

KAROL KLAUZA
MAŁGORZATA GRUCHOŁA

SPOŁECZNE WYZWANIA OSOBLIWOŚCI TECHNOLOGICZNEJ.
KOMPLEKSYFIKACJA RZECZYWISTOŚCI
HUMANO-WIRTUALNEJ

*...każda zaawansowana technologia
jest nie do odróżnienia od magii
[III prawo Arthura Charlesa Clarke'a
Eseje do „Profiles of the Future”¹*

Z pewnością intuicja cytowanego tu Arthura C. Clarke’a, pisarza science-fiction z wizjami nowych technologii, które w drugiej połowie XX wieku stawały się faktami, wpisuje się swym trzecim prawem² w problematykę osobliwości technologicznej (lub po prostu osobliwości, z ang. *singularity*). W semantyce tego pojęcia zawiera się, zgodnie z przyjętą współcześnie terminologią, hipotetyczny punkt w przyszłym rozwoju cywilizacji, w którym postęp techniczny stanie się tak szybki, że wszelkie ludzkie przewidywania staną się nieaktualne³. O tempie rozwoju tego procesu decyduje m.in. dynamizm

Prof. dr hab. Karol Klauza – Katolicki Uniwersytet Lubelski Jana Pawła II; adres do korespondencji: Al. Raławickie 14 20-950 Lublin; e-mail: karolklauza@gmail.com; ORCID: <https://orcid.org/0000-0002-0969-5537>.

Dr hab. Małgorzata Gruchola, prof. KUL – Katolicki Uniwersytet Lubelski Jana Pawła II; adres do korespondencji: Al. Raławickie 14 20-950 Lublin; e-mail: malgorzata.gruchola@kul.pl; ORCID: <https://orcid.org/0000-0002-2367-0416>.

¹ Arthur C. CLARKE, *Profiles of the Future* (London: Indigo, 2000) (tłum. własne).

² CLARKE, *Profiles of the Future*. Trzy prawa Arthura Clarka: „I. Kiedy poważany, a sędziwy naukowiec twierdzi, że coś jest możliwe, prawie na pewno ma rację. Gdy twierdzi, że coś jest niemożliwe, prawdopodobnie się myli. II. Jedyнным sposobem poznania granic możliwego jest ich lekkie przekroczenie i wejście w niemożliwe; III. Każda zaawansowana technologia jest nie do odróżnienia od magii”.

³ Vernor VINCE, „*The Coming Technological Singularity: How to Survive in the Post-Human Era*”. Referat wygłoszony na sympozjum VISION-21 (30-31 marca 1993) sponsorowanym przez

rewolucyjnego fenomenu sztucznej inteligencji, ukierunkowany na osiągnięcie poziomu superinteligencji, istotnego elementu transhumanizacji społeczeństw. Na ten dynamizm rewolucyjnego fenomenu składają się także: rozwój szybkich i silnych komputerów kwantowych, postęp w globalnej sieci urządzeń cyfrowych, postępy nauk medycznych w kierunku biologiczno-elektronicznych powiązań ludzkiego układu nerwowego z urządzeniami technologii cyfrowych⁴. Równolegle wpływają nań także procesy związane ze środowiskiem życia, nastawione na postęp zrównoważonego rozwoju i maksymalnego zachowania różnorodności biologicznej. W efekcie oczekuje się od nowych technologii, że osiągną potencjał superinteligencji będzie wyznaczać dziejowy status transhumanizmu cywilizacyjnego, mierzony skalą zachowania gatunku, odpowiednio do współczesnej interpretacji antycznych zasad: kosmologicznej (*Panton krematon metron anthropos* – ‘miarą wszystkich rzeczy jest człowiek’) oraz legislacyjnej (*Propter hominem omne ius constitutum est* (‘ze względu na człowieka wszelkie prawo jest stanowione’)). Oba te paradygmaty wyznaczają granice dyskursu etycznego i pragmatycznego związanego z nowymi technologiami, w odniesieniu do których sami ich twórcy zgłaszają nieodzowne wyzwania. Wystarczy odwołać się do sugestii Billa Gatesa, wygłoszonych w 2023 roku⁵. Jego zdaniem

Historia pokazuje, że możliwe jest sprostanie wyzwaniom stworzonym przez nowe technologie, a jeśli rządy i sektor prywatny zrobią to, co do nich należy, możemy to zrobić ponownie [...] nie powinno się powstrzymywać ludzi przed wdrażaniem nowych rozwiązań w zakresie sztucznej inteligencji, jak niektórzy proponowali⁶.

NASA Lewis Center oraz Ohio Aerospace Institut, dostęp 24.11.2024, <https://edoras.sdsu.edu/~vin-ge/misc/singularity.html>.

⁴ Tim URBAN, „The AI Revolution: The Road to Superintelligence”, dostęp 15.11.2024, <https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html>. Dokumentuje rozróżnienie skali zastosowań sztucznej inteligencji na: ANI (ang. *Artificial Narrow Intelligence*) – dziedzinową sztuczną inteligencję; AGI (ang. *Artificial General Intelligence*) – ogólnodziedzinową sztuczną inteligencję, zdolną do rozumienia, planowania, rozwiązywania problemów z elementami abstrakcyjnego myślenia ideami, szybkiego przyswajania danych i uczenia się w miarę działania; ASI (ang. *Artificial Superintelligence*) – sztuczną Superinteligencję, która zgodnie z poglądem Nicka Bostroma będzie charakteryzować się sprawniejszą inteligencją niż ludzka, w tym także naukową kreatywnością.

