztuczna ewolucja działa na populacji żywotnych genotypów (chromosomów), które są krzyżowane pod presją selekcyjną, podobnie jak w standardowej genetyce. W rezultacie pojawiają się bardziej ewolucyjnie uprzywilejowane architektury sterowania. Połączenie tych innowacyjnych technik spowoduje rozwój robotów, które uczą się, pamiętają, podlegają ewolucji.

Jeśli tylko uda się rozbić każde z zadań usług medycznych na proste, powtarzalne czynności czy decyzje i będziemy potrafili je zalgorytmizować, to będzie można zlecić te zadania urządzeniom, komputerom, robotom. Trzeba dobrze zrozumieć składowe procesy wykonywania danej usługi medycznej i obiektywnie opisać „ścieżki realizacji” – będziemy musieli tego nauczyć sztuczną inteligencję. Przez praktykowanie w relacji mistrz–uczeń, przez próby dążenia do doskonałości, uzyskuje się odpowiednie kwalifikacje. Uczucie maszynowe może pełnić podobną rolę jak doskonalenie studenta-praktykanta w czasie szkolenia w szpitalu. Nie należy się tego bać, ponieważ podobnie jak w przypadku szkolenia ludzi, tak i w przypadku ML oraz AI możemy, jeśli nie spełniają zdefiniowanych standardów, wykluczyć je z systemu. To człowiek włącza i wyłącza AI. Przynajmniej tak powinno być.

Do tej pory urządzenia i narzędzia medycyny były zorientowane na wykonawców usług, czyli na lekarzy i pielęgniarki. Obecnie mogą być zorientowane na pacjenta. To zmiana sposobu wykonywania usług medycznych. To także inne podejście projektowe – zasada ergonomicznego dopasowania narzędzia do użytkownika. Tarka do prania jest zorientowana na człowieka (musi być ergonomiczna i dobrana do praczki, czyli człowieka pracującego przy jej pomocy), zaś pralka automatyczna (posiadająca algorytmy sterowania oparte na AI śledzi nasze przyzwyczajenia, sama dobiera wagę i procedurę do materiału) jest zorientowana na pranie. W ten sposób wchodzimy w realizację nowego paradygmatu wykonywania usług medycznych. Od pacjent<math>\leftrightarrow</math>lekarz, przez pacjent>technologia>lekarz, do pacjent>technologia zdrowia, w którym to systemie lekarz pełni jedynie rolę kontrolną i innowacyjną. Wyzwania przyszłości nie będą wcale mniejsze niż te, które spotkały nasze pokolenie. Kreatywność, nauka będą niezbędne dla utrzymania dobrostanu jak największej liczby ludzi oraz rozwoju usług medycznych i opiekuńczych. Brakuje milionów pracow-

ników usług zdrowia, by utrzymać obecny poziom sprawowanej pomocy medycznej. Oczekiwania i potrzeby są znacznie większe. Kontakt kompetentnego pracownika z pacjentem, czyli relacja człowieka z człowiekiem ciągle jest wyznacznikiem komfortu w czasie, gdy mamy problemy zdrowotne. Automatyzacja sprzyja medycynie, bo uwalnia zasoby ludzkie do pełnienia chyba najważniejszej roli – udzielania pomocy człowiekowi i sprawowania opieki medycznej. Już w niedalekiej przyszłości większość zatrudnionych będzie pracowała w usługach zdrowia. Bo rzeczy naprawdę wielkie robimy z ciekawości, a rzeczy ważne – z potrzeby pomocy drugiemu człowiekowi. Lecząc chorego, przywracając go do życia i odczuwania radości, dajemy „wolność pominiętym”<sup>24</sup>. W medycynie nie chodzi o to, by żyć najdłużej, tylko by każdy miał szansę je przeżyć – wbrew chorobom, wypadkom czy innym zdarzeniom losowym. Stąd konieczność prowadzenia prac nad rozwojem robotów w medycynie.

To, co tworzymy, powinno spełniać zarówno kryteria piękna (sztuka), dobra (działalność publiczna), prawdy (nauka), jak i użyteczności (wynalazczość). Lekarz podejmuje decyzje, posiadając ograniczony zasób wiedzy dotyczący danego pacjenta. Dzięki wsparciu diagnozy i aktualizowanym analizom danych podczas zabiegu przez systemy AI zmniejszy się ryzyko popełnienia błędu. Medycyna oparta na AI i faktach będzie z założenia spersonalizowana (albo przez AI skolonizowana). Im mniejszy dostęp do informacji, im przykładowo mniejsza inwazyjność operacji, tym bardziej wzrasta rola sztucznej inteligencji.

Roboty medyczne stanowią jedyną szansę na spełnienie w niedalekiej przyszłości szerokiego, demokratycznego dostępu do dobrej jakości usług w zakresie opieki zdrowia. Zachodzące zmiany demograficzne, brak właściwej liczby personelu medycznego i konieczność podnoszenia standardów wykonywanych usług w szpitalach, przychodniach oraz w domach potrzebujących pomocy, to trendy światowe, które wymuszają większe zaangażowanie w dziedzinę robotyki medycznej, w tym w rozwój sztucznej inteligencji. Dzięki harmonijnemu rozwojowi nauk technicznych i biologicznych oraz ich rozsądnemu wdrażaniu będzie możliwość wpływania na ewolucję gatunków i jakość życia na Ziemi. Roboty i sztuczna inteligencja to tylko kolejne przystanki (techno)ewolucji. W znacznej mierze o jakości naszego życia już dzisiaj decydują roboty (roboty przemysłowe). Jednak wszystko wskazuje na to, że roboty medyczne (w tym sztuczna inteligencja) są niezbędne dla człowieka w przyszłości. R o b o t y

---

<sup>24</sup> Z. N a w r a t, *Dlaczego rzeczy naprawdę wielkie robimy z ciekawości, a rzeczy naprawdę ważne z potrzeby pomocy drugiemu człowiekowi?*, w: *Wystąpienia i wykłady inauguracyjne rok akademicki 2020/2021*, Instytut Problemów Współczesnej Cywilizacji im. Marka Dietricha, Warszawa 2021, s. 112-113.

będą stróżami jakości naszego życia. Staną się niezbędne jak bieżąca woda i Internet; będą gwarancją bezpiecznej przyszłości homo sapiens. Dlatego, w odpowiedzi na pojawiające się wyzwania, medycyna jest sceną pojawienia się (wykluwania) i tworzenia (wykuwania) się etosu robotów. Mogą one stać się dla nas szansą na uwolnienie się od wielu ograniczeń człowieka zdrowego i chorego. Roboty – czy też ogólnie: innowacje – powstają, aby pomnażać wolność człowieka.

Na zakończenie można się odnieść do myśli Jana Pawła II na temat efektów ludzkiej pracy. Jeżeli traktować usługę medyczną jako przejaw miłości bliźniego (zarówno jako cel pracy, jak i sposób jej realizacji), to tworzenie narzędzi podnoszących jakość i ułatwiających dostęp do pomocy medycznej jest ze wszech miar godnym pracy celem. Zgodnie z duchem i słowem encykliki *Laborem exercens* odślaniamy nowe znaczenia pracy ludzkiej w kontekście wykorzystania AI. Znajdziemy ślady prognozujące sposób przyjęcia, akceptacji, tych nowych narzędzi człowieka w słowach Jana Pawła II, bo roboty są najbardziej zaawansowanym osiągnięciem techniki, a ona stanowi przecież „zrodzony z myśli ludzkiej sprzymierzeniec pracy człowieka”<sup>25</sup>. „Technika rozumiana w tym wypadku nie jako podmiotowa umiejętność czy sprawność pracy, ale jako zespół narzędzi, którymi człowiek posługuje się przy pracy, jest niewątpliwie sprzymierzeńcem człowieka. Ułatwia mu pracę, usprawnia ją, przyspiesza i zwielokrotnia [...] Tak więc owocem pracy jest wszystko to, co ma służyć pracy, co stanowi – przy dzisiejszym stanie techniki – jej rozbudowane «narzędzie»”<sup>26</sup>. Wszystkie nadzieje i ryzyka związane z upowszechnieniem i nadaniem specjalnej roli robotyce i AI są związane z tym, że wykonują one realną pracę, a „praca ludzka stanowi klucz, i to chyba najistotniejszy klucz, do całej kwestii społecznej, jeżeli staramy się ją widzieć naprawdę pod kątem dobra człowieka”. Jan Paweł II zwrócił uwagę, że problem pracy musi być ciągle uwspółcześniany, bo decyduje o godności i warunkach życia człowieka (tu i teraz). Robot (w tym sztuczna inteligencja) jest narzędziem, wybitnym osiągnięciem ludzkiej pracy i jego niezwykłość, ryzyko, polega na tym, że może wykonywać pracę samodzielnie. Czyli ma nie tylko czysto praktyczne znaczenie (profesjonalne) ale również społeczne (socjologiczne i filozoficzne). Stoi przed nami zatem obecnie wyzwanie: jak umiejętnie korzystać z narzędzi, które wiążą się z robotami medycznymi i AI.

<sup>25</sup> Jan Paweł II, Encyklika *Laborem exercens*, nr 5, [https://www.vatican.va/content/john-paul-ii/pl/encyclicals/documents/hf\\_jp-ii\\_enc\\_14091981\\_laborem-exercens.html](https://www.vatican.va/content/john-paul-ii/pl/encyclicals/documents/hf_jp-ii_enc_14091981_laborem-exercens.html).

<sup>26</sup> Tamże.

## BIBLIOGRAFIA / BIBLIOGRAPHY

- Čapek, Karel. *R.U.R.: Rossum's Universal Robots*. Translated by Martyna M. Le-  
mańczyk. Warszawa: Fundacja Nowoczesna Polska, 2021.
- Encyklopedia Polskiego Wydawnictwa Naukowego*, s.v. "Praca." <https://encyklopedia.pwn.pl/haslo/praca;3961600.html>.
- Jan Paweł II. Encyklika *Laborem exercens*. [https://www.vatican.va/content/john-paul-ii/pl/encyclicals/documents/hf\\_jp-ii\\_enc\\_14091981\\_laborem-exercens.html](https://www.vatican.va/content/john-paul-ii/pl/encyclicals/documents/hf_jp-ii_enc_14091981_laborem-exercens.html).
- Kant, Immanuel. *Uzasadnienie metafizyki moralności*. Translated by Mściśław War-  
tenberg. Warszawa: PWN, 1953.
- Kaku, Michio. *Wizje czyli jak nauka zmieni świat w XXI w.* Warszawa: Prószyński  
i S-ka, 2000.
- Nawrat, Zbigniew. "Dlaczego rzeczy naprawdę wielkie robimy z ciekawości, a rzeczy  
naprawdę ważne z potrzeby pomocy drugiemu człowiekowi?" In: *Wystąpienia  
i wykłady inauguracyjne: rok akademicki 2020/2021*. Warszawa: Instytut Proble-  
mów Współczesnej Cywilizacji im. Marka Dietricha, 2021.
- . "Etyka ery sztucznych narządów." In: *Etyka wobec współczesnych wyzwań*.  
Edited by Aleksander Bobko. Rzeszów: Uniwersytet Rzeszowski, 2013.
- . "Roboty medyczne w systemach teleinformatycznych." In *Informatyka  
w medycynie*. Edited by Marek Kurzyński et al. Warszawa: Akademicka Oficyna  
Wydawnicza Exit, 2019.
- . "The Ethics of Artificial Organs." In *Implant Expert*. Edited by Zbigniew  
Nawrat. Zabrze: M-Studio 2011.
- Nawrat, Zbigniew, and Marcin Kawalski. "Ethics – Challenges and Questions in Per-  
sonal and Medical Robotics." In *Etyka wobec współczesnych wyzwań*. Edited by  
Aleksander Bobko. Rzeszów: Uniwersytet Rzeszowski, 2013.
- . "Questions about a New Ethics – Ethics of Medical Robots Era." *Medical  
Robotics Report* 2(2013): 22–25.
- Organisation for Economic Cooperation and Development. *Recommendation of the  
Council on Artificial Intelligence*. OECD Legal Instruments. [https://legalinstru-  
ments.oecd.org/en/instruments/oecd-legal-0449](https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449).
- Raporty OECD. [https://www.oecd-ilibrary.org/\\_Etyczna-Sztuczna-Inteligencja-  
rekomendacje-OECD.pdf](https://www.oecd-ilibrary.org/_Etyczna-Sztuczna-Inteligencja-rekomendacje-OECD.pdf).
- Sant'Anna, Roberto T., et al. *Robotic Systems in Cardiovascular Surgery*. "Revista  
Brasileira de Cirurgia Cardiovascular" 19, no. 2 (2004): 171–8. [https://www.scielo-  
br/j/rbccv/a/GDNJJkcTGxDZVcfj8cxhxgR/?lang=en](https://www.scielo.br/j/rbccv/a/GDNJJkcTGxDZVcfj8cxhxgR/?lang=en).
- Satava, Richard M. "Future Direction in Robotic Surgery." In *Surgical Robotics.  
System application and Visions*. Edited by Jacob Rosen, Blake Hannaford, and  
Richard M. Satava. New York, Dordrecht, Heidelberg and London: Springer,  
2011.
- du Sautoy, Marcus. *Kod kreatywności: Sztuka i innowacje w epoce sztucznej inteli-  
gencji*. Kraków: Copernicus Center Press, 2020.



- Sejnowski, Terrence J. *Deep Learning: Głęboka rewolucja*. Translated by Piotr Cypriański. Warszawa: Wydawnictwo Poltext, 2019.
- Topol, Eric. *Medycyna głęboka: Jak sztuczna inteligencja może ponownie uczynić opiekę zdrowotną ludzką*. Edited by Adrian Horzyk and Tomasz Stompór. Translated by Anna Boniszewska, Warszawa: ITEM Publishing, 2020.
- Warwick, Kevin. "The Diminishing Human-Machine Interface." *Medical Robotics Reports* 2 (2013): 4–11.
- [*Wykonywanie zawodu lekarza*] – *Zawody lekarza i lekarza dentystry*. Dz.U.2021.790 Art. 2. <https://sip.lex.pl/akty-prawne/dzu-dziennik-ustaw/zawody-lekarza-i-lekara-dentystry-16798282/art-2>.

## ABSTRAKT / ABSTRACT

Zbigniew NAWRAT – Medycyna i ethos robotów

DOI 10.12887/36-2023-4-144-12

W artykule autor analizuje zagadnienie medycyny innowacyjnej opartej na inteligentnych urządzeniach oraz relacjach między głównymi bohaterami tych (r)ewolucyjnych zmian: sztucznej inteligencji i robotyce. Do tej pory urządzenia i narzędzia medycyny były zorientowane na wykonawców usług. Obecnie urządzenia są elementem bezpośredniej relacji między lekarzem i pacjentem. W przyszłości będzie to relacja między pacjentem a robotem. Robot (w tym sztuczna inteligencja) jest narzędziem, które jest ważnym osiągnięciem ludzkiej pracy. Ryzyko związane z jego działaniem wiąże się z tym, że może on wykonywać pracę samodzielnie. Wykorzystanie robotów medycznych daje szansę na spełnienie (w niedalekiej przyszłości) szerokiego, demokratycznego dostępu do dobrej jakości usług w zakresie terapii i szybkiego powrotu do zdrowia. We współczesnej medycynie w coraz większym stopniu docenia się rolę robotów. Roboty poszerzają zakres ludzkiej wolności.

Słowa kluczowe: sztuczna inteligencja (AI), uczenie maszynowe (ML), roboty, roboty medyczne (MR)

Kontakt: Zakład Biofizyki, Wydział Nauk Medycznych w Zabrze, Śląski Uniwersytet Medyczny, Katowice, ul. H. Jordana 19, 41-808 Zabrze; Fundacja Rozwoju Kardiologii im. prof. Zbigniewa Religi, ul. Wolności 345a, 41-800 Zabrze  
E-mail: znawrat@sum.edu.pl; nawrat@frk.pl  
Tel. 32 3735626  
ORCID: 0000-0003-0638-3789



Zbigniew NAWRAT, *Medicine and the Robotic Ethos*

DOI 10.12887/36-2023-4-144-12

In the article, the author analyses the issue of innovative medicine based on intelligent devices and the relationship between the main protagonists of these (r)evolutionary changes: artificial intelligence and robotics. Until now, medical devices and tools have been oriented towards service providers. Currently, devices are part of the direct relationship between the doctor and the patient. In the future, it will be the relationship between the patient and the robot. A robot (including artificial intelligence) is a tool, an outstanding achievement of human work. One of the risks related to robots is their ability to work autonomously. The use of medical robots creates a chance to achieve (in the near future) broad, democratic access to good quality services in the field of therapy and recovery. Medicine is an area where the robotic ethos is developed. Robots, or innovations in general, are created to increase human freedom.

Keywords: artificial intelligence (AI), machine learning (ML), robots, medical robots (MR)

Contact: Zakład Biofizyki, Wydział Nauk Medycznych w Zabrze, Śląski Uniwersytet Medyczny, Katowice, ul. H. Jordana 19, 41-808 Zabrze, Poland; Fundacja Rozwoju Kardiologii im. prof. Zbigniewa Religi, ul. Wolności 345a, 41-800 Zabrze, Poland

E-mail: znawrat@sum.edu.pl; nawrat@frk.pl

Phone: +48 32 3735626

ORCID: 0000-0003-0638-3789



KAROL WOJTYŁA – JAN PAWEŁ II  
INSPIRACJE



Władysław ZUZIĄK  
Barbara ŻMUDA-FRYDYRCHOWSKA

## MAN'S PARTICIPATION IN THE WORLD WOJTYŁA—RORTY

*Rorty limits his considerations to the pragmatic dimension and—referring to a liberal vision of man—aims to maintain the status quo of American democracy. Wojtyła, whose proposal is based on transcendence and personalistic assumptions, wants to develop an axiological model aimed at forming a responsible human being who, in the process of self-perfection, transcends both his internal and external limitations (the latter being conditioned by a community). The confrontation of both thinkers' reflection reveals the limitations of all attempts at creating a synthesising and one-dimensional approach to communal reality.*

The philosophical thought of Richard Rorty and Karol Wojtyła represent two diametrically different approaches to the world and man. Although they are based on opposite ontological and anthropological assumptions, they do share certain points, which justifies their comparison and facilitates demonstrating their theoretical and practical strengths and weaknesses. The elements they share include the assumptions in the axiological sphere regarding the communal “nature” of man, both philosophers' sensitivity to suffering as well as their proposals how to reduce its scale.

Wojtyła tries to save Thomism, which seems to be a theory which is resistant to the changing reality and prone to dogmatic solutions, by developing and dynamizing it. It is worth emphasizing that his new approach to the dynamic condition of man in the world led to, among others, the creation of the Solidarity movement.

Rorty is more radical in his opposition to dogmatism. His aim is to eliminate a dogmatic—in his opinion—definition of truth as correspondence to reality. He believes that the only criterion of truth is the criterion of utility. It is worth noting, however, that his pragmatism is not instrumental and egoistic, because, if it were such, it would not be able to justify how sensitivity to the suffering of others can be useful to man.

In this article, we attempt to argue that Rorty limits his considerations to the pragmatic dimension and – referring to a liberal vision of man—aims to maintain the *status quo* of American democracy. Whereas Wojtyła, whose proposal is based on transcendence and personalistic assumptions, wants to develop an axiological model aimed at forming a responsible human being who, in the process of self-perfection, transcends both his internal and external limitations (the latter being conditioned by a community).

The confrontation of both thinkers' reflection reveals the limitations of all attempts at creating a synthesising and one-dimensional approach to communal reality. The category of "participation" (Latin: *participatio*) has a rich history. This term was popularised by the Stoics, and within Christian philosophy it was analysed by, among others, St. Augustine, St. Ambrose, and St. Thomas Aquinas. Karol Wojtyła addressed this problem and developed his own theory of participation. For both of them, the notion of participation is related to the issue of solidarity.

#### RORTY A COMMUNITY OF THE ADMITTED

The sources of contemporary disintegration of the axiological order include the processes which detach man from his environment, technological advancements, industrialisation, marketization processes, as well as an increasing impact of the media which nowadays mediate man's experience of reality. These factors have led to a situation in which man has lost his natural support in reality. Participating in various spheres of social life, fulfilling various functions connected with them, man does not become attached to these functions, nor does he form bonds with people with whom he has to interact. As a result, he both alienates himself from others and disintegrates himself.

At the same time, the process of disenchantment, i.e. desacralisation of the world, as Max Weber called it, continues and deepens. For some people, this process is a sign of progress and man's liberation from metaphysical and religious superstitions, while for others it is an opportunity to reflect on the emergence of new superstitions. Richard Rorty, a very interesting postmodernist thinker, not only criticises previous "superstitions" but also tries to show new moral perspectives and possibilities of integration in the world devoid of all certainties.

The postmodern fear of the Enlightened deification of Reason and its extreme outcomes, fascism, and communism, directs Rorty towards rejecting all previous metanarratives, which, in his opinion, were created in the past by religion, science, and poetry to equal extents.

Each of these areas adopts some initial assumptions, some values, and some criteria that distinguish them from the others, but it is difficult to find a metacriterion that would allow deciding which of them has the primacy over other cultural proposals. As Rorty observes: "Defenders of the idea that there is a methodological difference between artistic, political, and scientific revolutions typically adopt a strong, criterial notion of rationality, one in which rationality is a matter of abiding by explicit principles. They thus find them-

selves, willy-nilly, questioning the ‘rationality’ of the rest of culture.”<sup>1</sup> Each culture produces narratives through which it is more or less successful, but nothing more.<sup>2</sup>

In the absence of a conclusive criterion, Rorty claims that there is no basis which can give us definitive answers to the questions what truth is, what goodness is, or what constitutes the true nature of the world. Any attempt at finding an answer leads to sterile discussions, since “there is no natural order of justification of beliefs, no predestined outline for argument to trace”<sup>3</sup>. Nor are we able to decide whether our fate is determined by historical necessity or purposefulness or a series of coincidences.

Therefore, we can present “both intellectual and moral progress not as a matter of getting closer to the True or the Good or the Right, but as an increase in imaginative power. [...] Imagination is the source both of new scientific pictures of the physical universe and of new conceptions of possible communities.”<sup>4</sup> Imagination in intellectual and moral development is, in Rorty’s opinion, supposed to be a defence against overbearing attempts of specialists in the only truth and the only right choices. Such a defence is to be provided by “philosophical superficiality and light-mindedness” that will make people “more pragmatic, more tolerant, more liberal, more receptive to the appeal of instrumental rationality.”<sup>5</sup> Another tool used in defence against dogmatism is irony,<sup>6</sup> which offers a possibility of experiencing community in multiplicity and diversity.

Rorty emphasizes that the objectivist tradition of seeking truth for its own sake prevents man from becoming attached to community. Consequently, man

---

<sup>1</sup> See Richard R o r t y, *Objectivity, Relativism, and Truth* (Cambridge: Cambridge University Press, 1998), 48. A similar view on rationality in science is expressed by Józef Tischner: “The model of rationality proposed by the scientific reason and elevated by it to the rank of the only obligatory model creates the illusion that all other models of rationality are irrational. The scientific reason does not want to allow for the pluralism of rationalities, which corresponds to the pluralism of reality” (Józef T i s c h n e r, *Myślenie według wartości* (Kraków: Znak, 1982), 446). Unless indicated otherwise, the translations are ours.

<sup>2</sup> “There is no synoptic view of culture which is more than a narrative account of how various cultures managed to get to where they now are” (R o r t y, *Objectivity, Relativism, and Truth*, 92).

<sup>3</sup> *Ibidem*, 193.

<sup>4</sup> Richard R o r t y, “Ethics without Universal Obligations,” in: *Pragmatism as Anti-Authoritarianism* (Harvard: The Belknap Press of Harvard University Press, 2021, 142). “From an ethico-political angle, however, one can say that what is characteristic, not of the human species but merely of its most advanced, sophisticated subspecies – the well-read, tolerant, conversable inhabitant of a free society—is the desire to dream up as many new contexts as possible” (R o r t y, *Objectivity, Relativism, and Truth*, 110).

<sup>5</sup> *Ibidem*, 193.

<sup>6</sup> See Andrzej S z a h a j, *Ironia i miłość* (Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, 2012), 92.



is convinced that what is contingent must be suspended in order to reach the ultimate truth, which lies outside the world in which we live. Rorty disagrees with this view and argues that we should seek answers to the questions regarding truth and rationality with reference to a particular community in which these values will be for us “a compliment paid to the beliefs which we think so well justified that, for the moment, further justification is not needed.”<sup>7</sup> Here we could mention Kuhn’s concept of paradigm, in which a community, thanks to shared practices and the assumptions underlying their implementation, participates in its own development and in achieving goals stemming from a particular paradigm.

By rejecting all hypotheses which explain human desires and even the legitimacy of the existence of such desires, Rorty’s pragmatism would, as he himself believes, protect the existing structure of liberalism and justify the value of “good old” American democracy. An important assumption of this idea is that “it is no more evident that democratic institutions are to be measured by the sort of person they create than that they are to be measured by anything more specific than the moral intuitions of the particular historical community that has created those institutions.”<sup>8</sup> In this approach, man is the product of a historical community, and so, in Rorty’s view, the dignity of man can be legitimately reduced to “the dignity of a group with which a person identifies herself.”<sup>9</sup> In his opinion, “words like ‘kindness’ or ‘dignity’ do not form a vocabulary which all human beings can reach by reflection on their nature.”<sup>10</sup> This approach, as Marek Czachorowski aptly observed, is linked to the assumption that truth should not be required in understanding man, as each individual can develop such a model of “I” that would suit him and be adapted to his individual worldview, political preferences, or religion.<sup>11</sup> In this view, “the only ‘we’ we need is a local and temporary one: ‘we’ means something like ‘us 20th-century Western social democrats.’”<sup>12</sup> However, adopting this understanding of the notions of “I” and “we” will not allow us to explain any of the changes that have taken place in human history and have led to the formation of the “20th-century social democrats.” Here Rorty reduces the layers of the human being to those that are useful in a particular historical moment. The human person is but a point in the mass, a chain of molecules.

<sup>7</sup> R o r t y, *Objectivity, Relativism, and Truth*, 24.

<sup>8</sup> *Ibidem*, 190.

<sup>9</sup> *Ibidem*, 200.

<sup>10</sup> Richard R o r t y, *Contingency, Irony, and Solidarity* (Cambridge: Cambridge University Press, 1995), 93.

<sup>11</sup> See Marek C z a c h o r o w s k i, “Personalistyczna etyka Karola Wojtyły wobec zarzutów współczesnego antypersonalizmu,” *Logos i Ethos* 53, no. 1 (2020): 103.

<sup>12</sup> R o r t y, *Objectivity, Relativism, and Truth*, 214.

The philosopher opposes absolutizing and granting a universal status to what is merely local. His vision of the person is a vision of a concrete individual who is firmly embedded in the social reality that shapes her.<sup>13</sup>

Rorty does not seem satisfied with the outcome of this reduction, as he acknowledges that “even if the typical character types of liberal democracies are bland, calculating, petty, and unheroic, the prevalence of such people may be a reasonable price to pay for political freedom.”<sup>14</sup> The question is whether we would want to live in a society of bland, calculating and petty people and whether living among such people would provide us with a sense of security. Could participating in the life of such a society provide an impulse for development, for respect for others, for self-respect?

While reducing human dignity to the “dignity of the group” and, at the same time, moving beyond narrow pragmatism, Rorty views the issue of participation in moral social development in an interesting way.

In his opinion, morality “consists of those beliefs and desires and emotions which overlap with those of most other members of some community with which for purposes of deliberation, she identifies herself, and which contrast with those of most members of other communities with which hers contrasts itself”.<sup>15</sup> Thus, participation in a community would be accomplished through identification with a group or nation that is confronted with a different community. Similar distinctions were made in ancient Greece, where communities of Hellenic city-states were distinguished from “the barbarians” surrounding Greece.

Such identification may justify the existence of the social status quo, however, in the contemporary, dynamic, multicultural world it does not allow neither understanding of social relations nor expanding a community based on some historical-axiological basis by admitting other participants into its social life. Thus, Rorty accepts that “the idea of human beings striving for something supra-human should be replaced by the idea of welcoming more and more people into our community—taking into account the needs, interests and views of as many different people as possible.”<sup>16</sup> However, expanding a community simply for the sake of its expansion would be problematic for a historical community if there were no criterion for such expansion.

Lacking the possibility to decide what is good and true, Rorty presents a programme of moral development based on a specific criterion: “The problem with the pursuit of truth is that it is impossible to know when one has reached it,

---

<sup>13</sup> See Zbigniew Ambrożewicz, “Holizm, indywidualizm, personalizm: Od dualizmu platońskiego do sporu o źródło i status ludzkiej jednostkowości (N. Elias, I. Berlin, R. Rorty),” *Analiza i Egzystencja* 30 (2015): 103n.

<sup>14</sup> Rorty, *Objectivity, Relativism, and Truth*, 190.

<sup>15</sup> *Ibidem*, 200.