⁵ „Bill Gates przedstawił 5 zagrożeń związanych ze sztuczną inteligencją”, dostęp 11.01.2025, <https://forsal.pl/lifestyle/technologie/artykuly/8781146,bill-gates-przedstawil-5-zagrozen-zwiazanych-ze-sztuczna-inteligencja.html>.

⁶ „Bill Gates przedstawił 5 zagrożeń związanych ze sztuczną inteligencją”.

Jego zdaniem m.in. hakerzy lub osoby, które chcą wykorzystać AI w złych celach, nie zaprzestaną nad tym pracować: „[...] Wysiłki mające na celu ich powstrzymanie muszą być kontynuowane w tym samym tempie”⁷. Jak dodaje

[...] rządy powinny rozważyć utworzenie globalnego organu zajmującego się sztuczną inteligencją, podobnego do Międzynarodowej Agencji Energii Atomowej. Inne organizacje, w tym Alan Turing Institute i National Institute of Standards and Technology, pracują nad problemem stronniczości. Jednym z przybliżeń jest wbudowanie w sztuczną inteligencję wartości ludzkich i rozumowania wyższego poziomu. Jest to analogiczne do sposobu, w jaki działa samoświadomy człowiek. Dlatego Gates uważa, że należy sprawdzać informacje napisane przez AI [...]. Każdy, kto korzysta ze sztucznej inteligencji, musi zdawać sobie sprawę z problemu uprzedzeń i stać się świadomym użytkownikiem⁸.

Wyżej przywołane paradygmaty (kosmologiczny i legislacyjny) wyznaczają także granice naukowego dyskursu etycznego i pragmatycznego, związanego z nowymi technologiami. Zdaniem Luciano Floridiego (ur. 1964) to etyka wraz z zarządzaniem (*governance*) i regulacją stanowią główne elementy systemu normatywnego dotyczącego technologii (w tym AI), który określa, w jakim kierunku chcemy ją rozwijać. Podczas gdy regulacje prawne określają, które działania są zgodne bądź niezgodne z obowiązującym prawem, etyka wskazuje, jakie postępowanie jest pożądane z perspektywy dobra wspólnego i jakości życia społecznego. Innymi słowy, etyka pełni funkcję moralnej busoli, umożliwiającej krytyczną ocenę rozwoju i zastosowań technologii w taki sposób, który przynosi możliwie największe korzyści społeczeństwu. Trzy komponenty systemu normatywnego (etyka, regulacja i zarządzanie) pozostają ze sobą w ścisłej relacji, przy czym to etyka zajmuje pozycję nadrzędną, gdyż poprzez moralną refleksję kształtuje zarówno obszar regulacji, jak i praktyki zarządzania. Regulacja oddziałuje na zarządzanie poprzez wymóg zgodności z prawem, zarządzanie natomiast stanowi praktyczny wymiar realizacji zasad etycznych i norm prawnych⁹.

Istotny kolejny element – regulacje prawne Unii Europejskiej zawarte w akcie prawnym z 21 maja 2024 r. (z wejściem w życie w 2026 roku)¹⁰ podają

⁷ „Bill Gates przedstawił 5 zagrożeń związanych ze sztuczną inteligencją”.

⁸ „Bill Gates przedstawił 5 zagrożeń związanych ze sztuczną inteligencją”.

⁹ Podaję za: Maciej CHOJNOWSKI, *Etyka sztucznej inteligencji* (Warszawa: Centrum Etyki Technologii Instytutu Humanites, 2002), 14.

¹⁰ Tekst wraz z historią redakcji w: Conseil de l'Union Européenne, „Législation sur l'intelligence artificielle (IA): le Conseil donne son feu vert définitif aux premières règles mondiales en matière d'IA”, dostęp 23.05.2024, <https://www.google.com/>

klasyfikację technologii sztucznej inteligencji z uwzględnieniem różnorodności jej zastosowań i stopnia ryzyka, który należy uwzględnić w praktyce, w zależności od autoryzacji sugerowanych rozwiązań lub ich ewentualnego zakazu. Ogólny kierunek dokumentu charakteryzuje się troską o zachowanie prawnie gwarantowanych wolności oraz ograniczaniem tej innowacji w pewnych sektorach.

Już po opublikowaniu tego dokumentu John Hopfield i Geoffrey Hinton, laureaci nagrody Nobla z fizyki w 2024 roku za odkrycia w zakresie powiązań neurologicznych ze sztuczną inteligencją¹¹, wyrazili swoje zastrzeżenia co do nieuniknionych tendencji w zakresie pełnej kontroli nad zastosowaniami ich wynalazku. „Jako fizyk, niepokoję się w związku z przejawami niekontrolowanych procesów, których nie jesteśmy w stanie zatrzymać i poddać kontroli”¹².

1. ZAŁOŻENIA TEORETYCZNE

Rozwój sztucznej inteligencji (AI) stanowi nie tylko przełom technologiczny, lecz także głęboki impuls filozoficzny, który skłania do ponownego zdefiniowania pojęć świadomości, podmiotowości i odpowiedzialności. W literaturze przedmiotu to zjawisko jest analizowane w trzech powiązanych perspektywach: 1) koncepcji osobliwości technologicznej, 2) filozofii transhumanizmu i postczłowieczeństwa oraz 3) idei noosfery i kompleksyfikacji świadomości, które wspólnie tworzą ramy dla refleksji nad przyszłością ludzkiego poznania i etyki w epoce sztucznej inteligencji.

1.1. OSOBLIWOŚĆ TECHNOLOGICZNA I SUPERINTELIGENCJA

Koncepcja osobliwości technologicznej (ang. *technological singularity*), zaproponowana przez Vernora Vinge'a (1944–2024)¹³, opisuje moment, w którym

url?sa=t&source=web&rct=j&opi=89978449&url=https://www.consilium.europa.eu/fr/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/&ved=2ahUKEwjeqMjOu-WKAXVDAtsEHAYtA0YQFno-ECBUQAQ&usq =AOvVaw1lh99CCmPV3utzHCAwTQla. Zob. komentarz: „Intelligence artificielle: l'Union européenne pose des limites inédites”, dostęp 23.05.2024, <https://www.la-croix.com/international/intelligence-artificielle-l-union-europeenne-pose-des-limites-inedites-20240520>.

¹¹ „Intelligence artificielle: pourquoi les lauréats du Nobel de physique sonnent l'alarme”, dostęp 20.12.2024, <https://www.la-croix.com/international/intelligence-artificielle-pourquoi-les-laureats-du-nobel-de-physique-sonnent-l-alarme-20241009>.

¹² „Intelligence artificielle: pourquoi les lauréats du Nobel de physique sonnent l'alarme”.

¹³ Vernor VINGE, „The coming technological singularity: How to survive in the post-human era”, *Science fiction criticism: An anthology of essential writings* 81 (1993):352-363.

rozwój sztucznej inteligencji stanie się samonapędzający, prowadząc do przekroczenia granic ludzkiej zdolności poznawczej. Raymond Kurzweil (ur. 1948)¹⁴ rozwinął tę ideę, formułując „prawo przyśpieszających zwrotów”, zgodnie z którym postęp technologiczny przybiera charakter wykładniczy, a integracja człowieka i maszyny jest nieunikniona. W ujęciu Nicka Bostroma (ur. 1973)¹⁵ osobliwość może prowadzić do powstania superinteligencji – systemu przewyższającego człowieka we wszystkich domenach poznawczych. Badania współczesne wskazują, że eksplozja inteligencji wiąże się z ryzykiem systemowym oraz koniecznością opracowania strategii tzw. *superalignment*, czyli dopasowania celów maszyn do wartości ludzkich¹⁶. Koncepcja osobliwości technologicznej, przedstawiona przez V. Vinge’a¹⁷ i R. Kurzweila¹⁸ jest również krytykowana za quasi-religijny charakter – przedstawia technologię jako zbawcze narzędzie przekroczenia ludzkiej kondycji, co zbliża ją do sekularyzowanej eschatologii. Jak zauważa N. Bostrom¹⁹, problemem nie jest sama wizja przekroczenia człowieka, lecz brak refleksji nad społecznymi konsekwencjami tego procesu.

1.2. TRANSHUMANIZM I POSTCZŁOWIECZEŃSTWO

Filozofia transhumanizmu postuluje możliwość doskonalenia człowieka poprzez zastosowanie technologii zarówno biologicznych, jak i cyfrowych. W ujęciu Nicka Bostroma²⁰ oraz Maxa More (ur. 1964) i Natashy Vita-More (ur. 1950)²¹, transhumanizm stanowi program przekraczania ograniczeń biologicznych i poznawczych w kierunku postczłowieczeństwa (ang. *posthuman condition*). Francesca Ferrando²² podkreśla, że posthumanizm nie jest jedynie projektem technologicznym, lecz filozoficznym – redefiniuje relację między człowiekiem, naturą

¹⁴ Raymond KURZWEIL, *The singularity is near. Ethics and emerging technologies* (London: Palgrave Macmillan UK, 2005), 393-406.

¹⁵ Nick BOSTROM, *Superintelligence: Paths, Dangers, Strategies* (Oxford: Oxford University Press, 2014), 39.

¹⁶ Anthony M. BARRETT, Seth D. BAUM, „A model of pathways to artificial superintelligence catastrophe for risk and decision analysis”, *Journal of Experimental & Theoretical Artificial Intelligence* 29, nr 2 (2017):397-414.

¹⁷ VINGE, „The coming technological singularity: How to survive in the post-human era”, 355.

¹⁸ KURZWEIL, *The singularity is near*, 393-394.

¹⁹ BOSTROM, *Superintelligence*, 40.

²⁰ BOSTROM, *Superintelligence*, 40.

²¹ Max MORE, Natasha VITA-MORE, red., *The Transhumanist Reader: Classical and Contemporary Essays On the Science, Technology, and Philosophy of Human Future* (Malden–Oxford: John Wiley & Sons Ltd. Inc., 2013), 51.

²² Francesca FERRANDO, *Philosophical posthumanism* (London: Bloomsbury Academic, 2019), 12.

i technologią, odchodząc od antropocentryzmu. W tym kontekście AI jest nie tylko narzędziem, lecz partnerem poznawczym, współtworzącym nową epistemologię oraz formy komunikacji poznawczej w ramach rozszerzonej inteligencji.