<sup>16</sup> Rorty, *Ethics without Universal Obligations*, 135.

even if one has actually succeeded.... We can, however, strive for ever greater sensitivity to suffering and ever fuller satisfaction of ever more diverse needs".<sup>17</sup> Here we see a break in Rorty's pragmatism, which stems from his assumption that subjectivity is primarily defined by vulnerability to pain and humiliation and the satisfaction of needs. Thus, "it is best to think of moral progress as a matter of increasing sensitivity, an increasing responsiveness to the needs of a larger and larger variety of people and things."<sup>18</sup> Participation in a community would thus be a process associated with the ability to widen the circle of people with whom we spontaneously feel connected. Rorty assumes that each of us has the capacity to empathise with the other, or at least to be able to imagine the other's feelings in analogy with our own feelings.<sup>19</sup> This assumption, and the assumption of universal empathy, directs him towards idealised social constructs. As Szahaj, a prominent postmodern philosopher observes, Rorty sees man as someone who seeks love and—in its name—is capable of building a utopia of harmonious coexistence, devoid of injustice and domination.<sup>20</sup>

Summing up, the moral progress that is supposed to unite a dynamic society is a matter of ever-growing empathy, not rationality.<sup>21</sup> The morality that makes it possible to participate in a community consists in the ability to empathise with a person who is suffering. The attitude which makes it possible to identify with the sufferer is not a universal one, since we usually identify with a particular sufferer (possibly with a group) rather than with all sufferers, but it is possible, by stimulating sensitivity, to extend the group which we consider to be "our" group so as to extend our understanding of "we" to persons whom we previously considered to be "them". The basis for such inclusion is not abstract recognition of all people as members of humanity but direct interpersonal relationships that allow people to include themselves in their community.

#### WOJTYŁA

#### A COMMUNITY OF PERSONS

Arguably, Rorty would agree with Karol Wojtyła's view that participation in a community comes down to this feature of the human person thanks to which—by existing and acting "together with others" in a wide range of interpersonal and social relations—he is able to be himself and to fulfil himself.

<sup>17</sup> Ibidem.

<sup>18</sup> Ibidem.

<sup>19</sup> See Colin Koopman, "Rorty's Moral Philosophy for Liberal Democratic Culture," *Contemporary Pragmatism* 4, no. 2 (2007): 48–51.

<sup>20</sup> See Szahaj, *Ironia i miłość*, 95.

<sup>21</sup> See Rorty, *Ethics without Universal Obligations*, 135.

However, in Wojtyła's opinion, human social life takes the form of a community which realises the common good not only in the objective dimension but primarily in the subjective dimension. Man as a person is a specific subject of social life, and his spiritual development requires a communal form of this life. Rorty would probably argue with this point, claiming that a community itself is a fundamental value and that we are unable to decide whether any spiritual basis for an individual's development exists at all.

In Rorty's opinion, an individual's development takes place in the dimension of the relation between "We" and "I." The "I" identifies itself with the "We" and shapes its ideas regarding the aims of its existence on the basis of the models of survival developed within a historical community.

Wojtyła, however, notices two basic meanings of participation in society based on interpersonal relations: "I"—"you" and "I"—"we."<sup>22</sup> The first meaning is related to "participation in the humanity of other people."<sup>23</sup> Humanity is not an abstract idea of man but a unique personal "I," which most closely corresponds to the idea of a "neighbour." It is "humanity itself which is possessed by every 'other' man just as 'I' myself possess it."<sup>24</sup> Participation in "the humanity of every man" determines "the personalistic value in the community of being and acting."<sup>25</sup> Participation here means essential personalisation of man's relation to the other. When someone experiences the other as a person, he comes as close as possible to what constitutes his personal "I" as a unique and unrepeatable reality.<sup>26</sup>

Such participation reveals man's capacity for spontaneous and positive opening towards others. However, the existence and action of a particular man together with others is a task which, as Wojtyła observes, needs a certain impulse. This impulse can arise from the evangelical commandment of love, from which directly follows that "participation in the very humanity of every man is the core of all participation."<sup>27</sup> According to Wojtyła, the commandment of love allows a community to be fully human, moreover, "if any human community impairs this system of reference, it dooms itself to a disappearance of the fullness of participation, to an abyss between person and community."<sup>28</sup> It can be noticed that both Wojtyła and Rorty emphasize the essential value

---

<sup>22</sup> Karol Wojtyła, "Participation or Alienation," in Wojtyła, *Person and Act and Related Essays*, trans. Grzegorz Ignatik (Washington, D.C.: The Catholic University of America Press, 2021), 514–31.

<sup>23</sup> *Ibidem*, 520.

<sup>24</sup> Karol Wojtyła, "An Outline of the Theory of Participation," in Wojtyła, *Person and Act and Related Essays*, 409.

<sup>25</sup> *Ibidem*.

<sup>26</sup> Wojtyła, "Participation or Alienation," 521.

<sup>27</sup> Wojtyła, "An Outline of the Theory of Participation," 410.

<sup>28</sup> *Ibidem*, 414.

of sensitivity to the other in the process of participation in a community and in its development, however, Wojtyła's considerations in this area seem more profound than Rorty's.

Participation, according to Wojtyła, is a disposition that makes a human being a person. However, man can renounce this disposition. He can act "together with others" without necessarily co-participating. As Wojtyła points out, by participating, man "co-acts" while performing some acts and realising himself in them. There are two dimensions of participation: an interpersonal one and a social one. Rorty's thought misses the former, although participation in the humanity of one's neighbour ("I"—"you") seems close to the notion of empathy and to the building of a community through emotional bonds. Yet, even though community is a fundamental category for Rorty, the idea of the participation in the common good ("we") is absent from his considerations.

Wojtyła emphasizes that participation is a specific constitutum, an essential property of community, thanks to which a person and a community "adhere" to one another.<sup>29</sup> Although man is a member of various communities within which he lives and acts with others, membership in these communities is not the same as participation. From the point of view of a person and his act, it is not only an "objective community of acting" that is important (he acts with others although he remains outside this community, he does not identify with its aims, he does not fulfil himself in common action), but also a "subjective community of acting" which is linked to the choice taken.<sup>30</sup>

Wojtyła, with whom Rorty would certainly agree, states: "Man chooses what others choose or even when he chooses because others choose, seeing in this object of choice the value that is in some way homogeneous and his own. Self-determination is linked to this—and in the case of acting 'together with others,' self-determination includes and expresses participation."<sup>31</sup> Here we can see the pragmatic sources of communal thinking shared by both philosophers.

However, Wojtyła observes that without authentic participation human action can turn into *passio*—"the happening"—which takes place under the influence of others, of which the best example is the mechanism of a "lemming-like rush." Acting together with others without participating deprives a person's act of its personalistic value since it restricts the possibility of his self-determination.<sup>32</sup> The condition necessary for participation is personal ful-

<sup>29</sup> See *ibidem*, 393.

<sup>30</sup> See *ibidem*, 398.

<sup>31</sup> *Ibidem*, 387.

<sup>32</sup> See *ibidem*, 388n.

filment, while its opposite is alienation, which deprives man as a person of the opportunity to realise himself in communal action.

Both individualism and collectivism can be sources of alienation. Individualism advocates the good of an individual as the primary good and demands subordination of all communities and societies to it. In individualism acting together with others is assumed to be a necessity, but this form of co-acting does not develop an individual and is a source of constraints on his development.

Collectivism, the opposite of individualism, subordinates an individual and his wellbeing to society. It is built on the assumption that as an individual pursues his particular interests, it is necessary to secure the possibility of creating the common good and of harmonious cooperation with others by limiting individuals' rights. Alienation, which originates in individualism and collectivism, is an opposite of those structures within which an individual, through participation, can develop and shape himself by referring to common good.

Wojtyła finds a way of overcoming these forms of alienation in solidarity, which is a "natural" form of cooperation with others and, at the same time, the foundation of a community. For him, "a man who lives in solidarity performs what belongs to him not only on account of his membership in a community but also 'for the good of the whole.'"<sup>33</sup> Moreover, "in virtue of this attitude, man finds his own fulfilment in complementing others."<sup>34</sup> References to the common good, to others and to self-fulfilment are necessary elements of a person's authentic development and of his participation in a community.

However, such authentic participation does not rule out opposition when an individual finds the way in which the common good is defined and realised inappropriately. In this case, an attitude of opposition is constructive and even constitutes a condition for the proper functioning of a community.<sup>35</sup>

What links an attitude of solidarity to an attitude of opposition to a particular realisation of the common good is dialogue. Dialogue allows opposition to contribute to the formation and deepening of human solidarity, since it helps to bring out what is true and right, what can be a source of the good for people by overcoming subjective attitudes and prejudices.<sup>36</sup>

Wojtyła contrasts the attitudes of authentic solidarity and genuine opposition with the attitudes of conformism and avoidance. A conformist may support those in power if he sees in it his own advantage, or he may avoid doing anything if he decides that supporting them is disadvantageous for him. In conformism, authentic participation is replaced by "a pretense of participation,

<sup>33</sup> *Ibidem*, 401.

<sup>34</sup> *Ibidem*, 402.

<sup>35</sup> See *ibidem*, 401f.

<sup>36</sup> See *ibidem*, 402f.



a superficial adjustment to others, without conviction and without authentic commitment.<sup>37</sup> Conformism is a specific form of individualism, an escape from community which, by replacing authentic participation with a world of appearances, prevents an individual from fulfilling himself.

The same is true of avoidance, which is now often termed “internal emigration.” This attitude excludes any participation, even in its superficial form, as in the case of conformism. Avoidance is often justified by the belief that participation is impossible and opposition ineffective, but such an attitude leads to an individual’s abandoning his fulfilment in acting “together with others.”<sup>38</sup>

As a personalist, Karol Wojtyła emphasizes that the human person possesses natural dignity and unique individuality. Although—following Boethius—he understands the person as a *suppositum*, he also points out that man is not a being-for-itself,<sup>39</sup> because he is called to action and self-fulfilment, which can take place “together with others,” through co-existence and co-action.<sup>40</sup>

Wojtyła treats the human person as a contingent, potential, and not self-sufficient being, and hence, by his very nature, a being open to another person and a community of persons. A community, or rather the communion of persons, means living in a shared “for,” i.e., in a relationship of a mutual gift which replaces man’s original loneliness. Man fulfils himself through others and realises himself through other persons. In order to become oneself, man must participate in the life of a community and work for the common good, which is the axiological foundation of a community. Participation is realised only “when the person enters into a particular relationship with other persons and the common good.”<sup>41</sup>

Let us emphasize that what is meant here is authentic participation, in which the other person(s) is the target of action. The other person is a being endowed with interiority, a personal subject – a *neighbour*. The idea of a “neighbour” is related to the value of the person as such, independently of social references, and appears to us as the real good. Authentic attitudes, such as solidarity and opposition, are formed in this context. But if the person’s aim is self-interest, the other person becomes more of a competitor than a neighbour and, instead of working together for the common good, the members of a community begin to fight with one another. By separating his own good from the common good, such an individual “in a sense agrees to the fact that the community

<sup>37</sup> Ibidem, 405.

<sup>38</sup> Ibidem, 405f.

<sup>39</sup> See Karol Wojtyła, “Osoba – podmiot i wspólnota,” *Roczniki Filozoficzne* 24, no. 2 (1976): 13.

<sup>40</sup> See idem, “An Outline of the Theory of Participation,” 377n.

<sup>41</sup> Jan Gałarowicz, *Człowiek jest osobą: Podstawy antropologii filozoficznej Karola Wojtyły* (Kęty: Antyk, 2000), 275.



deprives him of himself,”<sup>42</sup> and, at the same time, “he deprives the community of himself.”<sup>43</sup> He becomes alienated from it and, at the same time, from the sources of his humanity. He deprives himself of the possibility of experiencing his own humanity in its fullness and of establishing a relationship of solidarity and communion with other people, for which he was created by God.

\*

Both thinkers, to a large extent, share communitarian views. They observe that the development of any social theory must begin with the protection of the sphere of human moral judgements. They also point out that there are better and more reliable fulcrums for this sphere than human will or a subjective belief that one is right. At the same time, they are against perceiving individuals as idealised, abstract entities, and argue that people should be treated as concrete entities, who exist in concrete political, cultural and social reality. They also emphasize the role played by a community in shaping man and the fact that particular individuals are rooted in tradition, religion and culture. They share the convictions that a community, in a way, precedes our projects and that sensitivity to the suffering of others is an impulse for the moral development of people and societies. And at this point it does not matter whether this empathy is innate to us or whether we create it.

However, unlike Rorty, Wojtyła emphasizes the relationship between the existence and functioning of communities and the existence in these communities of the specific common good against which an individual defines himself. Only when we acknowledge what the common good is, can we hope that, based on this good, a community will create the foundations of the axiological order in society.

Rorty is undoubtedly right when he argues that—to a great extent—we are the image of the existing social order. Society significantly “determines our consciousness”, including moral consciousness. It is a fact that man has no way of recognising his more deeply internalised spheres of the ‘I’ without having contact with others and being able to confront them. Without the belief that these spheres exist, man is condemned to deal with the most external sphere of the “I” and may even consider it to be his only reality.

What awaits us, citizens of the postmodern, democratic world, in the disenchanted world, devoid of transcendence and faith in human dignity?

---

<sup>42</sup> Wojtyła, “Person and Act,” 405.

<sup>43</sup> *Ibidem*.

Rorty's intellectual idealism offers us possible answers to this question. The defender of liberal democracy fails to see that democracy must be fought for, that the democratic status quo is under threat even in the United States and other fully-fledged democratic countries worldwide. Democracy cannot survive without people who voice their opposition in the name of the common good. Neither good-hearted, "bland, calculating, petty, and unheroic" people nor ironic intellectuals will make it possible. Contrary to what Rorty claims, man is more important for the development of societies than "moral institutions." Institutions are always vulnerable to manipulation and instrumentalisation of their message. At the end of the day, what really matters is the individual conscience of a person who is aware of his own dignity and of the importance of the values he is willing to defend. Decisions taken by many Ukrainians employed in Poland, who leave their secure jobs and return to their homeland to defend it against the aggressor, are indicative of this form of participation in the common good. Referring to the notion of "group dignity," especially within pragmatism, does not provide adequate explanation of a situation in which a person sacrifices his interests and decides to risk losing his life.

There is no room here for avoidance or compromise. Karol Wojtyła was an advocate of such an uncompromising form of participation. While Richard Rorty developed his projects in a somewhat illusory perspective of security, based on the belief in the power of the United States, Wojtyła based his thoughts on the experience of human confrontation with totalitarian regimes. We can see the validity of his conclusions in today's world in which democracy is at risk.

Rorty's thought allows including the people suffering in Ukraine in our social consciousness-sensitivity and treating them as "us," but it does not explain why Ukrainians give their lives for their country, which is the common good for them, regardless of the language they speak.

Rorty's social theory, on the one hand, deprives us of faith in truth, and on the other, appeals to hope.<sup>44</sup> However, without faith in bonds stronger than those pointed to by Rorty, without heroic, selfless people aware of their dignity, such hope seems to have no foundations. It is the vision of a perceptive but resigned man. Karol Wojtyła's idea of participation offers greater hopes for the survival of our world, based on the values to which we are attached.

---

<sup>44</sup> "Human solidarity is a matter of sharing a common selfish hope, the hope that one's world—the little things around which one has woven into one's final vocabulary—will not be destroyed." Richard R o r t y, "The Dangers of Over-Philosophication—Reply to Arcilla and Nicholson." *Educational Theory* 40, no. 1 (1990): 41.

## BIBLIOGRAPHY / BIBLIOGRAFIA

- Ambożewicz, Zbigniew. "Holizm, indywidualizm, personalizm: Od dualizmu platońskiego do sporu o źródło i status ludzkiej jednostkowości (N. Elias, I. Berlin, R. Rorty)." *Analiza i Egzystencja* no. 30 (2015): 89–108.
- Czachorowski, Marek. "Personalistyczna etyka Karola Wojtyły wobec zarzutów współczesnego antypersonalizmu." *Logos i Ethos* 53, no. 1 (2020): 97–117.
- Galarowicz, Jan. *Człowiek jest osobą: Podstawy antropologii filozoficznej Karola Wojtyły*. Kęty: Antyk, 2000.
- Koopman, Colin. "Rorty's Moral Philosophy for Liberal Democratic Culture." *Contemporary Pragmatism* 4, no. 2 (December 2007): 45–64.
- Rorty, Richard. *Contingency, Irony, and Solidarity*. Cambridge: Cambridge University Press, 1995.
- . "Ethics without Universal Obligations." In *Pragmatism as Anti-Authoritarianism*. Harvard: The Belknap Press of Harvard University Press, 2021, 126–143.
- . *Objectivity, Relativism, and Truth*. Cambridge: Cambridge University Press, 1998.
- . "The Dangers of Over-Philosophication—Reply to Arcilla and Nicholson." *Educational Theory* 40, no. 1 (1990): 41–44.
- Szahaj, Andrzej. *Ironia i miłość*. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika, 2012.
- Tischner, Józef. *Myślenie według wartości*. Kraków: Znak, 1982.
- Wojtyła, Karol. "An Outline of the Theory of Participation." In Wojtyła, *Person and Act and Related Essays*. Translated by Grzegorz Ignatik. Washington, D.C.: The Catholic University of America Press, 2021.
- . „Osoba – podmiot i wspólnota." *Roczniki Filozoficzne* 24, no. 2 (1976): 5–39.
- . "Participation or Alienation." In Wojtyła, *Person and Act and Related Essays*. Translated by Grzegorz Ignatik. Washington DC: The Catholic University of America Press, 2021.
- . "Person and Act." In Wojtyła, *Person and Act and Related Essays*. Translated by Grzegorz Ignatik. Washington DC: The Catholic University of America Press, 2021.

## ABSTRACT / ABSTRAKT

Władysław ZUZIĄK, Barbara ŻMUDA-FRYDRYCHOWSKA, Man's Participation in the World: Wojtyła—Rorty

DOI 10.12887/36-2023-4-144-13

The paper presents two diametrically different approaches to the world and man, based on opposite ontological and anthropological assumptions. However, these have some common points that allow them to be compared and show the strengths and weaknesses (theoretical and practical) of both proposals. Com-

mon assumptions in the axiological sphere include, among others, the communal “nature” of man, the sensitivity of both philosophers to suffering, and their proposals how to reduce the scale of this phenomenon. Wojtyła attempts to save the traditional Thomistic theory, which seems resistant to the changing reality, and criticizes the modernist designs of the world. He also emphasizes that the act is what reveals the essence of the human being as a person. New approaches to his dynamic thought on man in the world have resulted, among others, in the emergence of the solidarity movement. Rorty, who also criticizes modernist projects, is more radical. His aim is to eliminate thinking about truth as being in conformity with reality and to remove the oppositions between subject and object as well as objective and subjective. According to him, the only important criterion is the criterion of utility. However, it is worth noting that this is not an instrumental and egoistic pragmatism, because it would be difficult to justify in this trend how sensitivity to the suffering of others can be useful for a person. Moreover, we try to show that Rorty’s proposal is limited to the pragmatic dimension, aimed at maintaining the status quo of American democracy, while Wojtyła’s thought (referring to transcendence) is aimed at shaping human beings who are perfecting themselves, heroic, and going beyond their internal and external limitations.

Keywords: Karol Wojtyła, Richard Rorty, participation, axiology, sensitivity

Contact: Faculty of Philosophy, Pontifical University of John Paul II, ul. Kanoniczna 9/203, 31-002 Cracow, Poland

Phone: +48 12 8898 608

E-mail: (Władysław Zuziak) [wladyslaw.zuziak@upjp2.edu.pl](mailto:wladyslaw.zuziak@upjp2.edu.pl); (Barbara Żmuda-Frydrychowska) [barbara.zmuda-frydrychowska@upjp2.edu.pl](mailto:barbara.zmuda-frydrychowska@upjp2.edu.pl)

ORCID: (Władysław Zuziak) 0000-0002-0483-0533; (Barbara Żmuda-Frydrychowska) 0000-0001-9383-6090

Władysław ZUZIAK, Barbara ŻMUDA-FRYDRYCHOWSKA – Uczestnictwo człowieka w świecie: Karol Wojtyła – Richard Rorty

DOI 10.12887/36-2023-4-144-13

Przedstawiamy dwa diametralnie różne ujęcia świata i człowieka, oparte na przeciwstawnych założeniach ontologicznych i antropologicznych. Mają one jednak pewne punkty wspólne, które pozwalają je porównywać i ukazywać silne i słabe strony teoretyczne i praktyczne obu koncepcji. Wśród wspólnych założeń, w sferze aksjologicznej są między innymi wspólnotowa „natura” człowieka czy wrażliwość obu filozofów na cierpienie i wskazywanie dróg do zmniejszenia skali tego zjawiska.

Wojtyła stara się ratować skostniałą i odporną na zmieniającą się rzeczywistość teorię tomistyczną oraz poddaje krytyce modernistyczne projekty świata. Podkreśla przy tym, że czyn jest tym, co odsłania istotę człowieka jako osoby. Nowe ujęcia jego dynamicznej koncepcji człowieka w świecie zaowocowały

między innymi powstaniem ruchu solidarności. Rorty, również krytykujący projekty modernistyczne, jest bardziej radykalny. Celem, jaki stawia sobie, jest wyeliminowanie myślenia o prawdzie jako zgodności z rzeczywistością a także usunięcie opozycji podmiot/przedmiot, obiektywne/subiektywne itp. Według niego, jedynym ważnym kryterium, jest kryterium użyteczności. Nie jest to jednak pragmatyzm instrumentalny i egoistyczny, gdyż trudno byłoby w takim nurcie uzasadnić w jaki sposób wrażliwość na cierpienie innych może być dla człowieka użyteczna? Staramy się wykazać, że pragmatyczna koncepcja Rorty'ego prowadzi do ugruntowania postaw konformistycznych, podczas gdy odwołująca się do transcendencji koncepcja Wojtyły ma na celu kształtowanie człowieka doskonalącego się, heroicznego, przekraczającego ograniczenia zarówno własne, jak i zewnętrzne.

Słowa kluczowe: Karol Wojtyła, Richard Rorty, uczestnictwo, wrażliwość, aksjologia

Kontakt: Wydział Filozoficzny, Uniwersytet Papieski Jana Pawła II, ul. Kanoniczna 9/203, 31-002 Kraków

Tel. 12 8898 608

E-mail: (Władysław Zuziak) [wladyslaw.zuziak@upjp2.edu.pl](mailto:wladyslaw.zuziak@upjp2.edu.pl); (Barbara Żmuda-Frydrychowska) [barbara.zmuda-frydrychowska@upjp2.edu.pl](mailto:barbara.zmuda-frydrychowska@upjp2.edu.pl)

ORCID: (Władysław Zuziak) 0000-0002-0483-0533; (Barbara Żmuda-Frydrychowska) 0000-0001-9383-6090



MYŚLAĆ OJCZYŻNA...





Barbara TRYGAR

„WIELKA JEST POTĘGA PAMIĘCI”...  
„Mój Ojciec” Kazimierza Brauna

*Kolejne lata to wybuch drugiej wojny światowej, czas najtrudniejszy, kiedy jeszcze z większą uwagą należało wsłuchać się w głos sumienia, być aktywnym w służbie narodu i nieść pomoc drugiemu człowiekowi. Czytelnik odkrywa twarz Juliusza Brauna jako gospodarza w Milechowie, oblicze nauczyciela i wychowawcy, który uczył swoje dzieci patriotyzmu i miłości do Ojczyzny wskazując na postaci wielkich Polaków, którzy potrafili łączyć miłość do Ojczyzny z siłą, odwagą i moralnością. Okres wojenny to dla Brauna czas działań konspiracyjnych.*

Tylko że, aby drogę mierzyć przyszlą,  
Trzeba-ć koniecznie pomnieć, skąd się wyszło!<sup>1</sup>  
Cyprian Norwid

Cyprian Kamil Norwid, pisarz, artysta, człowiek in statu viae, który przemierzał świat w znaczeniu dosłownym, jak i metaforycznym, odbywał podróże w głąb siebie, zadając pytania: Kim jestem? Kim byłem? Gdzie szukać źródeł tożsamości? Solidarność z tym, co minione, pozwala odkryć prawdę o sobie, narodzie, sztuce. Autor *Promethidionu* konstatuje, że pamięć ma ogromną wartość, ponieważ tworzy tożsamość człowieka. Przywołuje ważne osoby i wydarzenia z przeszłości. W niej jest ukryta siła poznania teraźniejszości i tworzenia przyszłości. Św. Augustyn w dziesiątej księdze *Wyznań* napisał: „Wielka jest potęga pamięci. Trwogą przejmuję, Boże mój, to niezgłębione, nieskończone bogactwo. I to tym właśnie jest duch, tym jestem ja [...] a cokolwiek jest w pamięci, to jest i w duszy. Wszystkie te rzeczy (obrazy pamięci) przebiegam i przelatuję, jak zechcę, a nieraz się w nich zagłębiam, dokąd zdołam, lecz nigdzie nie znajduję dna. Taka jest siła pamięci, taka jest siła życia w człowieku żyjącym – chociaż jest on śmiertelny”<sup>2</sup>. Francuski uczoney Pierre Nora w artykule *Czas pamięci* konstatował, że wkroczyliśmy w erę pamięci, a odzyskiwanie pamięci stanowi integralną część naszej tożsamości<sup>3</sup>. Jan Assmann w książce *Pamięć kulturowa. Pismo, zapamiętywanie i polityczna tożsamość w cywilizacjach starożytnych* podkreślał, że „wszystko przemawia

<sup>1</sup> C. N o r w i d, *Niewola. Rapsod*, w: tenże, *Pisma wszystkie*, oprac. J.W. Gomulicki, t. 3, *Poematy*, Państwowy Instytut Wydawniczy, Warszawa 1971, s. 38.

<sup>2</sup> Św. A u g u s t y n, *Wyznania*, ks. X, 17, tłum. Z. Kubiak, Znak, Kraków 2007, s. 275.

<sup>3</sup> Zob. P. N o r a, *Czas pamięci*, tłum. W. Dłuski, „Res Publica Nowa” 2001, nr 7, s. 37-43; zob. też: „Ethos” 34(2021) nr 3(135) [*Pamięć*].

za tym, że kwestia pamięci wpłynie na wytworzenie nowego paradygmatu nauk o kulturze, który w nowym świetle postawi rozmaite zjawiska i dziedziny – sztukę, i literaturę, politykę i społeczeństwo, religię i prawo<sup>4</sup>. Bardzo interesującego porównania pamięci do lustra użyła Clare Mac Cumhaill w artykule *Still Life, a Mirror: Phasic Memory and Re-encounters with Artworks*. Pamięć fazowa w odróżnieniu od pamięci epizodycznej pozwala nam przypominać nie tyle wydarzenia z przeszłości, co nas samych i doświadczać na nowo, „jak to było być sobą” we wcześniejszych latach swojego życia<sup>5</sup>. Takim lustrem są wspomnienia zatytułowane *Mój Ojciec*<sup>6</sup>, w których Kazimierz Braun odkrywa siebie na nowo, przygląda się bliskim, których nie ma, ale wciąż żyją w jego sercu.

Kazimierz Braun urodził się 29 czerwca 1936 roku w Mokrsku Dolnym, reżyser, pisarz, uczoney. W 1958 roku ukończył Uniwersytet im. Adama Mickiewicza w Poznaniu na kierunku polonistycznym i w 1962 roku Państwową Wyższą Szkołę Zawodową im. Aleksandra Zelwerowicza w Warszawie. W 1971 roku obronił pracę doktorską na Uniwersytecie Adama Mickiewicza w Poznaniu, stopień doktora habilitowanego uzyskał w 1975 roku na Uniwersytecie Wrocławskim, a w roku 1992 został mu nadany tytuł profesora. W roku 1989 otrzymał tytuł profesora uniwersytetu w Buffalo. Jako reżyser teatralny zadebiutował w roku 1961, a w 1962 miał miejsce jego debiut telewizyjny. Reżyserował w Teatrze Polskim w Warszawie, Teatrze Wybrzeże w Gdańsku, Teatrze im. Juliusza Słowackiego w Krakowie, Teatrze im. Wiliama Horzycy w Toruniu, Teatrze im. Ludwika Solskiego w Tarnowie, a także w teatrach w Stanach Zjednoczonych, Irlandii, Kanadzie i Niemczech. W latach 1967-1971 był dyrektorem artystycznym Teatru im. Juliusza Osterwy w Lublinie, a w latach 1971-1974 jego dyrektorem naczelnym i artystycznym; w latach 1975-1985 był dyrektorem Wrocławskiego Teatru Współczesnego. Od roku 1985 mieszka w Stanach Zjednoczonych, gdzie reżyserował między innymi w Guthrie Theater w Minneapolis, Odyssey Theater w Los Angeles, Chicago Actors Ensemble i Kavinoky Theater w Buffalo. Jego najwybitniejsze prace reżyserskie to sztuki Cypriana Kamilla Norwida, Tadeusza Różewicza, Williama Szekspira i Adama Mickiewicza (*Dziady*). Pracował jako wykładowca na amerykańskich uczelniach: New York University, Uniwersytecie Kalifornijskim i uniwersytecie w Buffalo. Prowadził warsztaty reżyserskie w USA, Polsce, Irlandii i Wielkiej Brytanii<sup>7</sup>. Wyreży-

<sup>4</sup> J. A s s m a n n, *Pamięć kulturowa. Pismo, zapamiętywanie i polityczna tożsamość w cywilizacjach starożytnych*, tłum. A. Kryczyńska-Pham, Wydawnictwo UW, Warszawa 2008, s. 27.

<sup>5</sup> Zob. C. M a c C u m h a i l l, *Still Life, a Mirror: Phasic Memory and Re-encounters with Artworks*, „Review of Philosophy and Psychology” 11(2020), nr 2, s. 423-446.

<sup>6</sup> Zob. K. B r a u n, *Mój Ojciec*, Wydawnictwo Adam Marszałek, Toruń 2023.