Jednocześnie należy podkreślić, że choć transhumanizm przedstawia wizję emancypacji człowieka poprzez technologię, wielu badaczy zwraca uwagę, że jego założenia niosą ze sobą ryzyko redukcjonizmu antropologicznego i technologicznego determinizmu²³. Transhumanizm często marginalizuje wymiar etyczny i społeczny postępu, skupiając się niemal wyłącznie na potencjale biologiczno-poznawczym jednostki²⁴. W efekcie wizja „ulepszonych człowieka” może prowadzić do nowego rodzaju nierówności – cyfrowego elitaryzmu, w którym dostęp do technologii staje się warunkiem uczestnictwa w pełni ludzkiego doświadczenia. Luciano Floridi wskazuje także na zjawiska *ethics shopping* i *ethics bluewashing*, które prowadzą do instrumentalizacji etyki²⁵. Współczesne reinterpretacje koncepcji noosfery podkreślają jej potencjał, ale także ryzyko utopizacji wspólnoty poznawczej²⁶.

1.3. NOOSFERA I KOMPLEKSYFIKACJA ŚWIADOMOŚCI

Koncepcja noosfery Pierre’a Teilharda de Chardin SJ (1881–1955)²⁷ stanowi filozoficzne tło współczesnych analiz świadomości zbiorowej. Noosfera rozumiana jest jako kolejna faza ewolucji, w której myśl i refleksja stają się główną siłą kształtującą rzeczywistość. W ramach tej wizji proces kompleksyfikacji świadomości prowadzi do powstania sieciowej formy umysłu, globalnej „warstwy myśli”, będącej rezultatem kumulacji wiedzy i technologicznej interakcji. We współczesnym ujęciu noosfera interpretowana jest jako emergentny system poznawczy, w którym AI pełni funkcję katalizatora ewolucji poznawczej ludzkości²⁸.

Filozoficzne ujęcia osobliwości, transhumanizmu i noosfery ukazują wspólne dążenie do przekroczenia granic ludzkiej kondycji poprzez złożoność i integrację. Współczesna etyka technologii staje się nie tylko dyscypliną refleksyjną,

²³ FERRANDO, *Philosophical posthumanism*, 12.

²⁴ FERRANDO, *Philosophical posthumanism*, 12.

²⁵ Luciano FLORIDI, „Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical”, *Philosophy & Technology* 32 (2019):185-193, <https://doi.org/10.1007/s13347-019-00354-x>.

²⁶ Clément VIDAL, „What is the noosphere? Planetary superorganism, major evolutionary transition and emergence”, *Systems Research and Behavioral Science* 41, nr 4 (2024):614-622; Hub ZWART, *Continental philosophy of technoscience* (Cham: Springer International Publishing, 2021), 207-227.

²⁷ Pierre Teilhard de CHARDIN, *The phenomenon of man* (New York: Harper & Row, 1959), 302.

²⁸ VIDAL, „What is the noosphere?”, 614-615; ZWART, *Continental philosophy of technoscience*, 208.

lecz praktycznym narzędziem projektowania przyszłości, w kierunku większej spójności, odpowiedzialności i współzależności globalnej.

Współczesne reinterpretacje koncepcji noosfery²⁹ podkreślają jej potencjał w tworzeniu globalnej świadomości poznawczej. Jednakże, jak zauważa Hub Zwart³⁰, ta idea niesie również ryzyko utopizacji wspólnoty poznawczej – przekonania, że rozwój technologiczny automatycznie prowadzi do wzrostu refleksyjności i etyczności. Tymczasem sztuczna inteligencja często przyczynia się do polaryzacji społecznej, dezinformacji i fragmentaryzacji świadomości. Teilhardowska wizja kompleksyfikacji świadomości zakładała wzrost integracji w miarę rozwoju złożoności, jednak w kontekście algorytmicznego społeczeństwa ta zależność może zostać odwrócona – złożoność technologiczna nie musi oznaczać wzrostu spójności poznawczej, a wręcz może prowadzić do chaosu epistemicznego.

2. ZARYS OSOBLIWOŚCI TECHNOLOGICZNEJ

W latach sześćdziesiątych XX wieku, wraz z postępowym doskonaleniem urządzeń cyfrowych i ich oprogramowania, rozwijała się refleksja futurystyczna na bazie analiz kulturoznawczych, filozoficznych i socjologicznych. Stawała się ona częścią ogólniejszych prognoz następujących w wyniku coraz szybszych zmian na polu projektowania i budowania urządzeń wspierających proces inteligentnego wykorzystania narzędzi elektronicznych w różnych dziedzinach i systemach społecznych. Wówczas też wskazano na perspektywę kolejnych etapów rozwoju sztucznej inteligencji, prowadzących do osiągnięcia tzw. osobliwości technologicznej (ang. *singularity*). Twórcą tej kategorii był Stanisław Ulam, który w 1958 roku wykorzystał ją w prasowej dyskusji z Johnem von Neumannem³¹. Wymiana opinii dotyczyła geometrycznego przyspieszenia postępu technologicznego, który w perspektywie czasowej w pierwszej połowie XXI wieku miał doprowadzić do osiągnięcia stanu, w którym sterowanie postępowaniem technologicznym przekroczy pułap inteligencji ludzkiej. W ślad za tym postępowaniem zmiany miały dokonywać się w pozostałych systemach składających się na cywilizację lokalne.

²⁹ VIDAL, „What is the noosphere?“, 614-622.

³⁰ ZWART, *Continental philosophy of technoscience*, 220-227.

³¹ Stanisław Ulam (ur. 1909 we Lwowie, zm. 1984 w Santa Fe), matematyk, przedstawiciel Lwowskiej Szkoły Matematycznej, współtwórca bomby wodorowej, twórca Metody Monte Carlo, komputerowej symulacji procesów fizykochemicznych. Tekst opinii: Stanisław ULAM, „Tribute to John von Neumann”, *Bulletin of the American Mathematical Society* 64, nr 2-3 (1958):1-49.

Dalsze prace nad osobliwością technologiczną prowadził także matematyk i kryptolog Irving Good (1916–2009)³². Przyjął on założenie, że w postępie inteligencji dużą rolę będzie odgrywała współpraca człowieka i inteligencji wynalezionych przez niego urządzeń. Przy czym na pewnym etapie znaczne przyspieszenie postępu będzie owocem działań pozaludzkich. Zwłaszcza gdy zmiany przekroczą punkt powstania ultrainteligentnej maszyny. Poglądy Gooda upowszechnił w latach osiemdziesiątych XX wieku Vernor Steffen Vinge, matematyk i informatyk, znany przede wszystkim z twórczości literackiej, popularyzującej wizję świata po przekroczeniu osobliwości technologicznej³³. Spopularyzował powszechne użycie kategorii „osobliwość”, semantycznie tożsamej z naukowym rozumieniem „osobliwości technologicznej”.

Dekady przełomu XX–XXI wieku to wyraźne zróżnicowanie kierunków badań nad sztuczną inteligencją i nad jej wdrażaniem w konteksty życia społecznego. Zaznaczył się bowiem nurt poszukiwań relacji inteligencji ludzkiej i sztucznego modyfikowania biologicznego układu poznania sensorycznego i nerwowego (głównie dzięki postępom bioinżynierii, genetyki i farmaceutyki nootropowej). Dla zwolenników tego typu wspierania fenomenu inteligencji wiodące znaczenie mają eksperymenty modyfikowania mózgu ludzkiego. Prowadzą one od symulacji anatomicznej i funkcjonalnej tego organu aż po wspieranie jego wykorzystania implantami sztucznymi (organoidami mózgowymi)³⁴.

Drugi nurt poszukiwań dotyczy uwzględnienia podmiotowości sztucznej inteligencji w obszarach dotychczas przynależnych człowiekowi i konsekwentnie określenia jej etycznej i prawnej rangi w odniesieniu do gatunku ludzkiego i jego ewolucji. Współczesne próby sprzęgania sztucznej inteligencji z ludzkimi mózgami mają miejsce w większości ośrodków badania procesów poznawczych z jej wykorzystaniem. Według Wodzisława Ducha z Wydziału Fizyki, Astronomii

³² Irving GOOD, *Good Thinking the Foundation of Probability and its Applications* (Minneapolis: University of Minesota Press, 1983), <https://www.booksfree.org/good-thinking-the-foundation-of-probability-and-its-applications-by-i-j-good-pdf/>.

³³ Najważniejsze opracowania, nagradzane prestiżowymi tytułami: Vernor S. VINGE, *Marooned in Realtime* (New York: Bluejay Books/St. Martin's Press, 1986); Vernor S. VINGE, *Deepness in the Sky* (New York: Books Editor New York, 1999); Vernor S. VINGE, *Otchłań w niebie* (Warszawa: Wydawnictwo Prószyński i S-ka, 2003); Vernor S. VINGE, *True Names and the Cyberspace Frontier* (New York: Tor Publishing Group, 2001).

³⁴ Wodzisław DUCH, „Ludzkość+ za rogiem? Prof. Duch o transhumanizmie: sztuczna inteligencja zmienia wszystko”, dostęp 24.02.2025, <https://www.tvp.info/65670974/sztuczna-inteligencja-zmienia-wszystko>. Wykład prof. Wodzisława Ducha, wygłoszony w grudniu 2022 roku, został zorganizowany przez UMK w Toruniu w ramach YUFE Academy. Akademia YUFE to seria wykładów, warsztatów i zajęć organizowanych przez uczelnie zrzeszone w jednym z pierwszych uniwersytetów europejskich – konsorcjum YUFE (Young Universities for the Future of Europe).

i Informatyki Stosowanej UMK w Toruniu coraz większy jest postęp w metodach pomiaru aktywności mózgu, analizie danych neuroobrazowych i elektrofizjologicznych oraz zrozumieniu procesów zachodzących w mózgu. „Droga do Ludzkości+”, czyli transhumanistycznego społeczeństwa bardziej rozwiniętych gatunków ludzkich, jest pełna niebezpieczeństw, ale może wcale nie być aż tak długa³⁵. W tym opracowaniu wskazał na futurologiczny, złożony i ambiwalentny charakter tych badań³⁶. Zdaniem W. Ducha od 2020 roku ludzie mają do czynienia z tzw. czwartą rewolucją przemysłową, która jest efektem rozwoju sztucznej inteligencji. Przejawia się ona w szybkim rozwoju technologii, przede wszystkim elementów składowych sztucznej inteligencji i robotyki, w osiągnięciach nauk behawioralnych, neuropsychologii, udoskonalania ludzkich możliwości dzięki bioelektronice, ingerencji w mózg, genetycznym modyfikacjom oraz rozwoju medycyny spersonalizowanej.

2.1. FILOZOFICZNE ASPEKTY (TEILHARDYZM)

Dyskusje filozoficzne na temat humanistycznego wymiaru alienacji, rozumianej jako zniewolenie człowieka przez jego wytwór, prowadzone w połowie XX wieku, pozwoliły odkryć na nowo pozytywne perspektywy w rozumieniu dziejowego procesu ewolucji. Stało się to m.in. w przypadku teilhardyzmu. Wypracował go, w oparciu o paradygmaty nauk stosowanych, i wyraził w semantyce filozoficzno-teologicznej paleontolog, geolog i antropolog Pierre Teilhard de Chardin³⁷. Syntezę jego poglądów określić można jako efekt rzetelnych badań XX-wiecznej geologii, paleoantropologii i historiozofii, wyrażony w kontekście paradygmatów filozofii i teologii chrześcijańskiej. Składają się na nią:

- przekonanie o ciągłości czasowej procesu ewolucji kosmosu od aktu stwórczego świata materialnego po przewidywaną anihilację materialnej formy istnienia, z jednoczesną perspektywą „nowego nieba i nowej ziemi” wolnych od ograniczeń czasu;

³⁵ DUCH, „Ludzkość+ za rogiem?”.