<sup>7</sup> Zob. B. T r y g a r, „Ja” w drodze do Wolności, Prawdy i Piękna. *Aspekty aksjologiczne twórczości prozatorskiej Kazimierza Brauna*, Wydawnictwo Uniwersytetu Rzeszowskiego, Rzeszów 2019.

serował ponad sto pięćdziesiąt widowisk teatralnych i telewizyjnych. Wydał ponad pięćdziesiąt książek, w tym prace naukowe, powieści, poezje i dramaty. W opublikowanej w ostatnio monografii *Mój Ojciec* wspomina swojego ojca Juliusza Henryka Brauna (1904-1990), wybitnego prawnika, ekonomistę i uczonego. We *Wprowadzeniu* do książki pisze: „Stoję na górze. Patrzę. Ojciec mknie na nartach w dół stoku, przy każdym skręcie wznosząc wielkie wachlarze śniegu prześwietlonego słońcem. Dzielny. Umiejętny. Piękny w tym słońcu i pędzie. Duma. Radość. Mój ojciec. Moment na zawsze zapamiętanej synowskiej miłości”<sup>8</sup>. Był to najpiękniejszy moment zapamiętany przez syna. Obraz ten zapisał się na trwale w jego pamięci. Mimo różnych trudnych doświadczeń życiowych, które spotykały rodzinę Braunów, Kazimierz zawsze z wielką dumą patrzył na swojego ojca, który posiadał ogromną wiarę w siłę Prawdy, Dobra, które tworzą Piękno życia. W życiu rodziny Braunów można zauważyć pewne napięcie między linią ziemską, horyzontalną, a linią nadziemską, pionową, prowadzącą od ziemi ku niebu, od nieba ku ziemi tworzącą podstawowy wymiar zawartego w jego wspomnieniach obrazu świata. W ich przecięciu rozgrywa się dramat człowieka, który ukazuje tęsknotę za utraconą „małą ojczyzną”. Kazimierz Braun odczuwał wewnętrzną potrzebę udokumentowania doświadczeń, w szczególności tych najboleśniejszych związanych z wybuchem wojny i późniejszymi latami powojennymi. Zgromadził wartościową merytorycznie literaturę na ten temat, dokumenty z Instytutu Pamięci Narodowej, przeprowadził rozmowy z wybitnymi osobami i odbył wiele podróży do miejsc związanych z przeszłością. Życiorys jego ojca, Juliusza Brauna, uporządkowany jest chronologicznie, akcentuje przede wszystkim okresy życia rodzinnego, edukacji, pracy zawodowej i naukowej. Ujmuje postać ojca wieloaspektowo i charakteryzuje go następująco: „«Wielowarsztatowiec», parający się czynnościami redaktorskimi obok właściwych zajęć – jako uczeń, później jako prawnik i ekonomista, jako pracownik samorządu gospodarczego i adwokat, jako więzień polityczny i wreszcie jako człowiek nauki. Et haec facienda et haec non omittenda»”<sup>9</sup>. Swoją pracą i postawą Juliusz Braun wpisał się w historię Polski. W życiu i pracy był bezkompromisowy, wierny swoim przekonaniom i ideałom.

Jak podkreśla Braun, jego pamięć złożona jest niejako z dwóch części: pierwsza obejmuje wydarzenia, w których uczestniczył i które przeżywał, czyli wspólne rozmowy z rodzicami, naukę w domu i szkole, modlitwę z najbliższymi, posługę ministrancką w bolomińskim kościele, święta spędzone w gronie rodzinnym. Druga część wskazuje na świat, o którym rodzice opowiedzieli

<sup>8</sup> B r a u n, dz. cyt., s. 8.

<sup>9</sup> Tamże, s. 20n.

młodemu Braunowi, życie ich dwojga i rodziny, historię kraju, dzieje kultury europejskiej.

Juliusz Braun urodził się 2 lipca 1904 roku w Dąbrowie Tarnowskiej. Jego rodzicami byli Karol Braun (1868-1935) i Henryka z domu Miller (1871-1954)<sup>10</sup>. Miał dwóch starszych braci: Kazimierza (1899-1920) i Jerzego (1901-1975), i młodszą siostrę Jadwigę (1907-1996). Na ich życie składały się trzy najważniejsze elementy: praca i nauka, atmosfera artystyczna oraz życie towarzyskie. Ważna była niedzielna Msza Święta i uroczystości w domu, takie jak święta Bożego Narodzenia i Wielkanocy. Ojciec Juliusza, Karol Braun był prawnikiem, angażował się również głęboko w sprawy społeczne. Z jego inicjatywy powstał budynek Sokolni w Dąbrowie Tarnowskiej<sup>11</sup>. Następnie został prezesem „gniazda” Sokoła w Dąbrowie Tarnowskiej. Kazimierz Braun przywołuje piękne wspomnienie z tych wydarzeń: „Na Złot Grunwaldzki, zwołany przez Ignacego Paderewskiego z okazji odsłonięcia Pomnika Grunwaldzkiego w Krakowie, który ufundował w 1910 roku na 400-lecie zwycięskiej bitwy z Krzyżakami, pojechała cała rodzina Braunów, tak że mały Juliusz widział swego ojca defilującego przez miasto, a potem swoją matkę biorącą udział w wielkim pokazie gimnastyki zbiorowej na błoniach. Na czele defilady szła grupa prezesów gniazd Sokoła. Prowadził ją Karol Braun – w uniformie sokolskim, w kapeluszu z orlim piórem. Jego fotografia z tej defilady stała się po latach wręcz rodzinną relikwią”<sup>12</sup>. Warto jeszcze przywołać wspomnienia spisane przez Juliusza zawarte w notatkach: „Pochód sokoli obserwowany z okien domu znajomych przy ulicy Wolskiej w Krakowie. Pamiętam jak dziś: [...] maszerował otwierający pochód oddział prezesów Gniazd Sokolich z czerwonymi opaskami na ramieniu. A na czele, kto na czele: mój ojciec Karol Braun, który był wówczas prezesem Sokoła w Tarnowie. Uznano go widocznie za najgodniejszego zaszczytnej roli honorowego dowódcy i jako *Primus inter pares* prowadził na tym historycznym Złocie oddział kolegów – prezesów”<sup>13</sup>.

Henryka Braun była instruktorką harcerską, w roku 1921 została hufcową w Tarnowie, w latach 1925-1930 pełniła funkcję komendantki Krakowskiej Chorągwi Harcererek. Wydarzenia te były dla rodziny Braunów doświadczeniem niepowtarzalnym, historycznym i ogromnym dowodem patriotyzmu, czyli wartości dla nich najwyższej.

<sup>10</sup> Ojciec Konstanty Miller (1816-1886), muzyk, emigrant, wyjechał do Ameryki Północnej, następnie do Paryża, gdzie poznał i przyjaźnił się z Cyprianem Norwidem.

<sup>11</sup> W budynku, który zbudował Karol Braun, powstał Tarnowski Teatr im. Ludwika Solkiego. Kazimierz Braun wyreżyserował tam: *Europę* (2003) (tekst Jerzego Brauna), *Upadek kamiennego domu* Romana Brandstaettera (2006), *Celę Ojca Maksymiliana* (2011) (tekst Kazimierza Brauna).

<sup>12</sup> B r a u n, dz. cyt., s. 29.

<sup>13</sup> Tamże.

Świat artystyczny Braunów tworzony był w szczególności przez matkę, która przekazywała miłość do literatury, muzyki, malarstwa i teatru. Uczyla swoje dzieci czytać i dogłębnie interpretować utwory w kontekście historycznym, „w dziele danego autora [...] odczytywać nie tylko to, co on sam w nim umieścił, ale również to, co pracą wieków na tym urosło”<sup>14</sup>. Każdy akt czytania jest również spotkaniem z autorem i ponownym aktem twórczym. Poprzez refleksję nad tekstem następuje odkrywanie prawdy o sobie, o byciu człowiekiem, katolikiem, Polakiem. Od „poezji dziejów”<sup>15</sup> oczekuje się utrwalenia wielkich postaci i doniosłych momentów przeszłości, jak i przedstawienia „domowego, artystycznego życia”<sup>16</sup> przodków: „Sercem domu była ogromna biblioteka [...] zawierająca klasykę polską i światową, książki z dziedzin historii literatury, kultury greckiej i rzymskiej, a także ogromny zestaw atlasów. Wszystkie dzieci zaczytywały się w tych książkach, studiowały mapy”<sup>17</sup>.

Życie rodzinne Braunów opierało się na dwóch silnych indywidualnościach rodziców: „Stosunek do autorytetu ojca i matki był jasny i oczywisty: miłość, posłuszeństwo, harmonia i pogoda”<sup>18</sup>. Młodość, czyli lata 1914-1921 Juliusz Braun spędził w Tarnowie, tu uczęszczał do szkoły, II Gimnazjum o profilu klasycznym (z rozszerzonym programem nauczania łaciny i greki). Był to również czas rozwoju jego życia duchowego, kształtowała się jego religijność, odkrywał Boga i wsłuchiwał się w Jego głos. Zainspirowany postawą kuzyna, Antoniego Żulińskiego, aktora, kompozytora, codziennie uczęszczał do kościoła Matki Bożej Szkaplerznej na poranną Mszę Świętą i przyjmował Komunię Świętą. W roku 1915 Juliusz został harcerzem, wstąpił do Trzeciej Tarnowskiej Drużyny Skautów im. Michała Wołodyjowskiego, która działała w II Gimnazjum, do którego uczęszczał<sup>19</sup>. 5 listopada 1911 roku po zaprezentowaniu wykładu na temat skautingu przez Andrzeja Małkowskiego podjęto decyzję o utworzeniu Drużyny Skautowej im. Zawiszy Czarnego. 23 stycznia

<sup>14</sup> J.W. G o m u l i c k i, *Norwid o książce i czytaniu*, w: C. Norwid, *Myśli o książce i czytaniu*, wybór i wstęp J.W. Gomulicki, Miejska Biblioteka Publiczna im. L. Waryńskiego, Łódź 1983, s. 5.

<sup>15</sup> B r a u n, dz. cyt., s. 28.

<sup>16</sup> Tamże.

<sup>17</sup> Tamże.

<sup>18</sup> B r a u n, dz. cyt., s. 31.

<sup>19</sup> Ruch skautowy w Polsce był określany harcerstwem. Jego twórcą był angielski generał Robert Baden-Powell. Zgromadził młodych chłopców w roli zawiadowców w kolonialnej wojnie prowadzonej przez Brytyjczyków w Rodezji, podczas obrony twierdzy w roku 1899 w Mafeking. Zasady organizacji skautowskiej przedstawił w książce *Scouting for Boys*, wydanej w 1908 roku. W Polsce ruch ten rozwinął się w 1910 roku. Jego propagatorem był student Politechniki Lwowskiej Andrzej Małkowski. Zaczęto tworzyć drużyny harcerskie na fundamencie Sokoła i innych organizacji i stowarzyszeń niepodległościowych. Por. M. Ż y c h o w s k a, *Harcerstwo Tarnowskie 1910-1975*, Komenda Tarnowska Chorągwi ZHP im. Generała Józefa Bema – Muzeum Okręgowe w Tarnowie, Tarnów 1984, s. 23; por. też: B r a u n, dz. cyt., s. 44n.

1916 roku Juliusz Braun złożył przyrzeczenie harcerskie na Górze Marcina: „Mam szczerą wolę całym życiem pełnić służbę Bogu i Ojczyźnie, nieść chętną pomoc bliźniemu, być posłusznym prawu skautowemu”<sup>20</sup>.

Był bardzo zaangażowany w służbę harcerską. Uczęszczał na zbiórki, zdobywał kolejne stopnie harcerskie: „Zwiadowca” (1916), „Ćwik” (1917), „Harcercz Orli” (1918), „Harcercz Rzeczypospolitej” (1919). Był zastępowym, następnie, w roku 1918, został drużynowym, jako sekretarz został przydzielony do trzyosobowej Komendy Dzielnicowej Harcerstwa w Tarnowie. W lecie 1919 roku wypełniał obowiązki hufca. Harcerstwo kształtowało jego postawę patriotyczną, uczyło miłości do ojczyzny, rozwijało jego duchowość i religijność. Jak napisał Norwid: „Chrześcijańska moralność pozostała na stanowisku swoim pierwszym...”<sup>21</sup>. Od moralności człowieka zależy rozwój postaw obywatelskich i kształtowanie się patriotyzmu. Obowiązkiem każdego harcerza było uczestnictwo w niedzielnej Mszy Świętej. Najważniejszą zasadą była służba Bogu, Ojczyźnie i Bliźniemu. „Mistrz Słowa”, Cyprian Kamil Norwid, którego twórczość Braunowie bardzo cenili, napisał w eseju *Zmartwychwstanie historyczne*: „Duchem-narodu jest moralna-całość narodu. [...] Wojsko nawet [...], o ile moralnej jest całości w tej lubo na pozór najmaterialniejszej organizacji, [...] jest właściwą sobie proporcją cnoty jakiejś duchowej, np. posłuszeństwa”<sup>22</sup>. W życiu każdego człowieka moralność wyznacza kierunki postępowania w życiu społecznym. Rozwój duchowy, kształtowanie charakteru, sumień, często odbywa się w samotności, w ciszy. Dlatego Juliusz Braun dużo wędrował samotnie po górach, była to okazja do wsłuchiwania się w swój wewnętrzny głos, czas głębokich refleksji nad swoim „ja”. Dbał również o rozwój naukowy uczestniczył w konferencjach, sympozjach, współpracował z harcerskim pismem „Czuwaj”. W roku 1920 udoskonalili profil czasopisma i zmienił jego nazwę na „Światowid”. Pisał artykuły, recenzje teatralne ze spektakli, które odbywały się w teatrze w Tarnowie: „Wydarzeniem w moim życiu gimnazjalnym było przygotowanie przedstawienia *Wesela* Wyspiańskiego na scenie teatralnej w gmachu Sokoła. [mieści się tam obecnie tarnowski Teatr im. Ludwika Solskiego – B.T.] [...] Mnie przypadła rola Stańczyka, a także – w wyniku koniecznych przetasowań obsady – rola Jaśka. Rzemiosła aktorskiego nigdy więcej się nie miałem. Występując zaś w życiu w różnych rolach, zawsze się starałem być sobą. A nie odgrywać kogoś innego. Czy mi się to zawsze udawało? *Nescio, Deus scit.* (Nie wiem, Bóg wie)”<sup>23</sup>. W tym sa-

<sup>20</sup> B r a u n, dz. cyt., s. 46.

<sup>21</sup> C. N o r w i d, List do Augusta Cieszkowskiego, Paryż, 1 marca 1871, w: tenże, *Pisma wszystkie*, t. 9, *Listy*, oprac. J.W. Gomułicki, Państwowy Instytut Wydawniczy, Warszawa 1971, s. 478.

<sup>22</sup> T e n ż e, *Zmartwychwstanie historyczne*, w: tenże, *Pisma wszystkie*, t. 6, *Proza*, cz. 1, oprac. J.W. Gomułicki, Państwowy Instytut Wydawniczy, Warszawa 1971, s. 613.

<sup>23</sup> B r a u n, dz. cyt., s. 49n.



mym roku Juliusz wstąpił do organizowanej przez jego ojca Ochotniczej Legii Obywatelskiej w Nisku, po zwycięstwie w Bitwie Warszawskiej powrócił do szkoły, gimnazjum w Tarnowie.

Kolejnym ważnym etapem w życiu Juliusza Brauna były studia prawnicze na Uniwersytecie Jagiellońskim w Krakowie rozpoczęte w roku 1921 roku, najpierw był „słuchaczem nadzwyczajnym”, od 1922 – studentem zwyczajnym. Jesienią 1921 roku rozpoczął również studia w Akademii Handlowej, a w kolejnym roku akademickim w Szkole Nauk Politycznych. Równocześnie ze studiami rozpoczął pracę w Biurze Prezydium Miasta w Krakowie jako pomocnik kancelaryjny i administrator Biblioteki Pracowniczej magistratu. Od roku 1923 roku pracował w Pocztovej Kasie Oszczędności. Uczestniczył też aktywnie w działalności różnych stowarzyszeń, organizacji i klubów studenckich, takich jak Akademickie Koło Tarnowiaków, najpierw jako jego członek, a potem przewodniczący Komisji Rewizyjnej; był również członkiem Komisji Rewizyjnej „Odrodzenie”. Uczestniczył w spotkaniach Koła Uczniów Szkoły Nauk Politycznych i Stowarzyszeniu Słuchaczy Prawa. Był członkiem zarządu i sekretarzem Klubu Społecznego. Uczestniczył w wieczorach literackich, organizowanych przez Krąg Młodych Literatów i akademicki klub literacki „Helion”, napisał artykuł do pierwszego numeru „Gazety Literackiej”<sup>24</sup>. Priorytetami młodego Brauna były: wszechstronny rozwój, zaangażowanie oraz praca społeczna, naukowa i literacka. Był ważnym czytelnikiem, w jego artykułach można zauważyć rudymenty licznych lektur: od starożytnych klasyków, filozofów po współczesne prace naukowe z wielu dziedzin (na przykład prawa czy ekonomii). W roku 1931 został samodzielnym adwokatem, uzyskując pełne uprawnienia do wykonywania tego zawodu. Rozpoczął pracę w Izbie Przemysłowo-Handlowej w Sosnowcu. Jego wydana przez Izbę książka *Oszczędność przymusowa i częściowe ubezpieczenie jako wytyczne reformy systemu ubezpieczeń społecznych* została przyjęta jako rozprawa doktorska na wydziale Prawa Uniwersytetu Jagiellońskiego. W roku 1931 poślubił Elżbietę Szymanowską. Jej ojciec Michał Korwin Szymanowski (1877-1950) wykształcenie ekonomiczne zdobył w Belgii, pracował na Śląsku w Towarzystwie Francusko-Włoskim Dąbrowskich Kopalń Węgla. W roku 1904 ożenił się z Marią z Tymienieckich. Ich córka Elżbieta, nazywana Lili, która ukończyła później studia z zakresu ogrodnictwa na Uniwersytecie Jagiellońskim, przyszła na świat w 1906. Szymanowski w 1908 roku kupił ziemię w Mokrsku Dolnym, a w roku 1936 Juliusz Braun z żoną rozpoczęli budowę domu letniskowego.

W kolejnych latach, 1926-1931, Juliusz Braun pracował w Prokuraturze Generalnej w Katowicach i współredagował „Dziennik Pracy” w Sosnowcu

<sup>24</sup> Por. tamże, s. 69n.

(współpracował z literatem Leonem Kruczkowskim). W roku 1931 rozpoczął pracę jako adwokat w Izbie-Przemysłowo-Handlowej w Sosnowcu. Wyzna-  
czał sobie wysokie cele – 26 maja tego samego roku otrzymał stopień doktora  
nauk prawnych. Jak podkreślił Kazimierz Braun: „Stawał się coraz bardziej  
znany na forum ogólnokrajowym, a następnie i międzynarodowym. [...] Na  
Międzynarodowym Kongresie Oszczędnościowym w Paryżu w 1935 roku  
ojciec miał referat (po francusku) o kontrolowanym systemie oszczędności-  
wym, którego był zdecydowanym zwolennikiem”<sup>25</sup>. W roku 1937 brał udział  
w ogólnopolskiej konferencji Izb Przemysłowo-Handlowych w Berlinie.

Kolejne lata to wybuch drugiej wojny światowej, czas najtrudniejszy,  
kiedy jeszcze z większą uwagą należało wsłuchać się w głos sumienia, być  
aktywnym w służbie narodu i nieść pomoc drugiemu człowiekowi. Jak napisał  
Norwid: „A każde prawe serce polskie jest jednym pulsu uderzeniem tej zbio-  
rowej osoby. Głosem tego Narodu jest harmonia ojczysta, mieczem – jedność  
i zgoda, celem – prawda. [...] Bo Ojczyzna [...] jest to moralne zjednoczenie”<sup>26</sup>.  
Czytelnik odkrywa twarz Juliusza Brauna jako gospodarza w Milechowie,  
oblicze nauczyciela i wychowawcy, który uczył swoje dzieci patriotyzmu i mi-  
łości do Ojczyzny wskazując na postaci wielkich Polaków, którzy potrafili  
łączyć miłość do Ojczyzny z siłą, odwagą i moralnością. Okres wojenny to dla  
Brauna czas działań konspiracyjnych. Poznajemy też jego twarz skupioną na  
modlitwie i kontemplacji Boga: „Kształtem miłości piękno jest; i tyle / Ile ją  
człowiek ogląda na świecie, /W ogromnym Bogu albo w sobie – pyle”<sup>27</sup>.

Odkrywamy oblicze Juliusza Brauna jako działacza niepodległościowego,  
współorganizatora „Unii”. W roku 1940, podczas tajnego zebrania w Warsza-  
wie, przywódcy „Nowej Polski” (Jerzy Braun), „Warszawianki” (Jan Hoppe)  
i „Grunwaldu” (Miroslaw Leśkiewicz) powołali do życia organizację o nazwie  
„Unia”. Przewodniczącym jej okręgu kieleckiego „Unii” został Juliusz Braun,  
który był również członkiem jej Rady Naczelnej. Opracowano wówczas dwa  
dokumenty: *Deklaracja ideowa „Unii”* i *Unionizm – podstawowe zasady dok-  
tryny*. Zawierały one koncepcję ustroju Polski po odzyskaniu niepodległości.  
W drugim wymienionym dokumencie pojawia się kluczowe stwierdzenie:  
„Nowy ustrój demokratyczny i nowe prawo międzynarodowe musi zapewnić  
autorytet obiektywnych, powszechnie obowiązujących norm moralnych”<sup>28</sup>.  
Szeroka i wieloaspektowa działalność kulturalna „Unii” zainicjowała powo-

<sup>25</sup> Tamże, s. 97.

<sup>26</sup> C. N o r w i d, *Przemówienie wygłoszone na jubileusz powstania listopadowego w Brukseli w 1846 roku*, w: tenże, *Pisma wszystkie*, t. 7, *Proza*, cz. 2, oprac. J.W. Gomulicki, Państwowy Instytut Wydawniczy, Warszawa 1973, s. 7n.

<sup>27</sup> T e n ż e, *Promethidion. Rzecz w dwóch dialogach z epilogiem*, Drukarnia L. Martinet, Pa-  
ryż 1851, s. 15.

<sup>28</sup> J. B r a u n, *Zarys historii i doktryny „Unii”*, „Chrześcijanin w świecie” 4(1985) nr 139, s. 32.

łanie Instytutu Europy Środkowej i Wszechnicy Narodowej, propagowała publikacje książek i prasy (na przykład „Kultury Jutra” redagowanej przez Jerzego Brauna). Sprawowała także opiekę na tajnym Teatrem Rapsodycznym w Krakowie, który 1 listopada 1941 roku wystawił *Króla Duchą* Juliusza Słowackiego. Reżyserem spektaklu i adaptatorem tekstów był Mieczysław Kotlarczyk. Warto przypomnieć, że w pierwszym przedstawieniu zagrał młody Karol Wojtyła.

„Unia” posiadała swoje godło, herb w postaci litery „U”, w który wpisany był krzyż. Słowa hymnu, do którego muzykę skomponował Stanisław Bukowski, napisane zostały przez Juliusza Brauna:

Unia swą mocą zmieni  
Cały ogromny świat.  
Prawdę wynurzy z cieni,  
Że bliźni to nasz brat.  
Ona połączy ludy  
W jeden braterski krąg.  
Ludzkość dostąpi cudu  
W jedności serc i rąk.  
Bo Unia wielką jest potęgą,  
Z miłości czerpie swoją moc.  
Po przebudować świata sięga,  
By w jasny dzień przemienić noc<sup>29</sup>.

W 1943 roku „Unia” połączyła się ze Stronnictwem Pracy i weszła do podziemnego parlamentu, który rok później został przekształcony w Radę Jedności Narodowej. Do organizacji „Unia” należały oddziały wojskowe, które w roku 1942 stały się częścią struktur Armii Krajowej.

Lata 1945-1948 dla Juliusza Brauna to czas intensywnej pracy w Izbie Przemysłowo-Handlowej w Częstochowie (powołał jej filie w Kielcach i Radomiu), gdzie wydawał on „Informator Gospodarczy Okręgu Izby Przemysłowo-Handlowej”. W roku 1946 zorganizował Wystawę Gospodarczą, uroczystego otwarcia której dokonał 7 września Józef Ozga-Michalski, Przewodniczący Wojewódzkiej Rady Narodowej w Kielcach. Przybyli na nią również Kardynał Adam Stefan Sapiecha, Kardynał Prymas August Hlond, nowy wyświęcony na biskupa ksiądz Stefan Wyszyński.

9 października 1945 roku Braun otrzymał pozwolenie z Ministerstwa Oświaty na powołanie Wyższej Szkoły Administracyjno-Handlowej w Częstochowie. Najtrudniejszy okres w życiu Juliusza Brauna i jego rodziny był to lata 1948-1953, kiedy przebywał w więzieniu w Częstochowie, Kielcach

<sup>29</sup> B r a u n, dz. cyt., s. 171.

i Warszawie za działalność w „Unii”. Pogarszający się stan zdrowia spowodował, że trafił do więziennego szpitala, który opuścił 8 listopada 1953 roku.

W 1959 roku podjął pracę jako radca prawny w biurze „Miastoprojekt” w Częstochowie, zaproponował projekt rozbudowy miasta. Podjął ponownie pracę na stanowisku adwokata, następnie został zaangażowany jako adiunkt w Zakładzie Badań Naukowych Górnośląskiego Okręgu PAN w Zabrze. Przygotował rozprawę zatytułowaną „Elementy ekologii miasta przemysłowego”. Zgłosił ją jako rozprawę habilitacyjną w Instytucie Geografii PAN. 30 czerwca 1963 roku otrzymał kolejny stopień naukowy – habilitację. Następnie został mianowany docentem w Polskiej Akademii Nauk. W 1973 roku uzyskał tytuł profesora nadzwyczajnego. Jego głównym miejscem pracy była Świętokrzyska Stacja Naukowa PAN w Kielcach.

Ważnym wydarzeniem w jego życiu okazało się spotkanie (w roku 1979) z papieżem Janem Pawłem II: „Przechowywana jak relikwia w moim archiwum fotografia ukazuje Papieża obejmującego ojca za ramię – jakimś ojcowskim, pełnym miłości gestem”<sup>30</sup>. Juliusz Braun zmarł 31 maja 1990 roku.

Kazimierz Braun w książce *Mój Ojciec* zaprasza czytelnika również do zapoznania się ze wspomnieniami swojej siostry Marii Braun-Gałkowskiej. Przytoczę myśl umieszczoną w zakończeniu jej rozważań: „Od dziecka [ojciec – B.T.] miał szczerą wolę pełnić służbę Bogu i Ojczyźnie i nieść chętną pomoc bliźniemu – i realizował to przez całe życie”<sup>31</sup>. Pamięć to most łączący świat żywych, z tymi, którzy odeszli, ale ciągle żyją wśród nas. Są światłem, które rozświetla mroki życia, są miłością, gdy jej brak, są radością, gdy panuje smutek. Żyją wciąż w naszej pamięci, a przede wszystkim w naszych sercach.

Pamięć to wartość wyjątkowa, czynnik sensotwórczy, pozwalający odbyć podróż w przeszłość, do źródła tożsamości, fundamentów wartości, zasad, na których opiera się świat Prawdy, Dobra i Piękną.

Książka *Mój Ojciec* Kazimierza Brauna odkrywa przed czytelnikiem świat rodzinnych wspomnień, pięknych postaw patriotycznych, wzorów do naśladowania i ideałów, według których należy żyć. Czytelniku, auscultate et perpende.

#### BIBLIOGRAFIA / BIBLIOGRAPHY

Assmann, Jan. *Pamięć kulturowa: Pismo, zapamiętywanie i polityczna tożsamość w cywilizacjach starożytnych*. Translated by Anna Kryczyńska-Pham. Warszawa: Wydawnictwo UW, 2008.

Św. Augustyn. *Wyznania*. Translated by Zygmunt Kubiak. Kraków: Znak 2007.

<sup>30</sup> B r a u n, dz. cyt., s. 171.

<sup>31</sup> M. B r a u n - G a ł k o w s k a, *O moim Ojcu*, w: K. Braun, dz. cyt., s. 319.

- Braun, Juliusz. “Zarys historii i doktryny ‘Unii.’” *Chrześcijanin w Świecie* 4, no. 139 (1985): 27–37.
- Braun, Kazimierz. *Mój Ojciec*. Toruń: Wydawnictwo Adam Marszałek, 2023.
- Braun-Gałkowska, Maria. *O moim Ojcu*. In Kazimierz Braun, *Mój Ojciec*. Toruń: Wydawnictwo Adam Marszałek, 2023.
- Gomulicki, Juliusz W. *Norwid o książce i czytaniu*. In C. Norwid, *Myśli o książce i czytaniu*. Edited by Juliusz W. Gomulicki. Łódź: Miejska Biblioteka Publiczna im. L. Waryńskiego, 1983.
- Mac Cumhaill, Clare. “Still Life, a Mirror: Phasic Memory and Re-encounters with Artworks.” *Review of Philosophy and Psychology* 11, no. 2 (2020): 423–46.
- Nora, Pierre. “Czas pamięci.” Translated by Wiktor Dłuski. *Res Publica Nova*, no. 7 (2001): 37–43.
- Norwid, Cyprian Kamil. “Niewola: Rapsod.” In Norwid, *Pisma wszystkie*. Vol. 3. *Poematy*. Edited by Juliusz W. Gomulicki. Warszawa: Państwowy Instytut Wydawniczy, 1971.
- . Norwid to August Cieszkowski, 1871. In Norwid, *Pisma wszystkie*. Vol. 9. *Listy*. Edited by Juliusz W. Gomulicki. Warszawa: Państwowy Instytut Wydawniczy, 1971.
- . *Promethidion: Rzecz w dwóch dialogach z epilogiem*. Drukarnia L. Martinet, Paryż 1851.
- . “Przemówienie wygłoszone na jubileusz powstania listopadowego w Brukseli w 1846 roku.” In Norwid, *Pisma wszystkie*. Vol. 7. *Proza*. Part 2. Edited by Juliusz W. Gomulicki. Warszawa: Państwowy Instytut Wydawniczy, 1973.
- . “Zmartwychwstanie historyczne.” In Norwid, *Pisma wszystkie*. Vol. 6. *Proza*. Part 1. Edited by Juliusz W. Gomulicki. Państwowy Instytut Wydawniczy, Warszawa 1971.
- Trygar, Barbara. “Ja” w drodze do Wolności, Prawdy i Piękna: *Aspekty aksjologiczne twórczości prozatorskiej Kazimierza Brauna*. Rzeszów: Wydawnictwo Uniwersytetu Rzeszowskiego, 2019.
- Żychowska, Maria. *Harcerstwo Tarnowskie 1910-1975*. Tarnów: Komenda Tarnowska Chorągwi ZHP im. Generała Józefa Bema – Muzeum Okręgowe w Tarnowie, 1984.