³⁶ W. DUCH, *Sztuczna inteligencja: wyzwania i szanse*, dostęp: 12.01.2025, <https://www.youtube.com/watch?v=All7XRgd2rU>.

³⁷ Czesław BARTNIK, *Teilhardowska wizja dziejów* (Lublin: Wydawnictwo Katolickiego Uniwersytetu Lubelskiego, 1975); Marcin KARAS, *Historiozofia Teilharda de Chardina wobec tradycyjnej myśli chrześcijańskiej* (Kraków: Księgarnia Akademicka, 2012); Regina LUBAS-BARTOSZYŃSKA, „Zrehabilitowany” Pierre Teilhard de Chardin, dostęp 04.12.2024, <https://kongreskk.pl/2023/09/30/zrehabilitowany-pierre-teilhard-de-chardin/>. W przestrzeni medialnej, zob. film dokumentalny *L'Unité créatrice du monde – Pierre Teilhard de Chardin*, reż. Caroline Puig-Grenetier, 2018, seria „Documentaire”, KTO TV, styczeń 2023.

- wyodrębnienie punktów granicznych dla kolejnych etapów ewolucyjnego rozwoju w kierunku pozytywnego zrealizowania uporządkowanych procesów ewolucyjnych. Czynnikiem porządkującym stanowi kompleksyfikacja świadomości. Obejmuje pięć punktów granicznych, wyznaczających przechodzenie na kolejne, coraz wyższe poziomy istnienia i świadomości, zmierzające do osiągnięcia celu złożonego w procesach rozwoju. Są to:
- planetyzacja kosmosu;
- epifenomen życia na gruncie materii rozproszonej we wszechświecie (w XX wieku upatrywano ten etap w pojawieniu się bakterii i wirusów);
- hominizacja (pojawienie się człowieka);
- cefalizacja (rozwój mózgu i układu nerwowego w kierunku Homo sapiens);
- noosfera (ponadjednostkowy poziom świadomości w procesach socjalizacji).

Osią scalającą uporządkowanie tych procesów jest rozumiany teleologicznie (celowościowo) globalny dynamizm złożoności i świadomości. Kres ewolucyjnego koncentrowania świadomości stanowi jedność uczestniczących w nim podmiotów w teilhardowskim projekcie punktu „Omega”. Według Teilharda to osobowe, rozumne i dynamiczne istnienie, transcendentne wobec czasu i przestrzeni, a jednocześnie immanentne w dziejach ewolucji jako czynnik, który nadaje procesowi ewolucji dynamizm osiągania coraz wyższych poziomów.

Komentując teilhardowską wizję dziejów ewolucji kosmosu, Czesław Bartnik wskazuje na prymat podmiotu antropologicznego i zauważa, że

przedmiot wszechdziejów stanowi człowiek, który jest zarazem ich podmiotem. Przy tym człowiek jest ujmowany nie tylko jako struktura już ukształtowana i dopełniona, ale przede wszystkim jako będąca w stadium tworzenia się, niejako w ruchu. Jest to więc rozumienie człowieka jako pewnego ludzkiego procesu, który dzięki czaso-przestrzeni materialnej zmierza ku coraz dalszemu ucłowieczaniu się, ku pełnej i finalnej realizacji człowieka³⁸.

Dodajmy od siebie, że teilhardowska wizja ewolucyjnych dziejów, oparta o kompleksyfikację świadomości w jej czasowo-przestrzennych wymiarach, może dotyczyć także etapu transhumanizacji. Jednakże nie oznacza ona nowego poziomu w ewolucji bytów ożywionych, ale stanowi nowe narzędzie, pozwalające w pełni wykorzystywać potencjał świadomości i istnienia, złożony w naturze bytów biologicznych. Konsekwentnie też w tej interpretacji etapów

³⁸ Czesław BARTNIK, *Problem historii uniwersalnej w teilhardyzmie* (Lublin: Wydawnictwo Katolickiego Uniwersytetu Lubelskiego, 1972), 39, 76.

ewolucyjnych sfera aktywności AI w ludzkim świecie stanowi narzędzie wspomagające doskonalenie operatywności intelektualnej w kierunku optymalizacji potencjału ludzkiego. Wszelkie próby zastąpienia (czasem wykluczenie) biologicznego podmiotu poznania i działania strukturami wirtualnymi nie mają logicznego i etycznego uzasadnienia³⁹.

Można więc z dużym prawdopodobieństwem przyjąć stwierdzenie, że w dyskusji nad kierunkami interpretacji transhumanizmu, zwłaszcza w kontekście przejmowania przez AI coraz szerszych zakresów relacji poznawczych i operatywnych współczesnego świata, teilhardyzm może okazać się przydatnym czynnikiem dla odkrywania pozytywnego wymiaru nieuchronnego postępu technologicznego.