## ABSTRAKT / ABSTRACT

Barbara TRYGAR – „Wielka jest siła pamięci”... *Mój Ojciec* Kazimierza Brauna

Celem artykułu jest omówienie wspomnień Kazimierza Brauna zatytułowanych *Mój Ojciec*. Autor przedstawia dzieciństwo, młodość, studia, pracę zawodową, rodzinę, okres wojny, działalność naukową Juliusza Brauna wybitnego prawnika, ekonomistę, uczonego. Bardzo ważnym źródłem informacji oprócz osobistych wspomnień pisarza było archiwum Juliusza Brauna, dokumenty, notatki, artykuły. Wspomnienia są uporządkowane tematycznie, co pozwala

odkryć drogę kształtowania się jego tożsamości. Najważniejszą kategorią jest pamięć, która stanowi fundament mądrości, poznania prawdy i rozumienia dobra i piękna, oraz inspiruje do działania.

Słowa kluczowe: pamięć, historia, tradycja, patriotyzm, wartości

Kontakt: Instytut Jana Pawła II, Wydział Filozofii, Katolicki Uniwersytet Lubelski Jana Pawła II, Al. Raławickie 14, 20-950 Lublin

E-mail: [ethos@kul.pl](mailto:ethos@kul.pl)

Tel. 81 4453218

Barbara TRYGAR, “Great is this force of memory”... *Mój Ojciec* [“My Father”] by Kazimierz Braun

The aim of the article is to discuss Kazimierz Braun’s memoirs *My Father*. The author presents the childhood, youth, education, professional work, family, the war period, and the scholarly activity of Juliusz Braun, an outstanding lawyer, economist, and scientist. A very important source of information, apart from the author’s personal memories, was Juliusz Braun’s archive, documents, notes, and articles. The memories are arranged thematically, which allows the reader to follow the path of shaping Juliusz Braun’s identity. The most important category is that of memory, which is the foundation of wisdom, of the knowledge of truth, as well as of the understanding of goodness and beauty, and provides inspiration for action.

Keywords: memory, history, tradition, patriotism, values

Contact: John Paul II Institute, John Paul II Catholic University of Lublin, Al. Raławickie 14, 20-950 Lublin, Poland

E-mail: [ethos@kul.pl](mailto:ethos@kul.pl)

Phone: +48 81 4453218

Łukasz MŚCISŁAWSKI

## WHEN WISDOM AND BIGTECH-AI COLLIDE

Where is the wisdom we have lost in knowledge?<sup>1</sup>

Thomas Stearns Eliot

In the age of information, it seems that we would be better off with more wisdom and a little less information. Switching over our smartphone and tablet to Plato and Epictetus, sometimes, may be a good start.<sup>2</sup>

Edward Spence

The subject of artificial intelligence (AI) provokes a wide variety of reactions—from extreme enthusiasm for the possibilities of this technology in the daily press to Cassandra—like visions almost straight out of the film *Terminator*. The fact that AI is becoming more and more present in new areas of human functioning provokes discussions on both the potential benefits and possible dangers.

In this context, *Stoic Philosophy and the Control Problem of AI Technology: Caught in the Web* by Edward H. Spence is a position worthy of attention.

It must be stressed that the title is somewhat misleading. Spence addresses a much broader range of issues than dealing with a philosophical analysis of narrow, technological issues related to AI, while ultima-

tely focusing indeed on issues related to this technology. Place them, however, in the much broader perspective of a specific view of human beings and how they function, in contact with technology in general and AI technology in particular.

Spence considers two most pressing issues affecting human existence: (1) climate change and (2) the rapid technological advances in the field of autonomous AI agents for information processing (see p. 1).<sup>3</sup> The author attempts to unravel the latter issue in a rather surprising way by asking what technology is good for and trying to explore the impact of technology on human well-being, especially when it comes to information processing technologies (including AI). Thus, it can be said that the starting point is not technology, but the human being affected by that technology. Human beings—and therefore, as one might suspect, thinking beings. It is precisely the defence of human thinking

<sup>1</sup> Thomas Stearns Eliot, *The Rock* (New York: Harcourt, Brace, and Company, 1934), 7.

<sup>2</sup> Edward Spence, *Stoic Philosophy and the Control Problem of AI Technology: Caught in the Web* (Lanham: Rowman & Littlefield, 2021), 93.

<sup>3</sup> The page numbers in parentheses refer to the discussed work by Spence.



and functioning, and the defence of human dignity, that the author is concerned with. All of this takes place against the backdrop of the question of who has power over information and the technologies of producing and disseminating information, above all AI, used by the giants of the IT market, referred to as BigTech.<sup>4</sup> When it comes to the immersion of modern man in the infosphere, the problem, as the author notes, is not so much the use of various types of information processing technologies themselves, but how, by whom and for what purpose they are used. By virtue of the amount of information and the speed and range of its propagation, the author seems to conclude that the problem is not one of an individual or local community, but a global one. Taking also into account that it is not a purely technical problem, he suggests that this issue has also a moral aspect and should be solved collectively by the entire global community<sup>5</sup> (see p. 59f.). The question that arises here is whether there are not already some means in the history of human thought that would make it easier, or, in fact, possible, to analyse such issues from an ethical perspective? After all, human nature has not fundamentally changed in the last hundreds or thousands of years. Does it automatically follow from the fact that AI is a contemporary technology that such means have yet to be developed? And if they are already available, could they be modified and applied to the issues mentioned? Does the rapid development of technology mean that it is only possible to analyse the situation ex

post, which, due to the dynamics of this development, results in perpetual lagging behind? Or, on the other hand, is it possible to develop tools of reflection that anticipate and modify the way advanced research on modern technologies is designed and, at the same time, make it possible to react in advance to the undesirable effects of the use of these technologies, especially in the area of information processing in the broadest sense? Edward H. Spence's proposal seems to be an interesting attempt to find answers to such questions.

A brief reconstruction of Spence's proposal is as follows. Because of the increasing accumulation of information and the increasing impact of technology on human beings and their well-being some kind of theoretical tool is needed to assess the impact of technology on human quality of life. Such a theory should be able to determine whether and how technology can assist humans in enhancing their quality of life (see p. 1). It is important to emphasise that Spence's posing the question in this way opens the way to the question of whether any technology is able to contribute to the well-being, ethical quality of a person's life or their quest for fulfilment. It is a question that, it seems, is too rarely asked in the context of contemporary technological developments.

The problem, of course, is to answer the question of what a good life is. Spence goes in favour of the eudaimonic direction here, stating that a good life is one that at least minimally allows a person to achieve self-fulfilment, well-being, or simply happiness (see p. 15). A theory that could be used for the aforementioned evaluation must meet the relevant formal and material conditions (see p. 15f).

The pursuit of the good life—or, ultimately, happiness—when confronted with an ocean of information is not so simple. Spence argues that there is no simple transition between taking in information and li-

<sup>4</sup> This usually includes technology giants such as Amazon, Apple, Facebook, Google, and Microsoft.

<sup>5</sup> Here is one of the weaker points of Spence's work. For it is difficult to consider that there is any homogeneous global community. Some of the inhabitants of our planet may not be interested in information processing problems at all.

ving well. Hence, a tool is needed to assess whether a given piece of information (or technology) will contribute to the good life. Such a tool, for him, is wisdom (see p. 18). He sees its role in multiple ways. Wisdom is thus a certain meta-knowledge, a principle of evaluation (*ibidem*), but it can also be seen as a meta-technology of the self (see p. 20).

In the context of the search for an appropriate criterion for evaluating technology, including communication and information processing together with AI, Spence presents the views of the Stoics and Neo-Stoics. In doing so, he emphasises that, according to the Stoics, a virtuous life and wisdom are necessary and sufficient conditions for happiness. He also juxtaposes the views of the Stoics and neo-stoics with those of Gewirth as a representative of the rationalist movement in the context of the key points: (a) the ultimate goal; (b) a good life in harmony with nature, (c) happiness, (d) the relationship between virtue and eudaimonia. In doing so, he notes that the latter's views coincide with those of the Neo-Stoicists. According to Spence, particularly significant here is the similarity between the concept of self-fulfilment and the Stoic concept of eudaimonia. In the case of (d), the juxtaposition of Gewirth and the Stoics is very interesting, as both are concerned to show the role of the conformity of life to the requirements of reason. Here, however, Spence emphasises the role of the principle of general consistency in Gewirth, which is to guarantee the achievement of the best possible version of oneself in conformity with human rights and leads to a view of man as someone whose nature is to act rationally, purposefully (see p. 39f.). Another important point for Spence is that, in order to avoid frustration or disappointment, the Stoics propose doing two things: controlling those things that are within our control, and remaining unmoved by the things we cannot control

(see p. 31). This explains, at least in part, why Spence is interested in the views of the Stoics, the Neo-Stoics and the views of Gewirth, which he believes to be strongly aligned with these two currents. This is because the fundamental issue at the heart of the author's interest is the problem of control, especially in the context of modern technology and its impact on humans (see p. 60). Here, Spence highlights the fact that by treating philosophy as a kind of therapy of the mind, they have noticed that there are things we can control and things we cannot control. In order to be able to achieve happiness, he proposes, in relation to technology, the application of the basic Stoic principles of the good life: (a) dealing only with things over which we have control; (b) considering the ultimate goal of life—that is, achieving happiness; (c) living a virtuous life and living wisely; (d) cosmopolitanism; (e) living in harmony with nature; (f) growing in virtue and internalising it, (g) treating philosophy as a way of life (see p. 62–69). As Spence points out, in the digital age, Stoic philosophy provides a systematically unified practical proposal for living fruitfully and achieving happiness. This is made possible by epistemic, ethical and eudaimonic principles centred on the concept of wisdom. It also provides the means, both theoretical and practical, needed to address the problem of control in the context of contemporary technology, especially in relation to information processing and AI (see p. 70). Spence devotes the first four chapters of his book (“Introduction: Who Is in Control?”, “What Is Technology Good For”, “Stoic and Neo-Stoic Philosophy”, and “Application of Stoic Philosophy to Technology”) to outlining the problems and preparing the relevant tools, based mainly on the two philosophical currents mentioned above.

In chapter five (“Wisdom and Well-Being: The Dual Obligation Information-Wisdom Theory”), Spence offers the reader

a particularly useful tool—the Dual Obligation Information-Wisdom Theory (DO-IT-Wisdom Model). It allows issues related to technology to be framed and analysed in two categories: epistemic and ethical. The author devotes the fifth chapter to this. Using previously developed tools provided by the Stoics and Neo-Stoics, he shows wisdom as a kind of meta-information and meta-knowledge, which include epistemic, ethical, and eudaimonic dimensions, and also provides the possibility to capture the link between information, knowledge, good life, and well-being. In general, it allows to combine into a conceptual scheme to find a direct link between information and communication technologies and AI technologies and well-being and good life. The proposed concept of freedom also provides normative criteria for assessing the dissemination and impact of information (see p. 74f.).

Central to this is the DOIT-Wisdom Model already mentioned (see p. 90f.). Within it, Spence proposes to link the concepts of information and wisdom with values into a coherent whole. The context of treating information as a message allows him to link issues of information processing to the values of truthfulness, reliability, and robustness. The context of treating the communication of information as a kind of action, he links to the preservation of the right to freedom and well-being. Treating wisdom as a kind of meta-informational action, he links it to epistemic, ethical and eudaimonic values. In doing so, it is noteworthy that he proposes a shift from epistemic to eudaimonic values (see p. 21) when dealing with issues related to the impact of technology on humans and the way they function. The whole is motivated by Spence's reference to three aspects. The metaphysical (or ontological) aspect is related to the fact that nowadays man lives not only in the biosphere, but also in the infosphere, and that informa-

tion and its processing increasingly permeate his action (see p. 91). He refers to the second aspect as technological: the digitalisation of information allows for the instantaneous creation and dissemination of information. This can have a positive impact, but also a negative one, e.g., due to the unreflective use of this possibility. The latter possibility instantly results in an impact on a person's existential situation, which Spence illustrates with numerous examples. This suggests the necessity for a culture to learn how to handle information, hence also his encouragement to use tools that are already present in human culture—e.g., his proposed use of Stoic and Neo-Stoic philosophy (see p. 93). A third aspect that Spence proposes to consider is that prudence and foresight are needed in dealing with information and the means of processing it (see p. 94). The entire chapter is interspersed with insights into the role of major corporations in information processing and uptake. Spence points out that, in fact, information or content that reaches us is very precisely selected with the tools used by the digital giants. The problem is that these tools operate in a non-transparent way, and it is not clear who could be held accountable for their actions. In places, his observations may seem too intrusive to the reader, but it is nevertheless difficult to refuse to be right (see p. 95).

Chapter Six (“Tech Media Corruption in the Age of Information”) is devoted to discussing the irregularities of the information processing giants. The starting point here is to present the implicit part of their business model, to compare them with the operation of traditional media, and to point out the places where the interests of users of information processing platforms and their owners are at odds with each other and the radical, albeit implicit, violation of the social order. These include, in particular, violations of the rights to: informed consent, privacy and autono-

my (see p. 111). The basic tool that the author uses is the DOIT-Wisdom model proposed earlier, supported by the identification of the basic characteristics of corruption: the possession of power, the disposition to exercise that power, the opportunity to use it, a kind of invisibility of actions or their concealment, consideration only of one's own gain, violation of the duty of the fiduciary—in this case, the data, provided by the users (see p. 112f.). Spence's analysis of the case of Facebook's involvement in the Cambridge Analytica affair leads the author to conclude that the business model of surveillance used has resulted in a corruption of communication and information, a corruption that violates democratic principles and strikes at society (see p. 117). In this context, he asks a very important question: what kind of society are such models of so-called Big-Tech functioning creating? He also stresses that the real problem is precisely the methods and practices of non-intrusive, non-transparent, unjust, and unaccountable (or accountable) collection and use of user data (see p. 120). An additional difficulty here is the invisibility of these practices (such as the transfer of user data to advertising companies or the preparation of information) to the user. Hence, the question is whether humans can be happy in a world ruled by Technological Oligarchs? Moreover, this raises the question of who actually controls our autonomous intentional action, which is a necessary condition for well-being (see p. 121). The problem is acute because factual knowledge of the world is part of the information on which we make decisions. If this information is appropriately and invisibly crafted, then we find ourselves in a situation where we are not even in control of our decision-making process. On the basis of Spence's reflections, another question can be posed: will we then have power over our own interiority, our own intellectual sphere, and

to what extent? However, Spence suggests that a collective multidisciplinary approach, involving the community and the relationships within it, could be a solution of the problem.

The next two chapters ("Normative Impact of ICT Technologies on Well-Being" and "The Normative Impact of AI Technologies on Well-Being") deal with this issue. In chapter seven, the author places particular emphasis on autonomy, integrity, and dignity, as seen in the context of achieving self-fulfillment or the pursuit of happiness. In practical terms, this means the well-understood self-control characteristic of the Stoics (see p. 127). This self-control is in some danger due to BigTech's actions, but it can be sustained at least at the level of our information generation. Therefore, one of the measures that can prevent an undesirable situation is paying attention to the dignity of the person or—alternatively—according to Spence—self-respect (see p. 129–33). The combination with Gewirth's principle of general consistency leads to very interesting conclusions, namely that the basic category to which attention now needs to be drawn is the category of human dignity, closely related to the dignity of a specific person. It is this category, in relation to issues related to the control of information, that Spence proposes to take special seriousness (see p. 137–46). Moreover, the previously introduced principles of Stoic philosophy and looking at the human situation from this very perspective give us the basis to see in the light of reason this inviolable dignity of people who are not alone—it is united by something that Gewirth defines as a community of rights (see p. 136).

The previously prepared tools allow the author in chapter eight to face the problem of controlling or influencing an individual or society, also on a global scale. This issue is discussed in the context of the impact of information processing technology and AI,

in particular the use of two techniques: machine learning and neural networks. With regard to the latter, Spence draws attention in particular to the black box problem—a situation in which it is not known exactly on what basis a given technology obtains its output data.<sup>6</sup> Such a situation means that the lack of transparency of the functioning of such technologies results in the impossibility of assigning responsibility for the decisions made. This, in turn, leads to a situation where it is impossible to monitor AI-based systems in order to eliminate the results of their actions potentially harmful to society (see p. 149).

In this context, Spence emphasizes (see p. 150) that the monopolistic practices of the IT industry giants undermine the individual power of individuals over their own will. People's sovereignty is threatened. Referring to the words of Luciano Floridi, the author emphasizes the need to rethink the very concept of sovereignty (see p. 151). In a very clear way, this ultimately leads to the issue of protecting our individual sovereignty, especially in the age of global information processing and dissemination. This is an urgent problem as individual sovereignty is one of our natural rights, essential to our freedom, well-being and dignity as persons.

A major problem with the lack of transparency with regard to machine learning is that: (1) there is some degree of loss of control over the activities of technology companies by their human workers; (2) through invasive persuasive algorithms, operating in an invisible way, Big-Tech companies have an impact on decision-making by users (see p. 153). Spence lists a whole list of dangers of such a situ-

ation, partly resulting from the very nature of the operation of this type of technology: inconclusive evidence, inscrutable evidence, misguided evidence, unfair outcomes,<sup>7</sup> transformative effects (see p. 153–62) traceability.

A separate problem of key importance that appears in the context of functioning in cyberspace is a kind of disintegration of identity, and more specifically of our identity data. Information processing systems, based on the data held, create a certain representation of the user, and a group digital identity over which these systems have control. Usually, it is not complete, and therefore it is in contradiction with our natural identity, which, unfortunately, also becomes dangerous under the influence of persuasive BigTech techniques. This causes a fundamental paradox: the autonomy of our natural identity is threatened by persuasive technologies that are callously used to shape the digital group identity (see p. 161f.).

In this context, Spence emphasizes the importance of the presented model of neo-stoic philosophy, based on the rationalist ethics of laws. This is because it enables us to identify and assess the violation of our autonomy (see p. 164). He also emphasizes that with the help of modern technologies, BigTech companies systematically undermine and violate our ethical and cognitive rights, as well as the common good and well-being—both for individuals and for society (see p. 165). Another very important issue arises here, introduced by, so to speak, immoderate enthusiasm for modern information processing technologies. This is the trend of trusting the results of programs more than individuals or the community. This leads to an increasing dehumanization of both individuals and

<sup>6</sup> This is a known problem in which AI algorithms can report the outcome of some decision-making process, but neither the systems nor their designers can justify this and not the other outcome.

<sup>7</sup> Outcomes of work of given neural network is strongly dependent on way in which this network is trained. Here appears another problem: who and how is training neural networks.



communities (see p. 165f.). It should also be emphasized here, which, unfortunately, Spence does not do, that unlimited trust in the results presented by IT systems is at best imprudence and getting rid of the habit of critical thinking.<sup>8</sup> In the whole of this puzzle, Spence emphasizes the fact that—based on the social commitment of the Stoic philosophy—greater social control of the functioning of BigTech companies is needed<sup>9</sup> (see p. 166–73).

Spence's very valuable contribution is drawing attention to the fact that in the era of mass digital information processing, the use of the concept of the dignity of a person in the context of functioning in the infosphere should be restored. He emphasizes that AI algorithms are based on statistical inference. Meanwhile, statistical correlations are not yet causal relationships. Hence, among other things, there is a conflict between our digital representation created by AI algorithms and our real identity. In this context, Spence suggests referring to the concept of the dignity of the person in order to provide us with certain digital rights that could not be appropriated by BigTech companies,<sup>10</sup> and the right to identity. The last postulate can be realized if the algorithms of artificial

intelligence, especially machine learning, are trained taking into account the norm of preserving the dignity of a person (see p. 173–85). It seems that another difficulty arises here: clarifying how the dignity of a person should be understood and how to implement their norms in neural network training procedures.

In chapter nine (“Smart Machines and Wise Guys”) Spence is outlining a certain future perspective, a kind of *modus procedendi* for the future. It begins with a short summary of the considerations so far (see p. 189–92). Next, Spence points to the predatory policy of BigTech, which tries to change our preferences and the basis of autonomous choices (see p. 192f.). This policy is based on the use of AI technology. So the question is, what will happen if super-AI technologies are developed? Interestingly, Spence shows a certain amount of optimism here and seems to have a somewhat similar approach to the matter here as Wooldridge.<sup>11</sup> He notes that we do not need to be particularly afraid of this, as long as we act prudently. More precisely, instead of being afraid of what AI will do, we need to teach AI what our goals are, learn to include and keep our goals in the solutions it searches for (see p. 193–97). However, the problem may arise at the meta-level: how to take into account the dignity of the person already when designing AI systems and their implementation, so that two classes of people are not formed: an elite, operated by people, and a lower class, managed by machines, what is the real threat (see p. 197f.)? From a technological point of view, a value-sensitive design (VSD) approach can be considered here (see p. 217). However, on the human side, perhaps our expectations and goals need to be changed. While the influence

<sup>8</sup> Suffice it to say that the program itself may work fine, but the basic ideas behind creating an algorithm, assumptions about what it should do and how, and the interpretation of its results, may be wrong. Interestingly, Spence points out to this, using his experience as an accountant and auditor, that it is important not only that the results match, but also that the methodology used by the company is correct (see p. 168–69).

<sup>9</sup> It can be seen as a kind of irony that similar statements can be found in the statements of BigTech representatives, such as Mark Zuckerberg, cited by Spence (see p. 166f.).

<sup>10</sup> Hence, some international regulations, such as GDPR-EU.

<sup>11</sup> Michael Wooldridge, *The Road to Conscious Machines: The Story of AI* (Dublin: Pelican Books, 2021).

of machines is essentially external to us, as Spence points out, we do have power over our inner self. Therefore, we need, in line with the indications of Stoic philosophy, to learn internal control, including the already mentioned modification of our expectations and goals (see p. 203). This in turn leads to the suggestion of a symbiotic relationship between us and machines (see p. 214) and thinking in terms of designing machines that accomplish our goals and defer to man (see p. 215).<sup>12</sup> Such an approach to our attitude, however, requires wisdom, able to comprehend our goals, expectations, laws and knowledge about the reality around us into one coherent whole. Due to the global scale of the impact of AI-based technologies used by BigTech, it would require some kind of cultural movement, also of a global nature, promoting the ethos of life based on wisdom—perhaps also based on the foundations of stoic thought (see p. 208).

Unexpectedly, Spence emphasizes the role of the concept of love in this kind of ethos and the relationship between people and machines,<sup>13</sup> by analogy to the literal understanding of the term philosophy, but also by emphasizing its importance in

various religious traditions, including the Christian vision of the relationship between God and man (see p. 224).

Edward Spence's book is undoubtedly unique. The author skillfully tries to apply the tools available in philosophical culture since antiquity to solve one of the most pressing problems of contemporary times, related to the development of AI technology and its almost ubiquitous presence. The solidity of building the arguments and the clarity of the whole should be emphasized. Undoubtedly, the enrichment of the whole is also the reference to abundant literature. However, the question arises whether, with all the solidity of the argument, it can be persuaded to change the way of functioning and thinking of a person who is accustomed to the current situation and who is simply comfortable in this situation? It seems that this sense of comfort, or simply the modern pace of life, can very effectively prevent both reflection on the problems related to the increasing interpenetration of human life and modern technology, as well as changing this situation to one more compatible with human nature.

It seems that for anyone who is familiar with the ethical side of the development of modern technologies and some sort of ordering and taming this extremely complex issue, Spence's book should be one of the basic readings.

Contact: Faculty of Management, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland  
E-mail: lukasz.mscislowski@pwr.edu.pl

<sup>12</sup> It is worth noting that similar conclusions were presented by Paweł Polak and Roman Krzanowski (see Paweł Polak and Roman Krzanowski, "How to Tame Artificial Intelligence (AI)? A Symbiotic AI Model for Beneficial AI," *Ethos* 36, no. 3 (143) (2023): 92–106).

<sup>13</sup> It also raises here the intriguing question, can AI, or super-AI, be taught some kind of love for humans?



Damian SZCZĘCH

## PLAN NA ŻYCIE W ŚWIECIE ZDOMINOWANYM PRZEZ SZTUCZNĄ INTELIGENCJĘ

Catriona Campbell jest psychologiem behawioralnym, przedsiębiorcą i doradcą biznesowym. Specjalizuje się w design thinking oraz tematach związanych z interakcją człowiek–komputer. Recenzowany tekst jest jej pierwszą książką. Cel autorki stanowi przedstawienie stosowanej przez nią metody dążenia do pożądanego efektu końcowego (metoda future-back) oraz pokazanie, jak tę metodę można wykorzystać do wykreowania pewnej społecznej rzeczywistości i stworzenia planu działań, który ma doprowadzić do jej realizacji.

Tekst składa się z krótkiej biografii, podziękowań, ośmiu rozdziałów i wniosków. Pierwszy rozdział, zatytułowany „Sleepwalking into Singularity” („Lunatykowanie w stronę Osobliwości”, s. 1-15), rozpoczyna się od banalnego stwierdzenia, że sztuczna inteligencja (ang. Artificial Intelligence, AI) używana jest wszędzie. Autorka przedstawia rozróżnienie słabej (wąskiej) oraz silnej (ogólnej) AI oraz krótki opis realnych i możliwych czy spodziewanych zastosowań. Następnie omawia koncepcję Osobliwości, zgodnie z którą maszyny sterowane przez AI mają zyskać świadomość i przewyższyć ludzi we wszystkich aspektach, co może stanowić dla nich zagro-

żenie. Przytacza ogłoszony w roku 2015 *Open Letter on AI* („List otwarty dotyczący AI”), którego sygnatariusze nawołują do podjęcia badań nad społecznym wpływem tej technologii. Rozdział kończy się stwierdzeniem, że musimy zaakceptować nadejście Osobliwości lub czegoś do niej zbliżonego i zaplanować, jak z nią żyć. Rozdział drugi, „AI by Design and the Future-Back Methodology” („AI według projektu i metoda future-back”, s. 17-25) rozpoczyna się od wskazania czterech obszarów, którym – według omawianego „Listu...” – należy poświęcić uwagę: (1) optymalizacja wpływu AI na ekonomię, (2) prawo i etyka, (3) bezpieczeństwo i weryfikacja, (4) kontrola. Zdaniem autorki należy skupić się na planowaniu działań wobec zagrożeń płynących ze strony technologii. Pomocna ma tu być metoda future-back, stosowana w marketingu i procesie myślenia projektowego (ang. design thinking), bazującego na tworzeniu rozwiązań opartych na zrozumieniu faktycznych potrzeb potencjalnych użytkowników. Metoda future-back polega na przyjęciu w punkcie wyjścia pożądanego przyszłego stanu rzeczy i z tej perspektywy opracowaniu strategii mającej doprowadzić do jego realizacji.

Rozważania rozdziału trzeciego, „Should We Be Afraid of the Current State of AI?” („Czy powinniśmy się obawiać AI w obecnym kształcie?”, s. 27-49), otwiera

<sup>1</sup> C. Campbell, *AI by Design: A Plan for Living with Artificial Intelligence*, CRC Press, New York 2022, ss. 164.

teza, że obecnie jesteśmy świadkami swobodnego „wyścigu zbrojeń”, analogicznego do tego z czasów zimnej wojny. Celem tego wyścigu jest stworzenie silnej AI, czyli systemu przewyższającego ludzki intelekt pod każdym względem. Autorka omawia problemy wynikające z zastosowania AI w wojnie cybernetycznej i konwencjonalnej. Rozdział kończy optymistyczne stwierdzenie, że obecnie mamy jeszcze wpływ na rozwój AI i możemy tym rozwojem kierować według naszego planu. W rozdziale czwartym, zatytułowanym „Current State of AI Governance & Regulation” („Obecny stan zarządzeń i regulacji dotyczących AI”, s. 51-62), autorka wymienia szereg instytucji zajmujących się przygotowaniem zaleceń odnośnie do rozwoju AI. Zauważa przy tym, że liderzy rozwoju technologicznego – USA i Chiny – praktycznie nie ponoszą konsekwencji za ich nieprzestrzeżenie. Wskazuje także na ogromną lukę w stanie wiedzy, występującą między organami tworzącymi regulacje a korporacjami tworzącymi technologie, bo to te ostatnie mają rzeczywistą wiedzę na temat powstających rozwiązań technicznych. Niestety autorka nie podaje rozwiązań, jak tę lukę można by zniwelować. Problem jest poważny także z tego względu, że zdaniem autorki przedsiębiorstwa – nastawione na zysk – nie są godne zaufania jeśli chodzi o wprowadzanie jakichś form samoograniczenia. Przy tworzeniu zasad i regulacji rozwoju AI konieczna jest wobec tego współpraca przedsiębiorstw z rządami. Niewiele jest w książce wskazówek, jak współpraca ta miałaby wyglądać, a to właśnie od jej formy i zasad zależy jej skuteczność przy opracowywaniu regulacji badań i wdrożeń AI oraz ewentualnych sankcji za ich złamanie. Rozwiązanie proponowane przez autorkę przypomina nieco teorię umowy społecznej, ponieważ zakłada, że możliwe będzie wypracowanie korzystnego porozumienia respektowanego przez wszystkich dla dobra wszystkich.

Autorka pomija natomiast rolę organizacji pozarządowych (ang. non-governmental organizations, NGOs) w procesie nadzoru i tworzenia regulacji dotyczących AI. To poważne przeoczenie, gdyż pozostawia całość władzy w rękach rządów i korporacji, których członkowie są powiązani biznesowo i politycznie – często się zdarza, że wysocy rangą urzędnicy państwowi po zakończeniu kadencji pracują w biznesie, a specjaliści ze świata biznesu są zatrudniani w organizacjach rządowych. Z jednej strony pozwala to w jakiś sposób wypełnić wspomnianą wyżej lukę w wiedzy, ale z drugiej strony rodzi poważny konflikt interesów.

W rozdziale piątym, „Current State of AI Ethics” („Obecny stan etyki AI”, s. 63-75), autorka opisuje słynne prawa robotów, sformułowane przez Isaaca Asimova: (1) Robot nie może skrzywdzić człowieka ani przez zaniechanie działania dopuścić, aby człowiek doznał krzywdy; (2) Robot musi być posłuszny rozkazom człowieka, chyba że stoją one w sprzeczności z pierwszym prawem; (3) Robot musi chronić samego siebie, o ile tylko nie stoi to w sprzeczności z pierwszym lub drugim prawem. Prawa te trudno jednakże stosować – twierdzi Campbell – ze względu na ich dużą ogólność. Dla ilustracji przytacza klasyczny w etyce dyalemat wagonika, odnosząc go do samochodów autonomicznych. Zachęca wobec tego do bardziej elastycznego, kontekstowego podejścia, które określa jako „podejście usamodzielniające” (ang. empowerment approach) (por. s. 65). Zakłada ono, że roboty mają podejmować decyzje samodzielnie w oparciu o kontekst danej sytuacji, a nie o sztywne reguły. Najczęściej postulowane zasady to: transparentność, brak dyskryminacji oraz prawo do prywatności (por. s. 67). Proponowane przez Campbell podejście rodzi oczywiście wiele problemów, poczynając od tego, jak uczyć roboty takiej „moralnej samodzielności”, aż po kwestię moralnej, ale i prawnej odpowiedzialności

za ich decyzje. Rodzi się też pytanie, jaki status miałyby mieć takie „usamodzielnione” roboty. Czy należałoby je traktować jako podmioty moralne? Konsekwencje postulowanego podejścia nie są przez autorkę analizowane. Autorka zauważa podobieństwo między etyką AI a etyką lekarską, ale podkreśla, że regulacje dotyczące AI są na bardzo wczesnym etapie rozwoju, podobnie jak sama AI.

W rozdziale szóstym, „Options for Our Future with AI” („Możliwe wersje naszej przyszłości z AI”, s. 77-111), Campbell przedstawia w formie opowiadań pięć scenariuszy. Mają one pomóc nam uchwycić wersje przyszłości, do których nie chcemy dążyć. Nie bierze natomiast pod uwagę możliwości powstrzymania rozwoju AI – uważa to za nierealne. Pierwsza historia przedstawia dzień w Londynie, gdy rozwój AI pozostawiono samemu sobie: kwitnie cyberprzestępczość, ataki dronów są codziennością, technologiczne bezrobocie wzrasta, a społeczeństwo powoli się rozpada. Druga historia opisuje sytuację, w której rozwinięta w wysokim stopniu AI utworzyła sztuczny świat, w którym żyją ludzie (jest to wersja hipotezy symulacji Bostroma<sup>2</sup>). Trzecia historia jest kolejną wizją pesymistyczną – AI zaczęła pustoszyć planetę, a pozostali przy życiu ludzie uciekają do założonej przez Elona Muska kolonii na Marsie. Czwarta opisuje osadę zamieszkaną przez neo-luddystów, którzy wyrzekają się wykorzystania sil-

nej AI oraz są odłączeni od publicznego Internetu i postawieni poza społeczeństwem. Ostatnia historia to opowiadanie o spotkaniu dwojga ludzi, którzy poddali się szeregowi zabiegów, by przekształcić się w cyborgi. Przedstawione scenariusze są dystopijne. Aby uniknąć realizacji któregoś z nich, konieczne jest regulowanie kierunków rozwoju i zastosowań AI. Warunek takiej regulacji stanowi wypracowanie pożądanej wizji przyszłości, a wtedy – właśnie metodą future-back – będzie można określić sposoby jej realizacji.

Rozdział siódmy, zatytułowany „Creating a Roadmap – a Plan for Living with Artificial Intelligence” („Tworzenie planu działania – projekt życia z AI”, s. 113-137), zawiera główne postulaty autorki. W formie wezwania do działania przedstawia cele i ramy czasowe, w których powinny one zostać osiągnięte. Cele krótkoterminowe (2022-2025) obejmują między innymi stworzenie wspólnego modelu (ang. framework) zasad i regulacji (przy czym należy tu uwzględnić możliwie wiele różnych punktów widzenia), zwiększenie poziomu wiedzy technicznej prawodawców oraz uczynienie AI bardziej sprawiedliwą poprzez usuwanie uprzedzeń (ang. bias) ze zbiorów danych, na których opiera się jej działanie. Cele średnioterminowe (2025-2028) to: określenie zasad mających zapobiec nadużyciom i chroniących ludzi przed krzywdą wyrządzoną przy pomocy AI oraz stworzenie międzynarodowego porozumienia analogicznego do Traktatu o Nierozprzestrzenianiu Broni Jądrowej (ang. Nuclear Non-Proliferation Treaty, NPT). Jako cele długoterminowe (2025-2030) Campbell wskazuje: prowadzenie edukacji pozwalającej wszystkim lepiej zrozumieć AI i z nią współpracować, przekwalifikowanie pracowników oraz zarządzanie ekonomicznymi skutkami wprowadzenia AI na rynek pracy. Autorka nie wyznacza środków do osiągnięcia tych celów ani nie wskazuje, kto ma prowadzić i nadzorować edukację prawodawców oraz obywateli.

<sup>2</sup> Hipoteza: zaawansowana cywilizacja mogłaby stworzyć symulację (tj. komputerowe odwzorowanie rzeczywistości) obejmującą nasze umysły oraz otaczający nas świat (podobnie jak w filmowej trylogii *Matrix*). Główny argument polega na tym, że jeśli żyjemy wewnątrz takiej symulacji, to nie poznajemy „prawdziwego” świata. Ponadto nie jesteśmy w stanie zweryfikować czy żyjemy w symulacji, czy też nie (por. N. B o s t r o m, *Are You Living in a Computer Simulation?*, „Philosophical Quarterly” 53(2003) nr 211, s. 243-255.

Rozdział ósmy, „A New Hope for Artificial Intelligence” („Nowa nadzieja dla AI”, s. 139-146), zawiera przemyślenia autorki dotyczące pożądanej wersji przyszłości. Campbell twierdzi, że odpowiednie zarządzanie AI może przynieść ludzkości ogromne korzyści. Pojawienie się Osobliwości może oznaczać nadejście nowego renesansu – ludzie będą mieć znacznie więcej wolnego czasu, dzięki czemu będą mogli rozwijać swoje zainteresowania, a ludzkość jako całość będzie dysponować znacznie większymi zasobami intelektualnymi, co znacząco przyspieszy jej rozwój we wszystkich aspektach.

W części pod tytułem „Conclusion” („Wnioski”, s. 147-148) autorka wskazuje, że AI jest jednocześnie największym wyzwaniem dla ludzkości oraz jej największą nadzieją. Jest „boska” (ang. godlike) i tak rozległa, że zrozumienie jej ogromu przekracza możliwości pojedynczego człowieka. Dzięki niej zyskamy możliwość decydowania, kim chcemy być (por. s. 148), a wobec tego musimy zaplanować, jak chcemy żyć w świecie z AI.

Książka ma charakter popularnonaukowy, ale jest interesująca ze względu na to, że autorka znajduje się w gronie osób mających wpływ na realnie powstające rozwiązania techniczne. Campbell podejmuje tematy pojawiające się często w popkulturze, stosuje język potoczny i powołuje się na opinie „technologicznych celebrytów” pokroju Elona Muska (por. s. 11-15). Układ treści jest nieco chaotyczny. Podziękowania oraz pierwsze dwa rozdziały de facto stanowią właściwy wstęp. Struktura książki zostaje zarysowana dopiero w drugim rozdziale. Zasadnicze rozważania znajdujemy w rozdziale siódmym – zawiera on najważniejsze tezy i postulaty. Rozdziały nie są symetryczne: najdłuższy ma trzydzieści siedem stron, a najkrótszy jednaście. Pierwszą częścią po spisie treści jest krótka i niezbyt informatywna biografia. Nie dowiemy się

z niej, czym konkretnie zajmuje się autorka. Z treści książki wynika, że zawodowo jest doradcą biznesowym, specjalizującym się w design thinking, nie zostało to jednak sformułowane explicite. W tekście pojawia się także marketingowy żargon (np. testy A/B – por. s. 18) oraz akronimy instytucji (np. RNIB, RNID – por. s. 20), które nie zostały wyjaśnione, co utrudnia czytanie tekstu. Na stronie sześćdziesiątej pierwszej pojawia się błąd rzeczowy – przytoczono kasus procesu Sokratesa, słynnego filozofa i mentora Platona, a w następnym zdaniu stwierdzono, że to Platon został skazany na śmierć poprzez otrucie. Oczywiście skazanym był Sokrates, a Platon dożył spokojnej starości.

Autorka nie jest deterministką techniczną. Przyszłość pełni w jej rozważaniach rolę swoistej zasłony niewiedzy – trzeba podejmować decyzje, nie znając faktycznego przebiegu rozwoju AI, ale mając na uwadze pożądany scenariusz przyszłości. Campbell twierdzi, że rozwoju nie da się zatrzymać, ale jesteśmy w stanie nim kierować poprzez regulacje wypracowane przez kooperację rządów, korporacji oraz instytucji pozarządowych (NGO, tj. wszelkiego rodzaju fundacji i stowarzyszeń). Jest to dość utopijne podejście. Zakłada uczciwość wszystkich stron oraz możliwość stworzenia rozwiązań sprawiedliwych i korzystnych dla wszystkich. Wydaje się nie zauważać znaczącego konfliktu interesów w obszarze regulacji dotyczących AI. Ponadto zgoda na poziomie postulatów nie gwarantuje zgody na poziomie środków realizacji – o środkach i sposobach wypracowania owych regulacji autorka nie mówi zbyt wiele, a przecież zastosowanie środków też niesie ze sobą konsekwencje. Co więcej, w rozważaniach widoczny jest „zachodni” punkt widzenia, zakładający, że rządy demokratyczne będą zainteresowane rozwojem AI dla wspólnego dobra. Campbell w rozważaniach ignoruje istnienie niedemokratycznych czy wręcz auto-

kratycznych rządów, które też inwestują duże środki w rozwój AI i raczej nie kierują się w swoich działaniach wspólnym dobrem.

Zamysł autorki dobrze wpisuje się w koncepcję budowy tak zwanej odpowiedzialnej AI (ang. responsible AI), choć sformułowanie to nie pojawia się w tekście. Takie postulaty zwykle marginalizują rolę filozofii. Podobnie jest i tutaj: Campbell podejmuje rozważania na istotne filozoficznie tematy, pomija jednak rolę samej filozofii w wypracowywaniu rozwiązań. Traktuje etykę jako inspirację i punkt odniesienia przy projektowaniu regulacji, ale nie jako narzędzie do rozstrzygania problematycznych kwestii. Heurystyczna rola filozofii została całkowicie zignorowana. Sama autorka zaznacza udział filozofów w procesie wypracowania rozwiązań wewnątrz korporacji oraz w konsultacjach rządowych (por. s. 68, s. 71), udział ten jednak zazwyczaj ogranicza się do formułowania ogólnikowych zaleceń i wskazówek. Jest to bardzo niekorzystne dla wszystkich stron. Książkę można odczytać jako wezwanie pod adresem filozofów, by włączyli się aktywnie w działania dotyczące regulacji AI. Przeważnie rzadko zajmują się bieżącymi problemami, a obecnie na świecie dzieją się rzeczy wymagające podejmowania decyzji już teraz. Zwlekanie z wypracowaniem konkretnych stanowisk może doprowadzić do przeciwnego momentu, w którym jeszcze mamy wpływ na to, co się dzieje. Bez konkretnych stano-

wisk i związanych z nimi postulatów filozofia zostaje sprowadzona co najwyżej do roli komentatora. Staje się jednym spośród wielu źródeł opinii.

Systematyczny namysł nad techniką i konsekwencjami jej rozwoju, nawet związanego z AI, ekscytuje na ogół jedynie wąskie grono pasjonatów. Społeczeństwo interesuje się raczej spektakularnymi wydarzeniami, zwłaszcza katastrofami (czy możliwościami katastrofy). Dlatego też książki stanowiące próby zainteresowania szerszego grona taką tematyką zasługują na uznanie. Autorka jest osobą „nietekniczną”, toteż stosowane przez nią opisy zawierają niewiele informatycznego żargonu, który mógłby odstraszyć czytelników. Książka stanowi ciekawą propozycję dla wszystkich zainteresowanych tematyką AI oraz nowych technologii – zawiera ogólny przegląd obecnego stanu regulacji oraz zachęca do podjęcia rozważań co do przyszłości człowieka i jego miejsca w świecie zdominowanym przez technologię. Jest również bardzo aktualna ze względu na opracowywane właśnie przez Parlament Europejski regulacje dotyczące zastosowań AI (tzw. AI Act), które uważane są za kluczowe, ponieważ będą wzorcem dla rozwiązań w innych częściach świata.

Kontakt: Wydział Filozofii, Katolicki Uniwersytet Lubelski Jana Pawła II, Al. Racławickie 14, 20-950 Lublin  
E-mail: [szczech.dam@gmail.com](mailto:szczech.dam@gmail.com)



## PROPOZYCJE „ETHOSU”

Joe Puckett Jr., *Apologetyka Radości. C.S. Lewisa argument z Pragnienia na rzecz istnienia Boga*, tłum. Piotr Bylica, Fundacja Prodoteo, Warszawa 2023, ss. 205.

Joe Puckett deklaruje, że celem książki „jest podważenie nacechowanego sceptycyzmem trendu (niekiedy określanego mianem «nowego ateizmu»), który uznaje publikacje starające się zachęcać do wiary w Boga za głupie. Wiara w Boga wcale nie jest głupia. Wiara w istnienie tego, co boskie, jest zakorzeniona w ludzkiej myśli i ludzkim doświadczeniu” (s. 14). Książka stanowi obronę tego ostatniego twierdzenia, a zatem coś więcej niż jedynie podważenie nowego ateizmu. Sednem obrony racjonalności wiary poprzez obronę twierdzenia o istnieniu Boga jest tak zwany argument z Pragnienia – zasadniczo w wersji zaproponowanej przez Clive’a S. Lewisa.

Książka, opatrzona przedmową przez Marka Linville’a, podzielona została na cztery części, które składają się z rozdziałów (mają one numerację ciągłą), całość poprzedzono *Wstępem* i *Wprowadzeniem*, a zakończono indeksem osobowym i rzeczowym oraz bibliografią. Punkt wyjścia rozważań stanowi problem rzekomej absurdalności świata: oto znajduje się w nas głęboka potrzeba sensu, a zdaniem tak wielu myślicieli, wszechświat jest go pozbawiony. Czy znajdujemy się w takiej sytuacji – pyta Puckett – że najgłębsze z pragnień ludzkości, pragnienie sensu, nigdy nie może zostać zaspokojone? Cokolwiek byśmy na ten temat sądzili, każdej potrzebie z pewnością odpowiada pragnienie jej zaspokojenia. Jako ludzie rozwinięliśmy wrodzoną umiejętność pragnienia rzeczy, które zaspokajają nasze potrzeby. Czy to, że odczuwamy naturalną potrzebę sensu, a nasz wszechświat nie może go zaspokoić, oznacza, że sensu w ogóle nie ma, czy też ta niespełniona tęsknota za sensem może wskazać na realne istnienie Transcendencji? Puckett opowiada się za drugim członem alternatywy i jego rozważania mają służyć uzasadnieniu tego stanowiska.

Część pierwsza książki poświęcona jest argumentowi z Pragnienia, ulokowanemu na tle innych argumentów na rzecz istnienia Boga. Argument ten głosi, że każdemu naturalnemu i powszechnemu pragnieniu odpowiada istniejący przedmiot, który je zaspokaja, i jeśli wykażemy, że wszyscy ludzie posiadają naturalne pragnienie Boga (czy Transcendencji), to Bóg musi istnieć. Argument ten odwołuje się do pierwszoosobowego doświadczenia, do tęsknoty za czymś

wykraczającym poza wszystko, co świat jest w stanie nam dać. Tło analiz autora stanowią próby wyjaśniania tego rodzaju pragnienia i tęsknoty przez Friedricha Nietzschego i Sigmunda Freuda. Puckett pokazuje szczegółowo, dlaczego próby te nie są satysfakcjonujące: ludzie tęsknią za czym więcej niż tylko za poczuciem bezpieczeństwa, a do tego właśnie odwoływali się wymienieni myśliciele. Puckett formułuje argument z Pragnienia, odwołując się do niebeletrystycznych dzieł Lewisa. Wychodzi od wyrażanego przez Lewisa przekonania, że gdybyśmy zajrzeli uczciwie w siebie, znaleźlibyśmy pragnienie, które na ziemi nigdy do końca nie może zostać zaspokojone. Zdaniem Lewisa ludzie radzą sobie z tym faktem na trzy sposoby. Pierwszy to sposób „głupca” (por. s. 32n.), który o niezaspokojenie tej potrzeby obwinia rzeczy i uważa, że gdyby posiadał rzeczy lepsze (lepsze auto, lepsze wakacje czy lepszą pracę), to jego tęsknota za Transcendencją zostałaby zaspokojona. Drugi to sposób „człowieka rozsądnego” (por. s. 34), który kontroluje swoje oczekiwania i „trzyma się” tego, co namacalne. I wreszcie sposób chrześcijanina, który odnajdując w sobie pragnienie, którego nie może zaspokoić żadne doświadczenie na tym świecie, przyjmuje, że najprawdopodobniej został stworzony dla innego świata. Argument z Pragnienia ma pokazać racjonalność i wartość tego trzeciego sposobu. To właśnie pragnienie Transcendencji jest formą Radości – różnej od „zwykłych” radości także dlatego, że zawiera ona ból aktualnego niezaspokojenia, niepewność co do tego Kim (Czym) jest to, za czym tęsknimy, docenienie wszystkiego, co piękne i inspirujące na ziemi, a jednocześnie niezadowolenie z rzeczy jako niezdolnych do zaspokojenia tęsknoty. Właśnie ze względu na te różnice, słowo „Radość” pisane jest dużą literą. Puckett wskazuje, że Radość u Lewisa określana jest niemieckim terminem „Sehnsucht”. Cytuje tu Lewisa, dla którego Sehnsucht to niezaspokojone pragnienie, które samo w sobie było bardziej godne pożądania niż cokolwiek, co mogłoby je zaspokoić: „Taki właśnie stan nazywam Radością. Określenie to należy potraktować jako pewien ścisły, techniczny termin – Radość oznacza tu coś zupełnie innego niż szczęście czy przyjemność. Łączy ją z nimi tylko jedno, a mianowicie to, że kto raz doświadczył Radości, będzie jej pragnął znowu” (s. 41). Sehnsucht nie oznacza po prostu tęsknot życiowych: „[Jest – A.L.K.] bolesną i jednocześnie przyjemną, intensywną tęsknotą za życiem, którego obecnie nie mamy, ale którego ludzie w sposób naturalny i powszechny pragną. Jest to uczucie utraty czegoś, co kiedyś mieliśmy, wywołujące w nas poczucie tęsknoty za domem i niezadowolenia nieidealnym światem, w którym obecnie żyjemy” (s. 46). Pragnienie to można jednakże zagłuszyć i odrzucić. Przywołując przykład wujka Andrzeja z Narni, który wbrew świadectwu zmysłów odrzuca własne doświadczenie, zgodnie z którym lew faktycznie mówi, Puckett pokazuje, iż – zdaniem Lewisa – o tym, w co człowiek będzie wierzył, decyduje coś więcej niż tylko intelektualna ciekawość i świadectwa na rzecz danego twierdzenia, w grę wchodzi bowiem także czynnik moralny. Ostatni rozdział części pierwszej poświęcono różnym sformułowaniom argumentu z Pragnienia w wersjach dedukcyjnych i indukcyjnych. Ostatecznie argument omawiany jest w następującej formie: Przesłanka 1: Każdemu naturalnemu lub wrodzonemu pragnieniu, jakie odczuwamy, odpowiada właściwy realnie istniejący przedmiot, który może to pragnienie zaspokoić.



Przesłanka 2: Istnieje w nas naturalne pragnienie, którego żadna istniejąca w czasie rzecz, nic, co spotykane na ziemi, żadna rzecz stworzona, nie może zaspokoić.

Wniosek: Istnieje coś poza czasem, ziemią i rzeczami stworzonymi, co może zaspokoić to pragnienie. To coś wiele osób nazywa Bogiem lub niebem.

Samo w sobie – zaznacza Puckett – powyższe wnioskowanie to nie dowodzi istnienia chrześcijańskiego Boga, ale pokazuje, że doświadczane przez ludzi naturalne pragnienie stanowi wiarygodne źródło racjonalnej wiary w istnienie nadnaturalnej rzeczywistości i nadnaturalnej Istoty.

Część druga zawiera analizę zarzutów Johna Beversluisa przeciwko argumentowi z Pragnienia w zaproponowanym przez Lewisa sformułowaniu. W kolejnych rozdziałach tej części Puckett analizuje następujące zarzuty Beversluisa: popełnienie błędu *petitio principii*, fałszywe uznanie pragnienia Boga za pragnienie naturalne (czy wrodzone), ucieczkę Lewisa przed pragnieniem Boga, niezgodność idei Boga jako powszechnego obiektu pragnienia ze stwierdzeniami Biblii oraz fakt istnienia ludzi, którzy nigdy nie doświadczają pragnienia Transcendencji. Puckett odpowiada na te zarzuty, pokazując materialistyczne założenia Beversluisa, ignorowanie świadectw (także biblijnych), a także błędne identyfikowanie racji Lewisa, w tym charakteru pragnienia i Radości. To ostatnie założenie okazuje się kluczowe. Pragnienia bowiem, które Bóg w nas umieścił, są naturalne, ale nie mają charakteru przymusu. Nasza wolność obejmuje także zdolność kontrolowania, a nawet tłumienia naszych naturalnych instynktów, w tym pragnienia Transcendencji. Ponadto z perspektywy chrześcijańskiej żyjemy w świecie „upadłym po grzechu”, w świecie, w którym pragnienie Boga zagłuszone jest przez pychę, strach czy bunt – nasze skupione na sobie ego, nasze pragnienie absolutnej autonomii sprawia, że stajemy się ślepi na wartość wiary. „Konsekwentnie – pisze Puckett – Bóg jest wystarczająco blisko, byśmy mogli Go znaleźć, ale jednocześnie wystarczająco daleko, abyśmy mogli wybrać, że nie chcemy tego zrobić. zawsze będziemy mogli w wolny sposób odrzucić ten boski głód lub przekonać sami siebie, że głód ten jest pragnieniem zupełnie innego obiektu” (s. 96). Co więcej, Pragnienie nie jest stanem trwałym – to raczej momenty przeblysku Transcendencji, uniesienia i zarazem tęsknoty. Chodzi o to, byśmy tych momentów nie przegapili i nie zagłuszyli, ale uczciwie i wnikliwie wejrzeli w głąb własnego „ja”, by odkryć właśnie to Pragnienie-Radość. I nawet jeśli wcześniej je zagłuszyliśmy, będziemy mogli je dostrzec.

Zarzuty Beversluisa dotyczyły zasadniczo pierwszej przesłanki argumentu z Pragnienia. Część trzecia książki Pucketta poświęcona została przesłance drugiej, która głosi, że istnieje w nas pragnienie, którego żadna istniejąca na tym świecie rzecz nie może zaspokoić. Puckett proponuje, aby myśleć o tym pragnieniu jako o „echu” Transcendencji – słysząc echo, zazwyczaj nie widzimy źródła dźwięku, ale echo wskazuje na jego istnienie: „Nasze własne życie ma być sposobem, w jaki ta «święta tajemnica» ma przemówić nie tylko do naszego intelektu, ale też do serca” (s. 115). Autor rozważa trzy „echa” Transcendencji: baśnie (wyobraźnia), piękno oraz wypoczynek. Analizując baśnie, przypomina, że – zdaniem Lewisa – jednym ze sposobów lekceważenia naturalnej tęsknoty za Bogiem jest trywializowanie jej jako jedynie dziecięcego

przekonania, podobnego do wiary w Świętego Mikołaja. Jako dorośli porzucamy „głupie bajki” i zanurzamy się w widzialny świat naszego codziennego doświadczenia. Puckett pyta za Lewisem: A jeśli to dzieci mają rację, że to, co uznajemy za „prawdziwe” życie, nie jest jedyną istniejącą rzeczywistością? Być może ludzie „dlatego przychodzą na ten świat z przekonaniem, że wszystko jest cudem, ponieważ wszystko rzeczywiście jest cudem” (s. 123), i stąd poczucie zależności od Kogoś (Czegoś) wyższego? Być może jest tak, że nie dzieci, ale właśnie „septyczni nie są w stanie dostrzec piękna boskiego lasu, bo zaślaniają im go molekularne struktury poszczególnych drzew” (s. 125). Drugie „echo” Transcendencji stanowi piękno. Puckett cytuje tu Lewisa (*Brzemienie chwały*): „Nie chcemy tylko oglądać piękna [...]. Chcemy jeszcze czegoś, co trudno ująć w słowa – zjednoczenia z obserwowanym pięknem, wnikięcia w nie, wchłonięcia go w siebie, pławienia się w nim, połączenia z nim” (s. 135). To „zachwycające” piękno – komentuje Puckett – motywuje nas do praktykowania miłości i sprawiedliwości, ale postrzegane jako cel sam w sobie, może stać się bałwochwalczym idolem. Ostrzeżeniu przed ubóstwieniem piękna towarzyszy inne: konsekwencją współczesnego odrzucenia obiektywnej prawdy jest odrzucenie obiektywnego piękna, a to pozbawia nas możliwości dostrzegania w napotykanym przejawach piękna ich Stwórcy-Artysty. Tymczasem piękno stworzonego świata wskazuje na piękno Stwórcy i piękno innego świata – co nie neguje i nie pomniejsza piękna dojrzanego w stworzonym świecie i w sztuce w tym świecie tworzonej. „«Piękno» w najszerszym sensie to nazwa pragnienia, które nie pochodzi z tego świata. W ten sposób piękno jest boskim posłannikiem, przekazującym nam wiadomość o innym świecie, który zawiera piękno większe od samego posłannika” (s. 146).

Trzecie „echo” to odpoczynek, rozumiany jako „odbicie” boskiego życia wypełnionego kontemplacją i wychnieniem. Odpoczynek jest raczej umysłową i duchową postawą niż prostym „nienierobieniem”. W odpoczynku nie chodzi o odtworzenie sił, by jeszcze wydajniej pracować i zdobywać kolejne rzeczy zaspokajające pragnienia, ale o wyłączenie się, o wewnętrzny pokój i ciszę – wówczas być może pojawi się ów przeblask Transcendencji. Odpoczynek – dodaje Puckett – może być najbliższy temu, za czym tęsknimy, ponieważ oferuje przynajmniej spokój umysłu i pełnię ducha, której inne działania nie pozwalają uzyskać. Bez wiary nasz świat staje się mały, a „gdy twój świat jest zbyt mały, to nie ma w nim miejsca, w którym można by zatańczyć” (s. 155).