2.2. SPOŁECZNE I ETYCZNE WYZWANIA

W społecznym wymiarze przewidywany potencjał superinteligencji aproksymatywnie wyznaczać będzie oprócz korzyści także, przewidywalne już obecnie, zagrożenia. Ich źródło stanowi m.in. dysproporcja IQ⁴⁰. Jak wiadomo, w przypadku człowieka IQ osiąga maksymalnie wartość około 130 (w przypadku utalentowanych jednostek zwykle wynosi około 85), a przewidywanym poziomem ASI wyliczonym na 12 952. Oznacza to, że wspomniana w definicji osobliwości technologicznej „nieaktualność przewidywań” ma znaczący charakter ilościowy. Najczęściej interpretacja tej dysproporcji prowadzi do eliminacji superrozumnych podmiotów, których inteligencja nie jest osadzona w istotach żywych,

³⁹ Z podobnymi opiniami można spotkać się m.in. w publikacjach: Themistoklis TZIMAS, „The Ontology of AI”, w: Themistoklis TZIMAS, *Legal and Ethical Challenges of Artificial Intelligence from an International Law Perspective* (Cham: Springer International Publishing, 2021), 33-67, [⁴⁰ W drugiej dekadzie XXI wieku podważono ten paradygmat, wskazując, że testy na iloraz inteligencji nie są jedynymi parametrami określającymi potencjał ludzkiej inteligencji. Oprócz tego ilorazu wskazano wartość inteligencji emocjonalnej człowieka, inteligencji moralnej \(MQ\) oraz inteligencji ciała \(BQ\). Pierwsze testy IQ przeprowadzono w 1905 roku. Na wzór Skali Inteligencji Bineta-Simonaa w ciągu XX wieku ustawicznie doskonalono to narzędzie badawcze. W wyniku postępu uznano m.in. Skalę Stanfordzką, Guilforda Strukturę Intelaktu, Psychometryczne Teorie Inteligencji oraz Grupowe Testy Inteligencji. Zob. Eliezer BERKOVITS, David J. ARMOR, *Maximizing Intelligence* \(New Jersey: Transaction Publishers New Jersey, 2003\) \(kolejne wydania: 2006, 2017\).](https://books.google.pl/books?hl=pl&lr=&id=twM7EAAAQBAJ&oi=fnd&pg=P-P7&dq=Tzimas+The+Ontology&ots=RfTOAx-K37&sig=iG-TakhwOre5fSEx7Urk1RfBpyU&redir_esc=y#v=onepage&q=Tzimas;ARAI, „Can a robot pass a university entrance exam?”, 1; Eliezer YUDKOWSKY, „Friendly Artificial Intelligence”, w: <i>Singularity Hypothesis. A Scientific and Philosophical Assessment</i>, red. Amnon H. Eden, et al. (Berlin, Heidelberg: Springer, 2012), 181-195; Dennis BRAY, „Brain versus Machine”, w: <i>Singularity Hypothesis. A Scientific and Philosophical Assessment</i>, red. Amnon H. Eden, et al. (Berlin–Heidelberg: Springer, 2012), 269-280.</p></div><div data-bbox=)

i wpływa na dziejowy proces ewolucji gatunku *Homo sapiens*. Ma to m.in. miejsce w kolejnych próbach kodyfikacji działania AI w imię aksjologii uwarunkowanej historycznie. Dlatego też wysuwane są postulaty etyczne i społeczne (w tym marginalizacja) dla ukierunkowania rozwoju technologicznego, z parametrem „przyjaznej relacji do człowieka”⁴¹.

W publikacji *Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical* Luciano Floridi⁴² omawia kluczowe zagrożenia związane z wdrażaniem zasad etyki cyfrowej, także etyki sztucznej inteligencji w praktyce. Podkreśla, że po fazie formułowania zasad („co” – „what”) konieczne jest przejście do etapu ich rzeczywistego stosowania („jak” – „how”). Wskazuje on na pięć istotnych ryzyk etycznych. Są to:

1. Etyczne „zakupy” (*Ethics shopping*) – organizacje wybierają lub tworzą takie zestawy zasad etycznych, które pasują do ich dotychczasowych działań, zamiast dostosowywać działania do uznanych norm etycznych. Powstaje „rynek etyki”, który sprzyja relatywizmowi i utrudnia porównywanie standardów. Autor proponuje przyjęcie wspólnych, publicznie uzgodnionych standardów (np. Wytyczne etyczne dotyczące godnej zaufania sztucznej inteligencji w Unii Europejskiej).

2. Etyczne „upiększanie” (*Ethics bluewashing*) – organizacje deklarują przywiązanie do etyki cyfrowej jedynie dla celów wizerunkowych, bez rzeczywistych działań. Rozwiązaniem problemu, w ocenie Floridiego, jest edukacja publiczna oraz opracowanie przejrzystych wskaźników i certyfikacji etycznych praktyk.

3. Etyczne lobbingowanie (*Ethics lobbying*) – zakłada zastosowanie retoryki etycznej do opóźniania lub osłabiania regulacji prawnych w zakresie technologii cyfrowych i algorytmicznych. Rozwiązaniem problemu są precyzyjne regulacje prawne, skuteczne egzekwowanie przepisów prawa oraz rozróżnienie między autentyczną samoregulacją a manipulacją etyczną.