W części czwartej przedstawiono analizę wniosku argumentu z Pragnienia. Wniosek ten głosi, że istnieje coś poza czasem i przestrzenią naszego wszechświata, co może zaspokoić Pragnienie – a najlepszym kandydatem do roli czegoś istniejącego poza czasem i przestrzenią jest Bóg. Puckett stawia jednak pytanie: A może to nie teizm, ale ewolucyjni naukowcy znaleźli bardziej odpowiednie „coś”, co ma zaspokajać pragnienie Transcendencji? Puckett analizuje tezę ewolucjonistów, iż religia wzięła się ze strachu przed śmiercią. Strach paraliżuje działania i wobec tego nie sprzyja przetrwaniu i to dlatego w drodze ewolucji powstały w naszym mózgu połączenia neuronalne, które powodują, że postrzegamy siebie jako istoty duchowe i wieczne. Ale – pyta rzeczowo Puckett – dlaczego z tego, że ludzie są zaprogramowani, by wierzyć w Boga, miałyby wynikać, że Bóg nie istnieje? Równie dobrze można by uznać, że

skoro Bóg istnieje, to stworzył ludzi tak, by mieli wrodzoną skłonność do wiary w Niego. Co więcej, trudno uznać za rację wiary strach przed śmiercią, skoro wedle obecnej wiedzy, większość ludzkich kultur zakładała, że ludzie nie przestają istnieć po śmierci. Puckett podsumowuje: „Nawet jeśli zgodzimy się, że odwołanie do ewolucji wyjaśnia naturalną skłonności do wiary w jakąś transcendentną istotę, to wciąż pozostaje nam ustalenie, jakie jest najlepsze wyjaśnienie tego, dlaczego tak się rzeczy mają. [...] Dlaczego ślepy dobór naturalny miałby faworyzować tęsknotę za czymś, co nigdy nie istniało? [...] Czy nie jest bardziej sensowne przekonanie, że tęsknota za transcendencją pochodzi od istoty transcendentnej? [...] Choć odwołanie do ewolucji może stanowić wyjaśnienie mechanizmu biologicznej przyczyny takiego pragnienia, to nie wyjaśnia ono, dlaczego w ogóle miałby istnieć taki mechanizm” (s. 164n.). Dlatego, wykorzystując tak zwaną abdukcję, należy uznać, że wyjaśnienie tęsknoty za Transcendencją przez wskazanie na istnienie Istoty Transcendentnej jest lepsze, bardziej poznawczo satysfakcjonujące. Puckett pokazuje, że również wyjaśnienie zaproponowane przez Richarda Dawkinsa, który twierdzi, iż pragnienie Transcendencji stanowi produkt uboczny zjawisk adaptacyjnych umożliwiających przetrwanie, nie obala wniosku argumentu z Pragnienia. Teoria Dawkinsa wyjaśnia istnienie ciągłości przekazywania prawdziwych lub fałszywych informacji, ale nie odpowiada na pytanie, skąd pochodzi idea transcendentnej rzeczywistości i dlaczego się pojawiła. Co więcej, ewolucyjny zarzut Dawkinsa nie wskazuje kryterium rozstrzygania, które przekonania są prawdziwe, a które stanowią nonsens. Tak jak ewolucja wytworzyła prawdziwe przekonanie, że pływanie w rzece pełnej krokodyli jest niebezpieczne – podsumowuje Puckett – tak mogła wytworzyć prawdziwe przekonanie, że istnieje Bóg, który zapoczątkował proces ewolucji. Autor przywołuje tu Swinburne’owską zasadę zawierzenia (zasadniczo należy wierzyć, że sprawy mają się tak, jak się wydaje, o ile nie istnieje jakiś przekonujący powód, by uważać, że jest inaczej, niż się wydaje) oraz zasadę świadectwa (powinniśmy wierzyć w to, co o stanie rzeczy mówią nam inni ludzie, o ile nie ma jakiegoś przekonującego powodu, by wierzyć, że jest inaczej niż mówią). W świetle tych zasad i poprzednio dyskutowanych przesłanek – twierdzi Puckett – argument ewolucyjny nigdy nie obali teizmu. W ostatnim rozdziale części czwartej autor stawia następujące pytania: Jeśli opisywana przez Lewisa tęsknota za Bogiem (Radość czy Sehnsucht) faktycznie jest naturalna, to czy możemy uczucie to wyjaśnić wyłącznie jako wytwór odziedziczonych genów? A jeśli tak, to czy wyjaśnienie takie podważa wniosek, że musi istnieć Bóg, który stanowi zaspokojenie naszej tęsknoty? Odpowiedź jest, jak można oczekiwać w świetle wcześniej przedstawionych rozważań, przecząca.

Ostatecznie Puckett stwierdza: „Argument z Pragnienia nie dowodzi, że Bóg istnieje, a już tym bardziej nie dowodzi, że istnieje Bóg teizmu chrześcijańskiego. Stara się ukazać, że wiara, iż musi istnieć coś poza tym światem, co jest źródłem naszej tęsknoty, jest racjonalna” (s. 183). Zaznacza też, że argument ten najlepiej ujmować nie tylko w kontekście całościowego wnioskowania na rzecz istnienia Boga, ale także w kontekście tego, co „odczuwamy w swoich ‘trzewiach’” (s.184). Tęsknota za Transcendencją jest ostatecznie tęsknotą nie za lepszym światem, ale za Osobą. Rozważania kończy odwołanie do Lewi-

sowkich *Opowieści z Narnii*. Jedną z postaci, mysz Rycypisk, przedstawia swój własny argument z Pragnienia. Rycypisk towarzyszy księciu Kaspianowi w misji odnalezienia siedmiu zaginionych lordów. Jest to długa i pełna niebezpieczeństw podróż, ale tym do czego naprawdę tęskni jej serce była Kraina Aslana. I do tej Krainy ostatecznie dociera. „Rycypisk – pisze Puckett – to wielka postać, która przypomina nam o głębokiej tęsknocie, która jest w nas. Powtórzmy, że nie jest to po prostu uczucie. Jest to odczucie, że potrzebuje się i pragnie czegoś więcej. Jest to niespokojny duch, który nie zazna pokoju, aż nie dotrze do przedmiotu, którego pragnie” (s. 186). Choć Puckett nie traktuje wypowiedzi bohatera jako osobnego argumentu, warto zwrócić uwagę, że jest to odpowiedź na zarzut – często formułowany przez społecznych aktywistów – że zwrócenie się ku Transcendencji prowadzi do zaniedbania tego świata. W ujęciu Lewisa nic takiego nie ma miejsca. Rycypisk prowadzi bogate, pełne przygód życie, wspiera Księcia Kaspiana w walce z podłym Mirazem, angażuje się w ratowanie lordów. „Myślą się ci – podsumowuje Puckett – którzy uważają, że ten świat jest tylko czymś przejściowym, miejscem, w którym jedynie czekamy na to, co prawdziwe. Ten świat jest miejscem, w którym uczymy się tańczyć przed rozpoczęciem dawno wyczekiwanego balu. Jest to świat, w którym kochankowie śpiewają, aż ich Radosna pieśń zostanie usłyszana w anielskich chórach. Ten świat to nie poczekalnia. To sala treningowa” (s. 187).

Książkę kończy „Dodatek”, w którym Puckett rozważa naszą pamięć o domu – ogrodzie Eden: gdyby nie grzech, Eden rozszerzyłby się na cały świat. Z powodu grzechu to, co pierwotnie było dla nas przygotowane, obecnie jest tym, o czym jedynie marzymy i myślimy. Sehnsucht to właśnie wyraz pamięci o domu. „Pamiętać” to czynić przeszłe zdarzenia niejako elementem naszego własnego doświadczenia. Pamięć ma swoje źródło w tradycji: odczuwania tęsknoty za pokojem zamiast wojny, za szczęśliwością zamiast chaosu, za sprawiedliwością. Także z powodu pięknych rzeczy tego świata pamiętamy o Edenie. Pamiętamy o nim również dzięki Biblii, która nam go opisuje. Z pamięci rodzi się pragnienie. Pragniemy powrócić do domu, który trwa w naszej pamięci. Istotnych wątków rozważań Pucketta jest oczywiście znacznie. Gdybyśmy uwzględnili współczesne standardy naukowości, książka nie ma charakteru naukowego. Zawiera uproszczenia, skróty myślowe, jednostronne oceny. I jest to zgodne z zamysłem autora, który wydaje się podejrzewać świat akademicki o tłumienie tytułowego Pragnienia w imię ciasno rozumianej empiryczności i instrumentalnej racjonalności. Być może dlatego o rozdziale poświęconym różnym sformułowaniom argumentu z Pragnienia pisze on w tonie usprawiedliwienia: „Choć niektórzy czytelnicy uznają ten rozdział za czysto akademicki, to służy on ważnemu celowi” (s. 53) – ostatecznie ocenie zasadności i siły argumentu z Pragnienia.

Na podstawie tytułu czytelnik mógłby się spodziewać książki z zakresu „lewisologii” i oczywiście książka jest skierowana do osób zainteresowanych twórczością Lewisa. Z pewnością też zachęci ona do lektury dzieł tego autora. Sądzę jednak, że to nie jego idee są ostatecznie ważne. Owszem, dzieła Lewisa dostarczają materiału, ale centralny przedmiot analiz stanowi sam argument. Rozważania Lewisa są dla Pucketta okazją do postawienia kluczowego pytania: Czy istnieje Bóg, który stworzył świat i któremu wypowiedzieliśmy naszą mi-

łość i posłuszeństwo, a który tak umiłował świat, że Syna swego zesłał byśmy mieli życie wieczne z Nim? Analizując argument z Pragnienia i jego krytykę, uzupełniając Lewisowską interpretację o idee współczesnych teistów, Puckett wykazuje, że odpowiedź pozytywna na to pytanie jest racjonalna i dobrze uzasadniona, ugruntowana w naturze rzeczywistości. Nawet jeśli argument z Pragnienia nie wskazuje na chrześcijańskiego Boga w Trójcy Świętej Jedyne, to Bóg taki jest doskonałym „kandydatem” do roli Transcendencji zaspokajającej powszechne ludzkie pragnienie. Nie musimy się więc wstydzić wiary jako czegoś irracjonalnego, naiwnego i nienaukowego ani traktować jej jako „prywatnego poglądu”. Po stronie tezy o istnieniu pragnienia Transcendencji stoją zresztą dane: osób wierzących w jakiś rodzaj Transcendencji wydaje się więcej niż zadeklarowanych ateistów; a teza, że religia zniknie wraz z rozwojem nauki i dobrobytu okazała się fałszywa. Swe rozważania Puckett kieruje wobec tego do – jak ich nazywa – „myślących chrześcijan” (s. 13), by pokazać, że wiara jest czymś wsobnym dla natury ludzkiej, a rozsądnie można przypuszczać, że „zasianym” przez Boga. „Z tego nasienia wiary (w połączeniu z objawieniem się Boga, z jakim mamy do czynienia w Biblii) – pisze – może rozwinąć się życie przepełnione czymś o wiele wspanialszym niż wszystko, co może zaoferować świat w jego obecnym stanie” (s. 13). Kieruje swą książkę także do poszukujących – by pokazać, dlaczego warto starać się o wiarę i z nadzieją, że książka pomoże „uczynić pierwsze kroki w stronę wiary (s. 14).

*Apologetykę Radości* można czytać jako rekonstrukcję myśli Lewisa albo jako rekonstrukcję argumentu z Pragnienia. Być może pojawi się wówczas rodzaj dyssatysfakcji – że książka jest niewystarczająco analityczna, że nie uwzględnia rozmaitych nasuwających się aspektów problemu, że upraszcza trudne kwestie. Czytając ją w taki sposób, pominiemy jednak coś istotnego. Puckett lokuje argument z Pragnienia na tle współczesnych dyskusji – stąd odwołania do znaczących myślicieli, takich jak Alvin Plantinga, Richard Swinburne, Josef Pieper, Alister McGrath, Peter Kreeft, Richard Dawkins, Mortimer J. Adler, Gilbert K. Chesterton, Sigmund Freud, Friedrich Nietzsche, Rudolf Otto czy Clifford Williams, by wymienić jedynie niektórych. Sięga też do klasyków filozofii: Platona, Arystotelesa, św. Augustyna, św. Tomasza z Akwinu. Nie chodzi mu jednak o rekonstrukcję i wyczerpującą analizę poglądów tych myślicieli. Swobodnie wykorzystuje ich idee, by wzmocnić argument z Pragnienia, także przez ukazanie niezasadności skierowanej przeciwko niemu krytyki. Pokazuje przy tym długą i bogatą historię namysłu nad problemem wiary i istnienia Boga. Najwyraźniej jest to problem przypisany do naszej ludzkiej kondycji. Dlatego właśnie książka Pucketta jest dla nas, ale i o nas. Jest ona dla nas, bo jest zaproszeniem, by we własnym życiu uczciwie i odważnie poszukiwać doświadczeń, w których przebłykuje pragnienie Transcendencji. Nie możemy mieć pewności, że je znajdziemy – ale z pewnością nie znajdziemy ich, jeśli nie spróbujemy. Jest też i o nas. Pokazuje przemiany kultury – na przykład oddzielenie faktów od wartości, nauki od religii, piękna od prawdy i dobra – które sprawiają, że trudniej nam pragnienie Transcendencji w sobie identyfikować, a w świecie dostrzegać „echa” Transcendencji. Zapewne też wielu z nas odkryje, że z pragnieniem, którego nie zaspokoi żadna rzecz tego świata, często radzimy sobie na sposób głupców – poszukując coraz „lepszyc” rzeczy, i na

sposób człowieka rozsądnego – odmawiając wagi temu pragnieniu. I zapewne wielu odkryje siebie w opisach tych, którzy tkwią w „wiecznym niedoczasiu”, nie wiedząc, co to odpoczynek, którzy nie potrafią odkryć prawdy w baśni, którzy nauczyli się żyć głusi na głos Sehnsucht, którzy nie potrafią sięgać myślą ponad gwiazdy. Analizy argumentu z Pragnienia okażą się dla wielu „portretem własnym”. Swoje rozważania Puckett kończy następująco: „Ostatecznie argument z Pragnienia skłania nas do tego, by zadać właśnie to pytanie: «A może tak właśnie jest?». Jeśli jesteśmy otwarci na możliwości, jakie kreśli ten argument, możemy uzyskać na to pytanie zniewalająco piękną odpowiedź” (s. 194). Warto postawić to pytanie, odpowiedź bowiem ma fundamentalną wagę dla naszego c a ł e g o życia. I być może w świetle argumentu z Pragnienia warto zastanowić się nad radykalnym twierdzeniem Lewisa: „Ostatecznie ludzie dzielą się tylko na dwie kategorie: tych, którzy mówią Bogu «bądź wola Twoja i tych, do których Bóg powie na końcu: «bądź wola twoja» (s. 96). To dlatego i wierzący, i poszukujący, i niewierzący powinni po tę książkę sięgnąć.

*A.L.K.*





Dorota CHABRAJSKA

## GWARANCJA DLA ODROBINY UMYŚLU

Zamówiony w polskiej odnodze największego sklepu internetowego na świecie, wąż prysznicowy dotarł pod drzwi odbiorcy w oszałamiającym tempie niecałych dwudziestu czterech godzin. Tempo było w istocie imponujące, zważywszy, że od momentu dokonania jego zakupu wąż pokonał trasę wiodącą z południa Francji (gdzie znajdował się jego macierzysty magazyn), przez Niemcy i Warszawę, aż do Lublina. Drogę węża i jego wszystkie (chwilowe) przystanki można było na bieżąco śledzić, transportem zajęła się bowiem jedna z największych międzynarodowych firm kurierskich, udostępniająca klientom możliwość monitorowania w czasie rzeczywistym lokalizacji zdążającego do nich przedmiotu.

Równie zdumiewająca jak tempo drogi węża była przesłana przez sklep internetowy uprzejma informacja, że produkt sprzedawany jest „w pakiecie z pięcioletnią gwarancją dla odrobiny umysłu”. O ile wiadomość, że umysł nabywcy węża będzie w podstawowym zakresie funkcjonował jeszcze przynajmniej przez pięć lat (a w przypadku odchylenia w jego pracy sklep okaże się pomocny), można by uznać za budującą, o tyle w kimś posługującym się w miarę poprawną polszczyzną mogła ona wzbudzić pewne zdziwienie, wszak w języku polskim gwarancja jest „gwarancją czegoś” bądź „gwarancją na coś”. Gwarancja może też jednak być gwarancją czegoś dla kogoś. Czyżby tym kimś był umysł nabywcy, a nie sam nabywca? Dlaczego zatem mowa o odrobinie umysłu? Semantyczne konotacje informacji przysłanej przez sklep można by rozwijać, standardowy odbiorca komunikatu nie traci jednak czasu na jego szczegółową analizę: rozumie, że wiadomość, którą otrzymał, przetłumaczona została z (niewiadomego) obcego języka przez automat (jakąś postać sztucznej inteligencji) i nie należy jej przyjmować dosłownie, nad jej na pierwszy rzut oka intrygującą treścią „nadpisany” został bowiem przekaz, że gdyby wąż w ciągu pięciu lat się zepsuł, sklep przyśle nowy, zapewne również objęty pięcioletnią gwarancją dla odrobiny umysłu. Wziąwszy pod uwagę konkurencyjną cenę przedmiotu, jego darmowy, błyskawiczny transport i oferowaną przez sklep gwarancję, trudno narzekać. Jeśli zaś ceną wszystkich tych udogodnień okazuje się (być może jedynie chwilowy) powrót języka do jego funkcji podstawowych i komunikacji bazowej, za której symbol w języku polskim służą choćby wypowiedzi Kalego z powieści *W pustyni i w puszczy*, to zapewne tak właśnie musi być<sup>1</sup>. W przyszłości powstaną znacznie lepsze automatyczne translatory i opisany problem, a w zasadzie pseudoproblem, po prostu zniknie. „Nowe

<sup>1</sup> Podobnie działająca bądź wręcz ta sama postać sztucznej inteligencji odpowiedzialna jest w sklepie za tłumaczenie opisów dostępnych w nim przedmiotów. Można tam na przykład znaleźć zagadkową informację o produkcie, ujętą w słowach „doniczki bawełniane do stroju ludowego”, zachętę dla kupującego: „Dodatkowe śliniaki wewnętrzne. Wysokiej jakości wykonanie, aby dodać

jutro” sztucznej inteligencji wykuwa się powoli. Dziś tymczasem, czy więcej warta jest – komu właściwie potrzebna – poprawna polszczyzna, czy też możliwość sprawnego i bezpiecznego nabycia niezbędnego w życiu codziennym przedmiotu?

W istocie obecny postęp w dziedzinie technologii cyfrowej (często utożsamianej ze sztuczną inteligencją), przekładający się na rosnącą wygodę życia codziennego, dla człowieka z lat sześćdziesiątych czy siedemdziesiątych ubiegłego stulecia nadal jeszcze byłby (zapewne niepokojącą) oznaką ziszczania się scenariuszy opisywanych w literaturze science fiction. Myśl, że właściwie każdy z nas mógłby posiadać na własność „telewizorek” z kolorowym ekranem, nie tylko pozwalający na oglądanie programu telewizyjnego czy filmów, ale umożliwiający dostęp do wszelkich informacji, które ludzkość zgromadziła przez tysiąclecia, a także – w dowolnym momencie – kontakt z wybraną osobą, wydawałaby się wówczas nierzeczywista. W krajach „realnego socjalizmu”, w których jedynej dostępnej informacji właściwie nie dawało się odróżnić od propagandy, a dostęp do faktycznej informacji okupiony był godzinami wsłuchiwania się w „pracę traktorów”, jak określano dźwięki emitowane w celu zagłuszenia sygnału zagranicznych stacji radiowych, perspektywa świata otwartego miała w sobie coś z rzeczywistości bajkowej, marzenia nigdy nieziszczalnego. Zakup węża prysznicowego łączyłby się w ówczesnych warunkach z wielotygodniowym korowodem zabiegów (nie bez powodu używano w takim kontekście słowa „polowanie”), już samo posiadanie prysznicza nie było powszechne, a nowo nabyty wąż mógł równie dobrze okazać się zepsuty bądź po prostu nie pasować do instalacji.

Uderzające jest w tym kontekście to, jak niewiele uwagi poświęcamy dziś samemu faktowi kulturowego przeobrażenia świata, które dokonało się w okresie ostatniego półwiecza, i że uznając je za oczywiste, akcentujemy przede wszystkim jego aspekt pragmatyczny, nie zaś aksjologiczny. Także wówczas, gdy przedmiotem rozważań są zagrożenia, przed jakimi może stanąć ludzkość wskutek utraty kontroli nad rozwojem sztucznej inteligencji, rozwój ów determinującej, myślimy przede wszystkim w kategoriach praktycznych, jeśli nawet ostatecznie będą to właśnie przeobrażenia kultury. Bliższe i łatwiejsze do postawienia są pytania o to, czy roboty przejmą nad nami kontrolę, czy na równi z nami staną się członkami społeczeństwa, czy w wyniku coraz nowszych rozwiązań technologicznych dojdzie do masowej utraty pracy przez ludzi, czy należy kierować edukacją w taki sposób, by do minimum ograniczyć kształcenie humanistyczne, a w jego miejsce wprowadzić nauczanie przedmiotów ścisłych, których znajomość okazuje się niezbędna w pracy informatyka i programisty. Zawody związane z uprawianiem humanistyki już dziś są powszechnie deprecjonowane i często uważane za fanaberię, a najwyraźniej trend ten zarysowuje się w społeczeństwach uboższych, takich jak choćby społeczeństwo polskie<sup>2</sup>. W jakimś sensie zachwianiu ulega równowaga treści i formy w kulturze – ta druga zyskuje i nieustannie ewoluuje: nierzadko odnosimy wrażenie, że wygląd strony internetowej jest istotniejszy od udostępnianych przez nią informacji, efekty wykorzystane w filmie wydają się ważniejsze niż jego akcja, liczba przypisów i odniesień w artykule w większym stopniu decyduje o naukowości tego tek-

---

wewnętrzne śliniaki pozwalają ulepszyć tę zwisającą sukienkę” czy wręcz polecenie „Wydrukuj zwierzę”, które pojawia się przy opisie sztuki odzieży ze wzorem w cętki (ang. animal print).

<sup>2</sup> Zob. np. Skuteczna Nauka. Studia Wyższe, *Masowa produkcja półgłówek z dyplomami skończonych studiów*, Skuteczna Nauka. Jak Uczyc się Mądrze i Skutecznie, <https://www.uczsieucz.pl/masowa-produkcja-polglowkow-z-dyplomami-skonczonych-studiow>.

stu niż oryginalność myśli jego autora, a zewnętrzna, „punktowa” ewaluacja placówki naukowej znaczy więcej niż jej faktyczne dokonania. Fakt, że nabywanie węży przyszniciowego zajmuje trzy minuty i nie wymaga kontaktu z drugim człowiekiem, sprawia, że nie przywiązujemy wagi do bezsensownej wiadomości, którą w dobrej wierze, poprzez automat, wysyła nam sklep. Algorytmizacja życia czyni je łatwiejszym w jego warstwie formalnej. Przykłady można mnożyć, a wszystkie one wskazują na postępowanie procesu „wyplukiwania” z kultury treści, które od zawsze stanowiły o jej charakterze, a zarazem wyznaczały dalszy kierunek jej rozwoju. Podczas bowiem gdy algorytm nie przewiduje oryginalności, doświadczenie ludzkie jest zawsze jednostkowe i niepowtarzalne, zawsze angażuje świadomość i wolność, a więc również sprawczość (ale też pasywność), a jego rezultat pozostaje nieprzewidywalny. Zawsze jest ono także doświadczeniem siebie<sup>3</sup>.

W takim właśnie duchu mówi o oddziaływaniu sztucznej inteligencji na nasz świat John C. Lennox<sup>4</sup>. Twierdzi on, że podejmując kwestię kulturowych czy wręcz aksjologicznych implikacji rozwoju technologicznego, niejako źle rozkładamy akcenty. Sam termin „sztuczna inteligencja” jest mylący – chociaż w istocie mamy tu do czynienia z czynnikiem o charakterze sztucznym, nie jest nim jednak inteligencja, a jedynie symulacja inteligencji, polegająca na wykorzystaniu jako podstawy czy też podłoża aktywności określanej tym mianem substancji nieorganicznej (na przykład silikonu), która o tyle przewyższa naturalną, organiczną podstawę inteligencji, że nie ulega degeneracji i nie podlega śmierci. W przypadku człowieka inteligencja łączy się jednak ze świadomością – czynnikiem o znaczeniu podstawowym, którego ani filozofom, ani neurobiologom, ani

<sup>3</sup> „Przedmiotem doświadczenia jest nie tylko moment, ale i człowiek, który wyłania się ze wszystkich momentów, a równocześnie tkwi w każdym z nich [...]. I nie można powiedzieć, że doświadczenie jest sobą tylko w owym momencie, a dalej pozostaje już tylko praca umysłu «kształtujuącego» człowieka jak swój przedmiot na podstawie jednostkowego momentu empirycznego, czy też szeregu takich momentów. [...] Ostatecznie zrozumienie siebie samego składa się z wielu rozumień, poniekąd tak, jak doświadczenie składa się z wielu doświadczeń. Wydaje się, iż każde doświadczenie jest zarazem jakimś rozumieniem”. K. W o j t y ł a, *Osoba i czyn*, Polskie Towarzystwo Teologiczne, Kraków 1969, s. 6.

Intuicja którą wyraża Wojtyła, pisząc o kolejnych „rozumieniach” siebie przez podmiot, powraca w dziełach gatunku literackiego znanego jako Bildungsroman. „Na świecie było tyle dżum co wojen. Mimo to dżumy i wojny zastają ludzi zawsze tak samo zaskoczonych” (A. C a m u s, *Dżuma*, tłum. J. Guze, PIW, Warszawa 1966, s. 41) – mówi narrator powieści Alberta Camusa, a postać i przemyslenia doktora Rieux, ukazują, że żadne doświadczenie nie stanowi „domknięcia” bytu, jakim jest osoba ludzka. Niepokorny architekt Howard Roark z powieści *Źródło* Ayn Rand wskazuje z kolei na bezsens postawy, którą można by określić jako algorytmizację życia: „Proszę spojrzeć. [...] Te osławione nacięcia na kolumnach: czemu służą? Kiedy robiono kolumny z drewna, służyły ukryciu spoin. Te są marmurowe. A te tryglify, z czego je zrobiono? Z drewna. Drewniane belki, takie, jakie kładziono, gdy ludzie uczyli się budować chatki z drewna. [...] Grecy brali marmur i kopiowali te drewniane struktury, bo inni przed nimi tak robili. Po nich przyszli [...] renesansowi bogowie i robili gipsowe kopie tamtych marmurowych kopii tamtych drewnianych kopii. A teraz my – co robimy? Stałowe i metalowe kopie gipsowych kopii marmurowych kopii drewnianych kopii. Dlaczego?” (A. R a n d, *Źródło*, tłum. I. Michałowska, Zysk i S-ka, Poznań 2002, s. 29).

<sup>4</sup> Zob. John Lennox *Unlocks the Truth about AI, Consciousness, and God* (Z Johnem C. Lennoxem rozmawia Samuel Marusca), Practical Wisdom, YouTube, <https://www.youtube.com/watch?v=Undu9YI3Gd8>. Zob. też: J.C. L e n n o x, *2084: Artificial Intelligence and the Future of Humanity*, Zondervan Reflective, Grand Rapids, Michigan, 2020, EPUB.

żadnym innym naukowcom nie udało się do tej pory zdefiniować w sposób satysfakcjonujący. Wpisując się w pragmatyczny nurt kultury, wyznaczony właśnie dynamiką jej rozwoju technologicznego, twórcy sztucznej „inteligencji” faktycznie zarzucili ideę pracy nad stworzeniem sztucznej świadomości i ograniczają swoje wysiłki do symulacji inteligencji. Jeśli bowiem wziąć pod uwagę sam aspekt użyteczności, świadomość nie okazuje się niezbędna, skoro i bez niej udaje się osiągać cele, których realizacja normalnie wymagałaby udziału świadomej osoby. Tworzone dziś tak zwane maszyny myślące, mimo że stanowią istotne osiągnięcie technologiczne, nie są zatem bytami świadomymi: między człowiekiem a robotami istnieje nieprzekraczalna przepaść i nie sposób mówić o jakimkolwiek „iloczynnie” ich kompetencji. Skoro zaś nikomu, jak dotąd, nie udało się dokonać symulacji świadomości, mowa o inteligentnych maszynach, często powracająca w dyskursach popularnych i pseudonaukowych<sup>5</sup>, jest nonsensem<sup>6</sup>. Można zaryzykować stwierdzenie, że tak jak dla dawnej nauki wielkie wyzwanie stanowił problem ruchu, tak dla nauki współczesnej wyzwaniem tego rodzaju jest właśnie problem świadomości. Być może najlepiej ujął go Roger Penrose, podchodząc doń od strony mechaniki kwantowej i twierdząc, że jedynie wówczas zrozumiemy, czym jest świadomość, gdy będziemy potrafili ją włączać i wyłączać<sup>7</sup>. Już John Searle, wysuwając argument pokoju chińskiego, udowodnił, że świadomość można skutecznie symulować<sup>8</sup>, i tym samym obalił tak zwany test Turinga, mówiący, że jeśli komputer potrafi przekonać osobę wchodzącą z nim w interakcję o tym, że posiada świadomość, to znaczy, że faktycznie świadomość posiada<sup>9</sup>.

Nie tylko kwestia świadomości ma decydujące znaczenie dla pojmowania sensu sztucznej inteligencji. Jej dotychczasowe zastosowanie – twierdzi Lennox – opiera się na wykorzystywaniu algorytmów, czyli stosowaniu wysoce złożonych zestawów reguł w odniesieniu do dużej liczby danych, co pozwala prowadzić wielorakie analizy służące najróżniejszym celom. Sztuczna inteligencja nie jest jednak w tym sensie nawet symulacją pracy mózgu, umysł ludzki nie podlega bowiem – jak uważa również Penrose – algorytmizacji<sup>10</sup>, a sam ten fakt stanowi z jednej strony w jakimś sensie dowód na słuszność twierdzenia Kurta Gödla. Z drugiej – można dodać – potwierdza on również antropologicznofilozoficzną tezę personalistów o niepowtarzalności i wyjątkowości osoby

<sup>5</sup> Zob. np. T. D u d a, *Inteligentne maszyny odbiorą Ci życie. Wystarczy, że nie zrobisz nic*, benchmark.pl., [https://www.benchmark.pl/testy\\_i\\_recenzje/sztuczna-inteligencja-sprawi-ze-ludzie-stana-sie-niepotrzebni.html](https://www.benchmark.pl/testy_i_recenzje/sztuczna-inteligencja-sprawi-ze-ludzie-stana-sie-niepotrzebni.html).

<sup>6</sup> Jeden z inżynierów zaangażowanych w tworzenie sztucznej inteligencji ogłosił jednak w roku 2022, że możliwe jest, iż duże sieci neuronowe wykazują pewien stopień świadomości (zob. I. S u t k e v e r, Post, X, <https://twitter.com/ilyasut/status/1491554478243258368?lang=en>).

<sup>7</sup> Zob. Roger Penrose: „*Consciousness must be beyond computable physics*” (Z Rogerem Penrose'em rozmawia Michael Brooks), New Scientist, <https://www.newscientist.com/article/mg25634130-100-roger-penrose-consciousness-must-be-beyond-computable-physics/>. Zob. też: L e n n o x, dz. cyt., rozdz. 7, „Artificial General Intelligence: The Future Is Dark?”.

<sup>8</sup> Zob. J. S e a r l e, *Minds, Brains, and Programs*, „Behavioral and Brain Sciences” 3(1980) nr 3, s. 417-424. Zob. też: L e n n o x, dz. cyt., rozdz. 7.

<sup>9</sup> Zob. A. T u r i n g, *Computing Machinery and Intelligence*, „Mind” 59(1950) nr 236, s. 433-460.

<sup>10</sup> Zob. Roger Penrose: „*Consciousness must be beyond computable physics*”.

ludzkiej<sup>11</sup>. Żaden algorytm nie będzie w stanie rozwiązać problemu świadomości, a tym samym problemu człowieka, gdyż umysł ludzki algorytmom się po prostu wymyka: komputery i napędzane ich pracą urządzenia działają w zupełnie inny sposób niż on.

Lennox zwraca uwagę, jak bardzo mylący w dyskusjach nad istotą i przyszłością sztucznej inteligencji okazuje się antropomorfizujący język, którym się posługujemy<sup>12</sup>. Mowa o tworzących ją sieciach „neuronowych”, o funkcjonujących niczym synapsy połączeniach między węzłami (ang. nodes) tych sieci czy o zdolności maszyn do „uczenia się” głębokiego (ang. deep learning) sprawia, że zatracamy pojęciowy dystans między rozumieniem umysłu a rozumieniem matematycznego konstruktów, którym są „myślące” maszyny i który rozumem umysłu nie przypomina, mimo że w jakimś stopniu symuluje jego aktywność. Dobrym przykładem takiego konstruktów jest Chat GPT, w swojej najnowszej wersji potrafiący generować wypowiedzi na dowolny temat przypominające wypowiedzi człowieka. Analizując sposób funkcjonowania tej maszyny, Noam Chomsky, Ian Roberts i Jeffrey Watumull wskazują jednak, że uczy się ona języka zupełnie inaczej niż człowiek. „W przeciwieństwie do Chatu GPT i podobnych mu maszyn – piszą – umysł ludzki nie jest ociążoną maszyną służącą do obliczeń statystycznych, poszukującą zbieżności wzorców, kompilującą setki terabajtów danych i dokonującą wyboru najbardziej prawdopodobnej odpowiedzi w rozmowie czy też najbardziej prawdopodobnej odpowiedzi na pytanie o charakterze naukowym. Przeciwnie, umysł ludzki stanowi zdumiewająco wydajny, a nawet cechujący się pewną elegancją system, który potrafi funkcjonować, dysponując nawet niewielką liczbą informacji; jego praca nie polega na ujmowaniu uproszczonych korelacji między zbiorami danych, lecz na poszukiwaniu wyjaśnień. Małe dziecko, na przykład, ucząc się języka, mimo że dysponuje jedynie bardzo niewielkim zasobem danych, nieświadomie, automatycznie i w szybkim tempie rozwija w sobie jego gramatykę: niesłychanie złożony system zasad i parametrów logicznych. Gramatykę można w tym sensie pojmować jako wyraz przyrodzonego, genetycznie wmontowanego «systemu operacyjnego», który zapewnia istotom ludzkim możliwość generowania złożonych zdań i długich ciągów myślowych. Starając się wypracować teorię tłumaczącą, dlaczego dany język funkcjonuje w określony sposób («Dlaczego akurat te – a nie inne – zdania uważane są za poprawne w sensie gramatycznym?»), lingwiści świadomie, podejmując żmudny wysiłek, krok po kroku odtwarzają model gramatyki, który dziecko instynktownie buduje w sobie przy minimalnej ekspozycji na informację. System operacyjny dziecka zdecydowanie różni się od systemu operacyjnego maszyny, która uczy się programu”<sup>13</sup>. Badacze podkreślają, że słabym punktem

<sup>11</sup> Zob. np. R. Spemann, *Dlaczego osoby nazywamy osobami*, w: tenże, *Osoby. O różnicy między czymś a kimś*, tłum. J. Merecki, Oficyna Naukowa, Warszawa 2001, s. 23-43; tenże, *O identyfikowaniu osób*, w: *Osoby. O różnicy między czymś a kimś*, s. 44-51; T. Styczeń, *Osoba – podmiot we wspólnocie*, w: tenże, *Dzieła zebrane*, red. A.M. Wierzbicki, t. 3, *Objawiać osobę*, red. A.M. Wierzbicki, Towarzystwo Naukowe KUL – Instytut Jana Pawła II KUL, Lublin 2013, s. 41-68.

<sup>12</sup> Na temat języka jako narzędzia umożliwiającego, a jednocześnie zaciemniającego poznanie rzeczywistości zob. W. Egginton, *Borges and Heisenberg Converged on the Slipperiness of Language*, „Aeon”, aeon.co, <https://aeon.co/essays/borges-and-heisenberg-converged-on-the-slipperiness-of-language>.

<sup>13</sup> N. Chomsky, I. Roberts, J. Watumull, *The False Promise of Chat GPT*, „The New York Times” z 8 III 2023, <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>. O ile nie wskazano inaczej, tłumaczenie fragmentów obcojęzycznych – D.Ch.



tak zwanej sztucznej inteligencji jest jej brak zdolności mającej w przypadku inteligencji znaczenie kluczowe, a mianowicie zdolności do stwierdzenia, co nie jest faktem, dlaczego nim nie jest i dlaczego nigdy nim nie będzie. „Myślące” maszyny, mimo że kompilują ogromne ilości danych, pozbawione są zdolności do wyjaśniania zjawisk, będącej oznaką rzeczywistej inteligencji – i właśnie dlatego są w stanie wygenerować informację o gwarancji dla odrobiny umysłu. Również błędy popełniane przez faktyczną inteligencję mają całkowicie odmienną naturę i strukturę od błędów generowanych przez inteligencję jedynie symulowaną: cechujące ludzi myślenie twórcze zasadza się na stawianiu i obalaniu hipotez, na procesie intelektualnym, który – właśnie poprzez wykrywanie i eliminację potknięć intelektualnych – ostatecznie skutkuje poruczeniem tych dróg poznania, które nie prowadzą do prawdy. W przeciwieństwie do człowieka maszyny nie potrafią odróżnić tego, co możliwe, od tego, co niemożliwe: można je „nauczyć” zarówno tego, że Ziemia jest okrągła, jak i tego, że jest ona płaska<sup>14</sup>. I jeśli się mylą, to przede wszystkim właśnie dlatego, że nie rozumieją danych, którymi operują. Chomsky i jego współpracownicy – nie odbiegając w tej kwestii od stanowiska personalistów – zauważają też, że cechą prawdziwej inteligencji jest umiejętność myślenia w kategoriach moralnych, równoznaczna z wyznaczaniem granic naszej kreatywności przez zasady etyczne określające treść powinności moralnej, jakkolwiek stanowiące przy tym również przedmiot namysłu. Tymczasem, pytany o kwestię jego własnej perspektywy spojrzenia, Chat GPT odpowiada: „Nie mam świadomości, samoświadomości ani zdolności do posiadania osobistej perspektywy. Mogę dostarczać informacji i analiz w oparciu o dane, na których się uczę, ale nie potrafię formułować własnych poglądów ani przekonań. Chociaż potrafię rozumieć i interpretować język, nie mam subiektywnego doświadczenia właściwego ludziom”<sup>15</sup>.

Z tych wszystkich względów działające w oparciu o algorytmy sztucznej inteligencji maszyny mogą też jednak stać się potężnym narzędziem szerszenia dezinformacji – w tym sensie ich praca zależy od tego, jakimi danymi operują. Fakt, że pozbawione są subiektywnego doświadczenia oraz umiejętności myślenia, pozwala człowiekowi zaprogramować je – poprzez dobór bazy danych, którymi operują – w wybrany przezeń sposób. W tym kontekście inicjator pracy nad sztuczną inteligencją Elon Musk wskazuje, że charakterystyczne dla współczesnego Zachodu zdominowanie przestrzeni publicznej – świata kultury, mediów, polityki oraz korporacji przemysłowych – przez poglądy skrajnie lewicowe może skutkować sytuacją, kiedy kontrola nad programowaniem sztucznej inteligencji znajdzie się w rękach przedstawicieli tego rodzaju prądów umysłowych, jak choćby zwolenników działającego od roku 1991 proekologicznego Ruchu na rzecz Dobrowolnego Wyginięcia Ludzkości (Voluntary Human Extinction Movement), postrzegających ludzkość i cywilizację jako coś w rodzaju zarazy szerzącej się na powierzchni Ziemi i wymagającej likwidacji<sup>16</sup>. Musk twierdzi, że tego rodzaju przejmowanie kontroli

<sup>14</sup> Zob. tamże.

<sup>15</sup> Cyt za: Chomsky, Roberts, Watmull, dz. cyt.

<sup>16</sup> Zob. C. Buckley, *Earth Has Now 8 Billion Humans. This Man Wishes There Were None*, „The New York Times” z 23 XI 2022, The New York Times, <https://www.nytimes.com/2022/11/23/climate/voluntary-human-extinction.html>. Artykuł zawiera obszernie cytaty z wywiadu przeprowadzonego przez Carę Buckley, dziennikarkę zajmującą się problematyką zmian klimatycznych, z Lesem Knightem, założycielem Ruchu na rzecz Dobrowolnego Wyginięcia Ludzkości. Tekst ukazał się na pierwszej stronie gazety.

nad wykorzystaniem sztucznej inteligencji ma już precedens: media społecznościowe w Stanach Zjednoczonych – wskutek wykorzystywania przez ich platformy określonych algorytmów w celu eliminacji pewnych wypowiedzi czy też wypowiedzi na pewne tematy – stały się w ostatnim okresie ramieniem rządu i tubą rządowej propagandy, przyjmując funkcje podobne do tych, które w Związku Radzieckim pełniła gazeta „Prawda”. Problem leży w tym, że podczas gdy w normalnych warunkach ideologie ekstremalne mają zazwyczaj znikomy zasięg terytorialny, dysponowanie sztuczną inteligencją, w której uruchomione zostają wybrane algorytmy, umożliwia propagowanie tego rodzaju poglądów na całym świecie przy jednoczesnym eliminowaniu możliwości wyrażania stanowisk im przeciwnych<sup>17</sup>. Z kolei technika rozpoznawania tożsamości osoby na podstawie elektronicznego wizerunku jej twarzy (a nawet jedynie jej postury) może stać się sprawnym narzędziem wyszukiwania i eliminacji ze społeczeństwa jednostek próbujących w jakiś sposób stanowiska te mimo wszystko głosić<sup>18</sup>. Podobny problem dotyczy zmiany, jaką wszechobecność sztucznej inteligencji wprowadza do struktury współczesnego świata twórczości literackiej i odbioru literatury: w dobie poprawności politycznej i kultury unieważniania (ang. cancel culture), z którymi łatwo „sprząć” odpowiednie algorytmy wyszukiwania, fikcja literacka przestaje pełnić funkcję narzędzia moralnego namysłu nad postawami ludzkimi czy eksplorowania motywacji bohaterów. Współczesna krytyka literacka oczekuje raczej od autorów ukazywania moralności „matematycznie precyzyjnej”<sup>19</sup> – oczywiście, niepodważalnej, wieczystej; moralności właściwej dla świata epoki powszechnej cyfryzacji<sup>20</sup>.

Zagrożenie związane z możliwością wprowadzania w ten sposób globalnej inżynierii społecznej, a zarazem nowej odsłony totalitaryzmu<sup>21</sup>, sprawia, że zarówno Lennox, jak i Musk podkreślają, iż technologia nie powinna rozwijać się szybciej niż jej etyczna podbudowa. Tworzenie tego rodzaju fundamentu moralnego nie jest jednak w tym przypadku łatwe. Oczywiście, zdroworoządkowe hasła przywołujące imperatyw równości, dobra wspólnego czy też dobra człowieka, jak te wypracowane podczas zorganizowanej w Asilomar konferencji na temat etyki wykorzystania sztucznej inteligencji<sup>22</sup>, niewiele dziś znaczą w świecie, w którym wszystkie tego rodzaju pojęcia mają wiele różnych interpretacji i w którym ogólnie respektowane wytyczne działania moralnie dobrego nie istnieją. Musk otwarcie zastanawia się, czy w warunkach powszechnego wykorzystywania sztucznej inteligencji system demokratyczny będzie w ogóle możliwy do utrzymania<sup>23</sup>. Lennox twierdzi z kolei, że – ze względu na istnienie w dzisiejszym świecie licznych,

<sup>17</sup> Zob. *Elon Musk on Buying Twitter and Turning It into X* (Z Elonem Muskiem rozmawia Joe Rogan), YouTube, <https://www.youtube.com/watch?v=N8Nf56srwca>.

<sup>18</sup> Zob. *John Lennox Unlocks the Truth about AI, Consciousness, and God*. Zob. też: L e n n o x, dz. cyt., rozdz. 1, „Mapping out the Territory”.

<sup>19</sup> W. B l y t h e, *The Life, Death – And Afterlife – of Literary Fiction*, „Esquire” z 13 VII 2023, Esquire, <https://www.esquire.com/entertainment/books/a44496450/literary-fiction-death-digital-age/>.

<sup>20</sup> Zob. tamże.

<sup>21</sup> Zob. L e n n o x, dz. cyt.; zob. też: S. Z u b o f f, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, Profile, London 2019.

<sup>22</sup> Zob. Future of Life Institute, *Asilomar AI Principles*, <https://futureoflife.org/person/asilomar-ai-principles/>.

<sup>23</sup> Zob. *Elon Musk’s Brutally Honest Interview with Tucker Carlson* (Z Elonem Muskiem rozmawia Tucker Carlson), Smart Sense, YouTube, [https://www.youtube.com/watch?v=zaB\\_20bko-A4&ab\\_channel=SmartSense](https://www.youtube.com/watch?v=zaB_20bko-A4&ab_channel=SmartSense).



nierównoważnych sobie centrów decyzyjnych – ratunkiem nie może być dominujący obecnie światopogląd utylitarystyczny. Jedyne godne zaufania źródło etycznego fundamentu dla rozwoju technologicznego stanowić może w jego przekonaniu chrześcijańska wersja teizmu ze względu na wpisany w nią bezwarunkowy szacunek dla każdej istoty ludzkiej, którego ostatecznym gwarantem jest Bóg<sup>24</sup>.

Źródłem swoistej „gwarancji dla odrobiny umysłu” może też jednak okazać się we współczesnym świecie myślenie zdroworozsądkowe: ostatecznie rzecz biorąc, każdą technologię i każde narzędzie, które stworzył człowiek, można wykorzystać zarówno dla jego dobra, jak i w celu jego destrukcji, a nasza cywilizacja może przeminąć, tak jak przeminęły inne, bez względu na to, czy (bądź w jaki sposób) wykorzystywać będzie sztuczną inteligencję.

---

<sup>24</sup> Zob. *John Lennox Unlocks the Truth about AI, Consciousness, and God*.

## NOTY O AUTORACH

Małgorzata A b a s s y, doktor habilitowany, profesor UJ, kulturoznawca, literaturoznawca, rusycystka i iranistka. Studia z zakresu literaturoznawstwa, filologii orientalnej i filologii rosyjskiej na Uniwersytecie Jagiellońskim.

Od 2007 r. zatrudniona na UJ, obecnie profesor uczelni w Katedrze Kultury Rosyjskiej i Teorii Systemów Kulturowych Instytutu Rosji i Europy Wschodniej na Wydziale Studiów Międzynarodowych i Politycznych.

Członek Association for Slavic, East European and Eurasian Studies (ASEEES), Societas Iranologica Europea (S.I.E.), Komisji Kultury Słowian PAU, Polskiego Towarzystwa Kulturoznawczego.

Główne obszary zainteresowań badawczych: inteligencja i elity intelektualne, kulturowe uwarunkowania odmienności Rosji i krajów islamskich od Zachodu, stosunki rosyjsko-irańskie i polsko-rosyjskie na fundamencie kultury, literatura piękna i publicystyka jako narzędzia kształtowania świadomości narodowej oraz źródło wiedzy o historii i mentalności narodów, cybernetyka jako „myśl porządkująca kulturę”.

Najważniejsze publikacje książkowe: *Inteligencja a kultura. O problemach samoidentyfikacji dziewiętnastowiecznej inteligencji rosyjskiej* (2008); *Irańska inteligencja w XIX wieku i Rewolucja Konstytucyjna (1905-1911)* (2010); *Kultura wobec postępu i modernizacji. Rosja i Iran w perspektywie porównawczej* (2013); *Russian Mason on the Paths of his Native Culture. The Case Study of Nicolas Novikov* (2014); *Modernizacja w cieniu Allaha. Współczesny Iran* (2019).

Andrzej B o r k o w s k i, doktor habilitowany, profesor AP, literaturoznawca, językoznawca. Studia z zakresu filologii polskiej w Akademii Podlaskiej, z zakresu literaturoznawstwa na Uniwersytecie Warmińsko-Mazurskim (doktorat).

Od 2004 zatrudniony w Instytucie Filologii Polskiej Akademii Podlaskiej (dziś Uniwersytetu w Siedlcach), w latach 2006-2011 wicedyrektor tegoż instytutu, obecnie profesor uczelni w Instytucie Językoznawstwa i Literaturoznawstwa na Wydziale Nauk Humanistycznych (2018-2020 – dyrektor).

Członek Niemieckiego Towarzystwa Komeńskiego (Deutsche Comenius-Gesellschaft), Instytutu Kultury Regionalnej i Badań Literackich im. Franciszka Karpińskiego (od 2013 r. prezes), Siedleckiego Towarzystwa Naukowego, Towarzystwa Kultury Języka Oddziału Siedleckiego, Towarzystwa Literackiego im. Adama Mickiewicza (Oddział Siedlecki – w latach 2006-2007 sekretarz).

Główne obszary zainteresowań badawczych: literatura polska i powszechna głównie XVI, XVII i XX w. oraz najnowsza (zwłaszcza literatury słowiańskie) w perspektywie komparatystycznej i hermeneutycznej, glottodydaktyka.

Najważniejsze publikacje książkowe: *Imaginarium symboliczne Wacława Potockiego: „Ogród nie plewiony”* (2011); *Pejzaż współczesny z barokiem w tle. Szkice o literaturze i kulturze* (2012); *Labirynty dyskursów w sławiańskich literaturach epoki barokko. Religia – polityka – obywatelstwo* (2015); *Poezja polska dla cudzoziemców z ćwiczeniami. Skrypta* (2015); *„A ja jestem córką czarownicy”*. *Żywioł autportretowy w twórczości Kazimierza Illakowiczówny* (2018).

Paweł F o r t u n a, doktor, psycholog. Studia z zakresu psychologii na Katolickim Uniwersytecie Lubelskim Jana Pawła II.

Od 1995 r. zatrudniony na KUL, obecnie adiunkt w Katedrze Psychologii Eksperymentalnej Instytutu Psychologii na Wydziale Nauk Społecznych.

Główne obszary zainteresowań badawczych: antropologia potoczna, psychologia pozytywna, cyberpsychologia pozytywna.

Najważniejsze publikacje książkowe: *Obrona przed wpływem telewizji* (2007); *Jak być lepszym szefem. Inspiracje psychologiczne* (współautor, 2009); *Przygotowanie szkolenia, czyli jak dobry początek prowadzi do sukcesu* (współautor, 2009); *Studium przypadku w praktyce szkoleniowej, czyli jak uczyć się na doświadczeniach innych* (2010); *Perswazja w pracy trenera, czyli jak kształtować postawy uczestników szkolenia* (2011); *Pozytywna psychologia porażki. Jak z cytryn zrobić lemoniadę* (2012); *Metafory i analogie w szkoleniach* (współautor, 2013); *Subiektywna psychologia biznesu* (2014); *Animal rationale. Mądrość zwierząt w biznesie, edukacji i rodzinie* (współautor, 2015); *Filmowy leksykon psychologii* (współautor, 2017); *Sprzedaż bez sprzedawania. Psychologia dobrego wpływu na klienta* (2018); *Optimum. Idea cyberpsychologii pozytywnej* (2021).

Izabela K a c z m a r z y k, doktor habilitowany, profesor UIK, literaturoznawca, kulturoznawca.

Od 2008 r. pracownik Uniwersytetu Ignatianum w Krakowie, obecnie profesor uczelni w Katedrze Dziedzictwa Kulturowego Instytutu Kulturoznawstwa i Dziennikarstwa na Wydziale Filozoficznym.

Główne obszary zainteresowań badawczych: dziedzictwo kulturowe Śląska, w szczególności dawne kaznodziejstwo ewangelickie, górnośląskie dziedzictwo industrialne od końca XVIII do przełomu XIX i XX w., szkolna kultura pamięci.

Najważniejsze publikacje książkowe: *Adam Gdaczusz. Z dziejów kaznodziejstwa śląskiego* (2003); *W cieniu mitu. Rzecz o Karolu Goduli* (2007); *Krajobraz, portret, maska. Studia o górnośląskim imaginariu przemysłowym* (2019).

Barbara K l o n o w s k a, doktor habilitowany, profesor KUL, literaturoznawca. Studia z zakresu filologii angielskiej na Katolickim Uniwersytecie Lubelskim.

Od 1995 r. zatrudniona na KUL, obecnie profesor uczelni w Katedrze Literatury i Kultury Angielskiej Instytutu Filologii Angielskiej na Wydziale Nauk Humanistycznych.

Członek Utopian Studies Society – Europe (USSE), Polskiego Towarzystwa Badań nad Australią i Nową Zelandią, Stowarzyszenia Nauczycieli Akademickich Filologii Angielskiej (PASE).

Główne obszary zainteresowań badawczych: współczesna proza brytyjska, teoria literatury, narracje utopijne i dystopijne w literaturze i filmie.

Najważniejsze publikacje książkowe: *Contaminations: Magic Realism in Contemporary British Fiction* (2006); *Longing for Romance: British Historical Romances 1990-2010* (2014); *Margins and Centres Reconsidered* (współred., 2008); *Echoes of Utopia* (współred., 2012); *(Im)perfection Subverted, Reloaded and Networked: Utopian Discourse across Media* (współred., 2015). Współredaktor tomu „Roczników Humanistycznych” *Brave New Human in (Trans/Post)humanist Utopias* (66(2018) nr 11).

Barłomiej K n o s a l a, doktor, filozof. Studia z zakresu filozofii na Uniwersytecie Śląskim.

Od 2010 r. zatrudniony na Politechnice Śląskiej, obecnie adiunkt w Katedrze Stosowanych Nauk Społecznych na Wydziale Organizacji i Zarządzania.

Sekretarz Polskiego Towarzystwa Oceny Technologii. Członek Śląskiego Centrum Etyki Biznesu i Zrównoważonego Rozwoju.

Główne obszary zainteresowań badawczych: filozofia techniki, społeczny wpływ technologii kognitywnych, koncepcja postnatury, zmiany klimatyczne, ekologie integralne.

Autor książki *Projekt nauki nowej Marshalla McLuhana. Filozoficzne konsekwencje zmian form komunikacji* (2017).

Zbigniew N a w r a t, doktor habilitowany, profesor IPS FRK, fizyk, specjalista w zakresie robotyki medycznej, wynalazca. Studia z zakresu fizyki teoretycznej na Uniwersytecie Śląskim, doktorat na temat sztucznego serca w Śląskiej Akademii Medycznej, habilitacja na temat robota chirurgicznego w Śląskim Uniwersytecie Medycznym (SUM).

Od 1984 r. zatrudniony w SUM (Katedra Biofizyki), od 1988 r. w Katedrze Kardiochirurgii prowadzonej przez prof. Zbigniewa Religę. Od 2018 r. pracownik Katedry Biofizyki SUM. Współtwórca projektu polskiego sztucznego serca i robota chirurgicznego Robin Heart.

Współtwórca Instytutu Protez Serca (IPS) Fundacji Rozwoju Kardiochirurgii im. Prof. Zbigniewa Religi (FRK) (od 2007 r. do 2019 r. dyrektor IPS FRK). Obecnie dyrektor kreatywny FRK.

Żołęźyciel i prezydent Międzynarodowego Stowarzyszenia na Rzecz Robotyki Medycznej. Członek European Society for Artificial Organs, Komisji Biomechaniki Komitetu Mechaniki PAN, Komisji Nauk Przyrodniczych, Stacji Naukowej Polskiej Akademii Umiejętności w Katowicach, Rady Naukowej projektu fundacji SANO, Fundacji AI LAW TECH, Polskiego Stowarzyszenia „Sztuczna Inteligencja w Medycynie”. Redaktor czasopisma „Medical Robotics Reports”.

Główny obszary zainteresowań badawczych: robotyka medyczna, sztuczne narządy (głównie protezy serca, zastawki i pompy krwi), technologia membranowa, symulacja komputerowa w planowaniu zabiegów chirurgicznych, telemedycyna, sztuczna inteligencja w medycynie.

Najważniejsze publikacje książkowe: *Postępy technologii biomedycznych/Advances in biomedical technology* (2007); *Roboty Medyczne /Medical Robots* (2007); *Postępy technologii biomedycznych 2/Advances in Biomedical Technology 2.* (2008); *E-enjoy ICT Quality Book* (2010); *Implant Expert* (2011); *Postępy robotyki medycznej. Advances in Medical Robotics* (współred., 2013).

Grzegorz T r ę b i c k i, doktor habilitowany, profesor UJK, literaturoznawca. Studia z zakresu filologii angielskiej na Uniwersytecie Marii Curie-Skłodowskiej i na Uniwersytecie Gdańskim (doktorat).

Od 2002 r. zatrudniony na Uniwersytecie Jana Kochanowskiego w Kielcach, obecnie profesor uczelni w Zakładzie Literatur Obcych Instytutu Literaturoznawstwa i Językoznawstwa na Wydziale Humanistycznym.

Główny obszary zainteresowań badawczych: anglosaska literatura niemimetyczna (szczególnie literatura fantasy), literatura *science fiction*, literatura młodzieżowa, teoria i geneologia literatury.

Najważniejsze publikacje książkowe: *Fantasy. Ewolucja gatunku* (2007); *Synkretizm fantasy. Fantasy świata wtórnego: literatura, kultura, mit* (2014); *Worlds So Strange and Diverse: Towards a Genological Taxonomy of Non-mimetic Literature* (2015); *Świat wartości. Aksjologia fantasy świata wtórnego. Model podstawowy* (2020).

Barbara T r y g a r, doktor, literaturoznawca.

Główny obszary zainteresowań badawczych: literatura i filozofia, zagadnienia aksjologiczne w literaturze.

Najważniejsze publikacje książkowe: *Doświadczenie, przeżycie, kontemplacja. Pisarstwo Kazimierza Brauna* (współred., 2017); „Ja” w drodze do Wolności, Prawdy i Piękna. Aspekty aksjologiczne twórczości prozatorskiej Kazimierza Brauna (2019).

Zbigniew W r ó b l e w s k i, doktor habilitowany, profesor KUL, filozof. Studia z zakresu filozofii przyrody na Katolickim Uniwersytecie Lubelskim.

Od 1992 r. pracownik KUL, obecnie (od 2012 r.) profesor uczelni, kierownik Katedry Filozofii Przyrody i Nauk Przyrodniczych Instytutu Filozofii na Wydziale Filozofii.

Członek Towarzystwa Naukowego KUL, Polskiego Towarzystwa Filozoficznego (od 2004 r. członek zarządu Oddziału Lubelskiego), Sekcji Filozofii Przyrody i Nauk Przyrodniczych PTF.

Główne obszary zainteresowań badawczych: filozofia przyrody ożywionej, etyka ekologiczna, kognitywistyka.

Najważniejsze publikacje książkowe: *Katechezy ekologiczne* (współautor, 1993); *Wokół mnożnika cztery* (współred., 2000); *Ekologia a transformacje cywilizacyjne na przełomie wieków* (współred., 2000); *Natura i norma. Kontrowersje filozoficzne* (red., 2010); *Natura i cele. Dyskusja argumentu teleologicznego na rzecz ochrony przyrody* (2010); *Filozofować eksperymentalnie. Umysł w kulturze* (współred., 2018); *Lubelska szkoła filozoficzna w porównaniach* (współred., 2019).

Władysław Z u z i a k, profesor, filozof. Studia z zakresu teologii na Papieskim Wydziale Teologicznym w Krakowie, z zakresu filozofii na Katolickim Uniwersytecie w Louvain-la-Neuve.

Od 1990 r. zatrudniony na Wydziale Filozoficznym PAT (od 2009 r. Uniwersytet Papieski Jana Pawła II), obecnie profesor w Katedrze Etyki (w latach 2002-2023 kierownik).

Członek Rady Instytutu Dziedzictwa Solidarności (w latach 2018-2023 dyrektor Międzynarodowego Centrum Badań nad Fenomenem Solidarności – MCBFS), Polskiego Towarzystwa Filozoficznego (Oddział w Krakowie), Polskiego Towarzystwa Etycznego, Sekcji Problemów Historyczno-Społecznych Związku Podhalan w Polsce im. św. Jana Pawła II, European Society for the Study of Science and Theology (ESSAT), Rady Programowej Centrum Etyki Chrześcijańskiej im. Tadeusza Ślipko SJ, Rady Naukowej serii wydawniczej „Polska Filozofia Chrześcijańska XX wieku”, Rady Programowej czasopisma „Logos i Ethos”, rad naukowych czasopism „The Person and the Challenges”, „Episteme”, „Świdnickie Studia Teologiczne”.

Główne obszary zainteresowań badawczych: filozofia człowieka, etyka, filozofia społeczna (w tym problematyka związana z ideą solidarności), aksjologia, personalizm, ekologia jako filozofia odpowiedzialności.

Najważniejsze publikacje książkowe: *Dojrzewanie świadomości refleksyjnej w filozofii moralnej Georges’a Bastide’a i Jeana Naberta*, (1998); *Społeczne perspektywy etyki* (2006); *Aksjologia Louisa Lavelle’a wobec ponowoczesnego kryzysu wartości* (2012); *Nurty etyki. Od starożytności do nowożytności* (2018).

Barbara Ż m u d a - F r y d r y c h o w s k a, doktor, filozof. Studia z zakresu filozofii na Uniwersytecie Papieskim Jana Pawła II w Krakowie.

Od 2019 zatrudniona na Wydziale Filozoficznym UPJPII, obecnie (od 2021 r.) adiunkt w Katedrze Etyki na Wydziale Filozoficznym.

Główne obszary zainteresowań badawczych: etyka stosowana, etyka współczesna, etyka ponowoczesna, problematyka dylematów moralnych, współczesne spory etyczne.

Autorka książki *Personalizm Adama Rodzińskiego wobec wyzwań współczesnej kultury* (2020).

CONTENTS  
*Ethos* 36, no. 4 (144) (2023)

ARTIFICIAL INTELLIGENCE AND THE HUMAN UNIVERSE

From the Editors, Rozwijać czy nie rozwijać? Oto jest pytanie (A.L.K., K.W.) . . . . .	5
From the Editors, To develop or not to Develop, that is the Question (A.L.K., K.W.) . . . . .	11
J a n P a w e ł II, Misja Kościoła w erze komputerów (Orędzie papieskie na XXIV Światowy Dzień Środków Społecznego Przekazu 27 maja 1990 roku, Watykan, 24 I 1990). . . . .	19
J o h n P a u l II, The Christian Message in a Computer Culture (Message for the 24th World Communications Day 27 May 1990, Vatican, 24 January 1990). . . . .	23

THE QUESTION OF SUBJECTIVITY

Zbigniew W r ó b l e w s k i, Paweł F o r t u n a, Moral Subjectivity and the Moral Status of Artificial Intelligence: A Philosophical and Psychological Perspective . . . . .	29
DOI 10.12887/36-2023-4-144-05	
Małgorzata A b a s s y, Storytelling as a Tool to Tame Artificial Intelligence: Between the Modeling Function of Language and the Subjectivity of Machines . . . . .	49
DOI 10.12887/36-2023-4-144-06	

ARTIFICIAL INTELLIGENCE IN THE WORLD OF FICTION

Andrzej B o r k o w s k i, Artificial Intelligence and the Problem of Language and Text: A Reading of Stanisław Lem's <i>Golem XIV</i> . . . . .	67
DOI 10.12887/36-2023-4-144-07	
Barbara K l o n o w s k a, Ethical Machines: Representations of Artificial Intelligence in Ian McEwan's <i>Machines Like Me</i> and Kazuo Ishiguro's <i>Klara and the Sun</i> . . . . .	87
DOI 10.12887/36-2023-4-144-08	

- Grzegorz Trębicki, The Marriage of Two Worlds: A Quasi-Christian Vision of the Evolution of the Universe and Humanity in Dan Simmons' "Hyperion Cantos" . . . . . 107  
DOI 10.12887/36-2023-4-144-09

### FEARS AND EXPECTATIONS

- Bartłomiej K nosała, Environmental Management by Artificial Intelligence: Challenges and Limitations of the Concept of Postnature. . . . . 127  
DOI 10.12887/36-2023-4-144-10
- Izabela Kaczmarszyk, Reductive Pitfalls of the Sociology of "Thinking Machines" . . . . . 146  
DOI 10.12887/36-2023-4-144-11
- Zbigniew Nawrat, Medicine and the Robotic Ethos. . . . . 169  
DOI 10.12887/36-2023-4-144-12

### KAROL WOJTYŁA – JAN PAWEŁ II INSPIRATIONS

- Fr. Władysław Zuziak, Barbara Żmuda - Frydychowska, Man's Participation in the World: Wojtyła – Rorty . . . . . 193  
DOI 10.12887/36-2023-4-144-13

### THINKING ABOUT THE FATHERLAND...

- Barbara Trygarr, "Great is this force of memory"... *Mój Ojciec* ["My Father"] by Kazimierz Braun . . . . . 211

### NOTES AND REVIEWS

- Lukasz Mścistański, When Wisdom and BigTech-AI Collide (review of E. Spence, *Stoic Philosophy and the Control Problem of AI Technology: Caught in the Web*, Lanham, Rowman & Littlefield, Lanham 2021). . . . . 223
- Damian Szczęch, Planning a Life in a World Dominated by Artificial Intelligence (review of C. Campbell, *AI by Design: A Plan for Living With Artificial Intelligence*, CRC Press, New York 2022) . . . . . 231
- Books recommended by *Ethos* (J. Puckett Jr., *Apologetyka Radości: C.S. Lewisa argument z Pragmatyzmu na rzecz istnienia Boga*, trans. P. Bylica, Fundacja Prodeoteo, Warszawa 2023). . . . . 236



---

THROUGH THE PRISM OF THE ETHOS

Dorota C h a b r a j s k a, Warranty for a Little Bit of Mind . . . . . 245

Notes about the Authors. . . . . 253



ETHOS  
Kwartalnik Instytutu Jana Pawła II KUL  
ISSN 0860-8024  
Prefix DOI 10.12887

Instytut Jana Pawła II  
Wydział Filozofii  
Katolicki Uniwersytet Lubelski Jana Pawła II  
Al. Raclawickie 14, 20-950 Lublin  
tel. 81 4453218  
fax 81 4453217  
e-mail: ethos@kul.lublin.pl

#### INFORMACJE DLA AUTORÓW

Tematy kolejnych numerów monograficznych „Ethosu” są zapowiadane w wyprzedzeniu na czwartej stronie okładki oraz na stronie internetowej kwartalnika. Artykuły związane z tematami poszczególnych numerów można nadsyłać od chwili ukazania się zapowiedzi. Wszystkie artykuły naukowe publikowane w „Ethosie” poddawane są recenzji w trybie anonimowym dla obu stron, dokonywanej przez dwóch niezależnych recenzentów zewnętrznych. Jeśli jedna z recenzji okaże się negatywna, redakcja może podjąć decyzję o skierowaniu tekstu do trzeciego recenzenta. W przypadku, gdy trzecia recenzja jest pozytywna, autor zobowiązany jest do odniesienia się w zmienionej wersji artykułu do uwag wszystkich recenzentów. Jeśli obydwie recenzje są negatywne, artykuł nie zostanie opublikowany.

Redakcja zastrzega sobie prawo do decyzji o nieskierowaniu artykułu do recenzji, a także o rezygnacji z publikacji artykułu na każdym etapie prac redakcyjnych. W sprawach spornych ostateczną decyzję o publikacji artykułu podejmuje redaktor naczelny kwartalnika „Ethos”.

W celu ułatwienia procesu recenzji anonimowej, prosimy nie umieszczać w tekście informacji, które pozwoliłyby zidentyfikować autora.

Przed publikacją wszystkie teksty poddawane są adiustacji językowej i opracowaniu redakcyjnemu zgodnie ze stylem edytorskim czasopisma, a następnie autoryzowane.

#### ZASADY OGÓLNE

Objętość artykułu, wraz z bibliografią i abstraktami, nie powinna przekraczać 20 stron maszynopisu (1 arkusza wydawniczego – 40 tysięcy znaków ze spacjami), objętość recenzji i sprawozdań – 10 stron maszynopisu (0,5 arkusza wydawniczego – 20 tysięcy znaków ze spacjami).

Do artykułu należy dołączyć abstrakt (o objętości nie większej niż 1 strona maszynopisu – 1800 znaków ze spacjami) i słowa kluczowe w języku polskim i angielskim, a także bibliografię zawiera-

jąca pełne opisy cytowanych pozycji (z rozpisanymi imionami autorów, tłumaczy i redaktorów oraz pełnymi nazwami wydawnictw).

Abstrakt powinien być napisany w metajęzyku i zawierać informacje, jaki jest cel artykułu i jaki problem został w nim poruszony, jakie metody badawcze autor wykorzystał, do jakich wniosków doszedł i w jakich kierunkach badania przedstawione w tekście mogą być kontynuowane w przyszłości.

W przypadku tekstów nadsyłanych w języku angielskim prosimy stosować zasady edytorskie rekomendowane przez The Chicago Manual of Style.

Artykuły niedostosowane do stylu edytorskiego „Ethosu” nie będą przyjmowane.

Autorzy zobowiązani są do wypełnienia i przesłania (wraz z artykułem) do redakcji (1) oświadczenia o oryginalności dzieła oraz (2) umowy o dzieło i przeniesienie autorskich praw majątkowych, a także (3) formularza, na podstawie którego opracowywane są ich noty biograficzne.

Złożenie artykułu do publikacji jest rozumiane jako deklaracja, że tekst ten nie został nigdzie wcześniej opublikowany również w postaci elektronicznej oraz że nie jest on równocześnie złożony w redakcjach innych czasopism ani rozpatrywany jako fragment monografii.

Zasadniczo nie zamieszczamy materiałów, które ukazały się wcześniej w druku lub w postaci publikacji elektronicznej. Wyjątkiem są tłumaczenia z języków obcych tekstów dotąd niedostępnych w języku polskim.

Tekst wykładu lub odczytu powinien zawierać informację o tym, kiedy i na jakim forum był prezentowany.

Autorzy proszeni są o ujawnianie wkładu wszystkich osób, które przyczyniły się powstania składanego do publikacji artykułu (z podaniem ich afiliacji oraz sprecyzowaniem ich udziału), oraz informacji o źródłach finansowania badań przeprowadzonych w związku z jego przygotowywaniem.

Ghostwriting i guest authorship są przejawem nierzetelności naukowej, a wszelkie wykryte przypadki będą demaskowane, włącznie z powiadomieniem odpowiednich podmiotów. Wszelkie przejawy nierzetelności naukowej będą dokumentowane.

## FORMAT PLIKU

Prosimy o składanie artykułów w postaci elektronicznej, w formacie doc lub docx. W przypadku, gdy praca na pliku przygotowanym za pomocą innego edytora niż Microsoft Word okazuje się utrudniona ze względów technicznych, redakcja zastrzega sobie prawo do rezygnacji z publikacji tekstu.

## FORMAT DOKUMENTU

Formatowanie tekstu prosimy ograniczyć do minimum. Prosimy na żadnym etapie pracy na pliku nie stosować specjalnych „stylów” dostępnych w edytorze tekstu, nie zamieszczać komentarzy „w dymkach” ani nie posługiwać się opcjami: „porównaj dokumenty”, „pokaż adiustację” ani „śledź zmiany”.

W całym dokumencie prosimy stosować czcionkę Times New Roman 12 pkt i nie wydzielać cytatów za pomocą innego kroju, rozmiaru czcionki ani zmiany szerokości kolumny.

W artykułach należy stosować alfabet łaciński, a w przypadku słów pochodzących z języków posługujących się innym alfabetem używać transliteracji/transkrypcji.

Kursywę prosimy stosować wyłącznie w tytułach cytowanych publikacji. Słów obcojęzycznych prosimy nie zapisywać kursywą.

Preferowana interlinia: 1,5.

Wcięcia akapitowe prosimy wprowadzać za pomocą tabulatora (a nie spacji).

W całym tekście prosimy wprowadzić numerację stron.

Przypisy należy umieszczać na dole strony (przypisów końcowych nie stosujemy). Przypisy powinny zawierać informacje bibliograficzne; wszelkie informacje oraz argumentacja powinny znajdować się w tekście górnym.

## FORMAT PRZYPISÓW

Gwoli jednolitości edytorskiej, którą chcemy utrzymać, mimo że pismo ma charakter interdyscyplinarny, a w różnych dziedzinach stosowane są różne zapisy bibliograficzne, prosimy wszystkich autorów o stosowanie następujących reguł:

Przypisy, w których po raz pierwszy przywoływane jest określone dzieło, powinny zawierać pełne dane bibliograficzne, według wzoru:

pozycja książkowa (wymagane dane: autor, tytuł, tłumacz, redaktor, wydawnictwo, miejsce i rok wydania),

np. R. Buttiglione, *Mysł Karola Wojtyły*, tłum. J. Merecki SDS, Instytut Jana Pawła II KUL, Lublin 2010.

artykuł w czasopiśmie (wymagane dane: nazwisko autora, tytuł, tytuł czasopisma, tom/rocznik, (w nawiasie) rok, a następnie numer bieżący i (w nawiasie) ciągły oraz numery stron),

np. G.E.M. Anscombe, *Nowożytna filozofia moralności*, tłum. M. Roszyk, „Ethos” 23(2010) nr 4(92), s. 39-60.

artykuł w publikacji zbiorowej (wymagane dane: autor, tytuł, tłumacz, tytuł publikacji zbiorowej, tłumacz(e), redaktor(zy), wydawnictwo, miejsce i rok wydania oraz numery stron), np. S. Judycki, *Istnienie i natura duszy ludzkiej*, w: *Antropologia*, red. S. Janeczek, Wydawnictwo KUL, Lublin 2010, s. 121-177.

W przypadku cytatu prosimy o podawanie wyłącznie numeru strony, z której cytat pochodzi. W przypadku odwoływania się do fragmentu tekstu prosimy zamieszczać informację, na których stronach fragment ten się znajduje.

Kiedy dzieło cytowane jest po raz kolejny, prosimy stosować zapis skrócony: nazwisko autora, dz. cyt., numer strony,  
np. Buttiglione, dz. cyt., s. 10.

Jeśli wcześniej cytowane są również inne dzieła danego autora, po nazwisku autora prosimy podać tytuł dzieła,  
np. Buttiglione, *Myśl Karola Wojtyły*, s. 10.

W przypadku dzieł klasycznych, w których stosowana jest wewnętrzna notacja, prosimy zamieścić ją po tytule dzieła,  
np. Platon, *Timajos*, 86 B-C, w: tenże, *Timajos, Kritias*, tłum. W. Witwicki, Antyk, Kęty 2002, s. 85.

Cytaty z nauczania Jana Pawła II prosimy przytaczać za polskim wydaniem „L'Osservatore Romano”, *Nauczaniem papieskim* (red. E. Weron SAC, A. Jaroch SAC, Pallottinum) lub za wydaniem watykańskimi.

Skrót zob. odnosi czytelnika do całości dzieła, a skrót por. do jego fragmentu (wymagamy wówczas podania zakresu numerów stron).

W przypadku cytowania haseł ze słowników bądź wydawnictw o charakterze encyklopedycznym prosimy o następujący zapis: autor, hasło „Tytuł hasła”, w: tytuł publikacji, redaktor, wydawnictwo, miejsce i rok wydania, numery stron (kolumny, szpalty etc.).

W przypadku klasycznych zbiorów tekstów źródłowych, w których nie ma numeracji stron i we wszystkich wydaniach zachowana jest ta sama numeracja, np. kolumn, dopuszczamy podanie wyłącznie pełnego tytułu i numeru, np. kolumny.

Prosimy nie stosować skrótów: passim, j.w. ani znaku =.

Prosimy rozpisywać wszystkie skróty, zastępować cyfry rzymskie łańskimi i nie podawać numerów wydań cytowanych pozycji.

## CYTOWANIE Z TEKSTÓW OBCOJĘZYCZNYCH

Cytaty z tekstów obcojęzycznych powinny pochodzić z dostępnych przekładów na język polski. W przypadku braku źródła polskojęzycznego należy zaznaczyć, że tłumaczem tekstu jest autor artykułu bądź podać nazwisko tłumacza. Polskie przekłady (polskie wersje językowe) uważamy za podstawę odniesień i przestrzegamy zasady, że oryginalne wersje językowe przywoływane są tylko wtedy, gdy między nimi a polskimi przekładami zachodzą istotne różnice. W przypadkach takich dopuszczalne jest wprowadzenie tłumaczenia własnego przy zachowaniu informacji bibliograficznej na temat dostępnego polskiego przekładu.

W przypadku braku polskiego tłumaczenia danego dzieła prosimy cytować jego fragmenty za tekstem oryginalnym, a nie za jego obcojęzycznymi przekładami, np. nie należy cytować tekstów Hume'a za ich przekładami na język włoski.

## ILUSTRACJE

Autorzy zobowiązani są uzyskać zgodę na publikację towarzyszących artykułom materiałów objętych prawem autorskim (np. fotografii, faksymile, reprodukcji dzieł sztuki). Prosimy dołączać pliki graficzne o rozdzielczości 300 dpi (preferowany format TIF). Każda ilustracja powinna zostać opatrzona podpisem zawierającym informacje o jej treści, źródle oraz wszelkie inne informacje wymagane prawem autorskim.

## PRAWA AUTORSKIE

Złożenie tekstu do druku jest równoznaczne z przekazaniem wyłącznych praw autorskich kwartalnikowi „Ethos”. W przypadku zamiaru przedruku tekstu opublikowanego wcześniej w „Ethosie” autor (lub w jego imieniu wydawca) proszony jest o zwrócenie się o zgodę do Instytutu Jana Pawła II KUL. Prośby takie są z reguły rozpatrywane pozytywnie z zastrzeżeniem, że ponownej publikacji towarzyszyć będzie informacja bibliograficzna na temat pierwodruku.

Prawa autorskie na przedruk, przekład bądź wykorzystanie w formie elektronicznej materiału pierwotnie opublikowanego w „Ethosie”, jeśli występuje o nie podmiot inny niż autor materiału, udzielane są odpłatnie. Składając prośbę o udzielenie praw autorskich, prosimy określić charakter publikacji, jej język oraz nakład. „Ethos” nie we wszystkich przypadkach jest właścicielem praw autorskich do opublikowanych w piśmie materiałów. Przed złożeniem wniosku prosimy zwrócić uwagę na informacje zawarte w przypisach.





ETHOS  
QUARTERLY OF THE JOHN PAUL II INSTITUTE  
ISSN 0860-8024  
DOI prefix: 10.12887

The John Paul II Institute  
Faculty of Philosophy  
The John Paul II Catholic University of Lublin  
Al. Raławickie 14, 20-950 Lublin, Poland  
Phone: +48 81 4453218  
Fax +48 81 4453217  
E-mail: ethos@kul.lublin.pl

#### SUBMISSION GUIDELINES

The themes of the forthcoming volumes of *Ethos* can be found on the back cover of every volume and on the journal's website. Submissions are subject to double-blind peer review by two independent and external referees and to editorial revisions. One of the referees will represent an academic milieu from outside Poland, non-related to that of the author; the other will be a Polish scholar. In order to facilitate double-blind peer review process, authors are requested to eliminate all references that might identify them.

The editors reserve the right to decide whether a received manuscript will be considered for publication.

#### GENERAL GUIDELINES

The suggested length for *Ethos* articles is 4,000 to 8,000 words, including notes. Review and report articles should not exceed 4,000 words.

An article should be followed by a bibliography comprising a full list of the references, an abstract (up to 1,800 characters), and keywords.

The abstract should define the goal of the paper and outline the problem it discusses, as well as the methods applied throughout the scrutiny, the conclusions the author has reached, and the proposed further research on the subject.

Authors are requested to fill in a form to provide information necessary for the editors to prepare biographical notes about the contributors.

Submission of an article written in a language other than Polish should be accompanied by a declaration that the text in question has not been published elsewhere, either in print or online, and that it is not being considered for publication in any other periodical or monograph.

Authors are requested to disclose names of anyone participating in the preparation of the submission (together with their affiliation and contribution), as well as all the sources of research funding used in its preparation.

Ghostwriting and guest authorship are violation of the standard codes of scholarly conduct, and all its instances will be unmasked and documented, and the appropriate entities notified.

## FORMATTING THE FILE

Please submit your file as a Word document (doc or docx).

## FORMATTING THE DOCUMENT

Lines should be double-spaced, including quotations and notes, using Times New Roman font in 12 point size. Format your paper following the latest edition of the *Chicago Manual of Style*. Place page numbers on each page. Notes should be numbered consecutively and placed at the foot of each page.

Notes should be used primarily for citation purposes and include all the necessary bibliographic information. Notes should not contain any discussion that can be incorporated into the body of the paper.

## IMAGES

Authors are expected to obtain permission to reproduce any copyrighted materials (e.g., photographs, facsimiles, art reproductions) used in their articles. Please submit 300 dpi TIF files. Supply each illustration with a caption, accompanied by a source line and any required acknowledgments.

## QUOTATIONS FROM FOREIGN LANGUAGES

In cases where standard scholarly translations of quoted texts are available, scholars may rely upon them for both text and notes. If such editions are not readily available, the author must provide the original texts for the use of reviewers and translators.

Please use the Latin alphabet throughout your manuscript and apply transliteration/transcription in cases of words from languages using other alphabets.

## CITATIONS FORMAT

The first reference to a given work should include full bibliographic information, as in the following basic citation examples:

(book)

Rocco Buttiglione, *Mysl Karola Wojtyly*, trans. Jaroslaw Merecki (Lublin: Instytut Jana Pawla II KUL, 2010).

(journal article)

Edward Sapir, "Sound Patterns in Language," *Language* 1, no. 2 (1925): 37–51.

(article in a monograph)

Stanisław Judycki, "Istnienie i natura duszy ludzkiej," in: *Antropologia*, ed. Stanisław Janeczek (Lublin: Wydawnictwo KUL, 2010), 121–77.

In cases of direct quotations, please supply only the number of the page from which the given quotation comes.

When a given work has already been cited, please shorten the reference to: author, title, page number.

References to the classical texts must include the standard identification numbers for those texts, e.g., Plato, *Timaeus*, 86 B-C.

References to the speeches of John Paul II should include precise information on when and where a given speech was delivered. References to encyclicals and other Catholic Church documents should include the numbers of the sections referred to.

## COPYRIGHT AND PERMISSIONS

By submitting a text for publication, authors grant an exclusive license to the quarterly *Ethos*. Should authors wish to republish their article in another periodical or in a monograph, they are requested to apply for permission to the John Paul II Institute at the Catholic University of Lublin. Such permissions are generally granted on condition that all the required credits are included in the republished version.

If you wish to reproduce any material published in *Ethos* in a paper or in an electronic work that is not the author's own, please contact the John Paul II Institute at the Catholic University of Lublin. Permissions and rights are granted after processing your request and involve fee payment. Your request should include a description of the intended publication, its language and print run. Note that *Ethos* may not be the copyright holder of some of the material published in the journal. Please check the copyright or acknowledgment notes before making your request.

**Wydawca**

WYDZIAŁ FILOZOFII KUL  
OŚRODEK BADAŃ NAD MYŚLĄ JANA PAWŁA II  
„INSTYTUT JANA PAWŁA II KUL”  
Al. Raclawickie 14, 20-950 Lublin  
telefon: +48 81 4453217, telefon redakcji: +48 81 4453218  
e-mail: ethos@kul.pl

Instytut Jana Pawła II powołany został do istnienia uchwałą Senatu Akademickiego KUL 25 VI 1982 r. jako „międzywydziałowy ośrodek naukowy i dydaktyczny Katolickiego Uniwersytetu Lubelskiego”, przeznaczony do studiów „myśli i dzieła papieża Jana Pawła II” oraz do „budowania wspólnoty osób w duchu głoszonej przez Niego nauki Chrystusa” (Statut Instytutu). Działalność Instytutu koncentruje się na filozofii i teologii człowieka, moralności oraz misji duszpasterskiej św. Jana Pawła II. Instytut realizuje swoje zadania głównie poprzez doroczne sympozja naukowe, działalność badawczą i dydaktyczną (konwersatoria) oraz wydawniczą. Instytut Jana Pawła II stanowią: Rada Naukowa (przewodniczący – prof. dr hab. Andrzej Derdziuk OFMCap), Zarząd Instytutu (dyrektor – ks. dr hab. Marek Słomka, prof. KUL), pracownie oraz kwartalnik „Ethos”.

**Rada Naukowa  
Instytutu  
Jana Pawła II**

Prof. Dr hab. Andrzej DERDZIUK OFMCap – przewodniczący Rady Naukowej  
Ks. Prof. Dr hab. Mirosław KALINOWSKI – rektor KUL  
Ks. Prof. Dr hab. Mirosław SITARZ – prorektor KUL  
S. Prof. Dr hab. Beata ZARZYCKA – prorektor KUL  
Prof. Dr hab. Beata PISKORSKA – prorektor KUL  
Prof. Dr hab. Ewa TRZASKOWSKA – prorektor KUL  
Prof. Dr hab. Adam ZADROGA – prorektor KUL  
Prof. Dr hab. Barbara CHYROWICZ SSpS  
Ks. Prof. Dr hab. Tadeusz DOLA  
Prof. Dr hab. Alicja GRZEŚKOWIAK  
Prof. Dr hab. Wojciech KACZMAREK  
Prof. Dr hab. Karol KLAUZA  
Prof. Dr hab. Jarosław KUPCZAK OP  
Prof. Dr hab. Ewa Agnieszka LEKKA-KOWALIK  
Prof. Dr hab. Małgorzata U. MAZURCZAK  
Abp Prof. Dr hab. Alfons NOSSOL  
Prof. Dr hab. Robert PIŁAT  
Prof. Dr hab. Grzegorz PRZEBINDA  
Prof. Dr hab. Jacek SALIJ OP  
Prof. Dr hab. Stefan SAWICKI  
Ks. Prof. Dr hab. Jan SOCHOŃ  
Prof. Dr hab. Andrzej STOFF  
Ks. Prof. Dr hab. Stanisław SUWIŃSKI  
Prof. Dr hab. Władysław STRÓŻEWSKI  
Prof. Dr hab. Krzysztof WIAK  
Ks. Prof. Dr hab. Alfred M. WIERZBICKI  
Ks. Prof. Dr hab. Henryk WITCZYK  
Prof. Krzysztof ZANUSSI  
Prof. Dr hab. Zofia ZDYBICKA USJK

**Warunki  
prenumeraty****Cena prenumeraty krajowej na rok 2024: rocznie 100,00 zł.**

Prenumeratorzy indywidualni oraz instytucje mogą zamawiać prenumeratę, wpłacając wymienioną kwotę na konto 94 1050 1953 1000 0024 3072 5941 z zaznaczeniem nazwy rachunku: „Prenumerata ETHOSU”.

Prenumeratę zagraniczną można zamawiać, wpłacając odpowiednią kwotę na konto:

33 1050 1953 1000 0090 3079 8129 (USD)

95 1050 1953 1000 0090 8066 3355 (EUR)

z zaznaczeniem nazwy rachunku: „Prenumerata ETHOSU”.

Cena pojedynczego numeru – 17 USD (16 EUR), prenumerata roczna – 55 USD (50 EUR).

**Kolportaż**

Kwartalnik ETHOS (Tomasz Górka), Al. Raclawickie 14, 20-950 Lublin  
telefon: +48 81 4453217, e-mail: eth-kolp@kul.pl

## W poprzednich numerach:

Stanisław Wyspiański: dramat i ethos (48)  
Chrześcijaństwo na progu trzeciego tysiąclecia (49-50)  
Polskie przełomy (51)  
Przesłanie Zbigniewa Herberta (52)  
Etyka a demografia (55)  
Dzieje: nie-Boska komedia? (56)  
O solidarność rodziny narodów Europy (57-58)  
O solidarność rodziny narodów świata (59-60)  
Prawo niedoskonałe czy niesprawiedliwe? (61-62)  
Ethos Pielgrzyma (63-64)  
Dramat Judasza (65-66)  
Polska Jana Pawła II (67-68)  
Osoba w społeczeństwie informacyjnym (69-70)  
Wojna sprawiedliwa? (71-72)  
Muzyka – piękno, dobro, świętość (73-74)  
O nowej edukacji (75)  
O filozofii Karola Wojtyły (76)  
Ethos teatru (77-78)  
Białoruś i Ukraina – bliższa Europa (81)  
Ciało – osoba – kultura (82-83)  
„I będą oboje jednym ciałem” (84)  
Oblicza emigracji (87-88)  
Kino i transcendencja (89)  
O winie i karze (90-91)  
Między cnotą a niegodziwością (92)  
O radości (93-94)  
Filozof osoby – Tadeusz Styczeń SDS (1931-2010) (95)  
Ubóstwo dziś (96)  
Ethos słowa (97-98)  
Czas i wieczność (99)  
Filozofia Vaticanum Secundum (100)  
Świadomość i tożsamość (101)  
Nasi „bracia mniejsi” (102)  
O przyjaźni (103)  
Ethos przestrzeni (104)  
Świętość – ideał i fakt (105)  
O przemocy (106)  
Wobec utopii (107)  
Lęk (108)  
Przyszłość humanistyki (109)  
Szaleństwo (110)  
Transhumanizm (111)  
Ethos filozofii (112)  
O ciszy i milczeniu (113)  
Ethos wizerunku (114)  
Chrześcijaństwo w kulturze Zachodu (115)  
Koniec osoby? (116)  
Obcy (117)  
Wstyd (118)  
Ethos światła (119)  
Ból (120)  
Pieniądz (121)  
Krzywda (122)  
Książka (123)  
Upadek (124)  
Słuchanie (125)  
Muzyczność (126)  
Piękno (127)  
Hope (128, tom w języku angielskim)  
Rozmowa (129)  
Dom (130)  
Rzeczy (131)  
Władza (132)  
Głupota (133)  
Język (134)  
Pamięć (135)  
Ludzie i miejsca (136)  
O jedzeniu (137)  
Widzenie (138)  
Samotność (139)  
Ku interpretacji wolności (140)  
Wolność – kultura – wartości (141)  
Przypadek (142)  
Oswoić sztuczną inteligencję (143)

---

## W najbliższych numerach:

- \* Godność
  - \* Truth and Post-truth (tom w języku angielskim)
  - \* Harmonie świata
  - \* Rozum i wiara
  - \* Oczekiwanie
- 

Cena brutto: 25 zł  
w tym VAT: 8%

ISSN 0860-8024



9 770860 802403

www.ethos.lublin.pl