⁴¹ Eliezer Yudkowsky postulował rozpoczęcie badań nad stworzeniem przyjaznej sztucznej inteligencji, z uwzględnieniem wszystkich związanych z tym zagrożeń. Zwrócił uwagę, że jeśli pierwsza sztuczna inteligencja będzie przyjazna człowiekowi, będzie miała lepszy start przed wrogimi jej realizacjami. Organizacja Machine Intelligence Research Institute koncentruje się na stworzeniu takiej przyjaznej sztucznej inteligencji. Wykorzystywane w procedowaniu badawczym w tym kierunku są m.in. trzy prawa robotyki sformułowane przez Isaaca Asimova. Są one jednym z najwcześniejszych przykładów proponowanych zabezpieczeń dla maszyn myślących. W 2004 roku SIAI rozpoczął internetową kampanię pod hasłem *3 laws unsafe*, by podnieść społeczną świadomość kwestii bezpieczeństwa i niedoskonałości praw robotyki Asimova. Niektórzy anarchoprymitywiści i ekoanarchiści, jak John Zerzan i Derrick Jensen, postrzegają osobliwość jako kulminację kontroli maszyn i utratę możliwości swobodnej egzystencji poza cywilizacją. James John Bell wyraża obawy natury ekologicznej w odniesieniu do osobliwości technologicznej.

⁴² FLORIDI, „Translating Principles into Practices of Digital Ethics”, 187-188.

4. Etyczne „zrzuty” (*Ethics dumping*) – przenoszenie nieetycznych praktyk badawczych do krajów o słabszych regulacjach lub/i importowanie wyników takich badań. Autor podkreśla konieczność kontroli źródeł finansowania badań, wprowadzenie certyfikacji etycznego pochodzenia technologii oraz spójność między etyką badań a etyką konsumpcji.

5. Uchylenie się od etyki (*Ethics shirking*) – ograniczanie zaangażowania etycznego w sytuacjach, gdzie zyski z działań etycznych wydają się niewielkie. Wynika to z rozmycia odpowiedzialności prawnej i etycznej w środowisku cyfrowym. L. Floridi zaleca wzmocnienie „etyki rozproszonej odpowiedzialności” – jasnego przypisywania odpowiedzialności i konsekwencji potencjalnym współtwórcom AI.

L. Floridi podkreśla, że przejście od zasad do praktyki wymaga nie tylko wiedzy, ale i świadomości ryzyka uproszczeń oraz nadużyć. Każde z pięciu opisanych zagrożeń osłabia zaufanie społeczne do etyki sztucznej inteligencji. Zapobieganie im wymaga transparentności, edukacji, współpracy międzysektorowej oraz wypracowania jednolitych, publicznie kontrolowanych standardów etycznych w zakresie technologii cyfrowych i sztucznej inteligencji⁴³.

We Francji m.in. Narodowa Komisja Informatyki i Wolności (CNIL) ukarała w grudniu 2023 roku na kwotę 32 mln Euro firmę Amazon France za to, że wykorzystwała sztuczną inteligencję z jej „nadmiernie inwazyjnym systemem monitorowania aktywności i wydajności pracowników”⁴⁴. Te działania, w przekonaniu socjologa Juana Sebastiana Carbonella⁴⁵, odslaniają realny casus regulacji nie tyle etycznej, co raczej politycznej. I z pewnością właśnie w obszarze polityki należy w przyszłości oczekiwać źródeł zagrożeń społecznych AI (zwłaszcza na poziomie osobliwości technologicznej). Potwierdzenie realności tego wątku na szczelbu Unii Europejskiej w jej Komitecie Ekspertów ds. Sztucznej Inteligencji stanowi krytyczna ocena dokumentu o Wytycznych Etyki⁴⁶.

⁴³ FLORIDI, „Translating Principles into Practices of Digital Ethics”, 187-188. Zob. Małgorzata GRUCHOŁA, „Interreligijna współpraca etyków wokół zagadnień sztucznej inteligencji”, w: *Sztuczna inteligencja – rzeczywiste wyzwania. Nowe technologie w komunikacji społecznej*, red. Katarzyna Drąg (Tarnów: Wydawnictwo Biblos, 2024), 97-122.

⁴⁴ Zob. Juan Sebastián CARBONELL, „Régulation de l’IA: «Le problème n’est pas éthique, ni technologique, il est politique»”, dostęp 21.11.2024, <https://www.la-croix.com/a-vif/regulation-de-l-ia-le-probleme-nest-pas-ethique-ni-technologique-il-est-politique-20240521>.

⁴⁵ CARBONELL, „Régulation de l’IA: «Le problème n’est pas éthique, ni technologique, il est politique»”.

⁴⁶ Piszą na ten temat m.in. Eugenia STAMBOLIEV, Tim CHRISTIAENS [11 lutego 2024], „How empty is Trustworthy AI? A discourse analysis of the Ethics Guidelines of Trustworthy AI,” [Jak pusta jest wiarygodna AI? Analiza dyskursu wytycznych etycznych godnych zaufania sztucznej

O tym, że osobliwość technologiczna rozwija się w sposób ciągły i nieunikniony (zwłaszcza w przestrzeni aktywności ludzkiej), wiadomo było już w połowie XX wieku. W XXI wieku, chcąc zrównoważyć ten rozwój, niektórzy naukowcy proponują rozwój alternatywny, określane mianem „ludzkiej osobliwości”⁴⁷. Ryzyko potraktowania bytu ludzkiego przez superinteligentne technologie, poprzez możliwą redukcję roli biologicznej inteligencji do poziomu zwierząt (co zachodzi w praktyce podczas zbliżania się do osobliwości technologicznej) wyznacza realną granicę zachowania tożsamości humanistycznej. Przekroczenie tej granicy oznaczać może zastąpienie świata humanistycznego światem apersonalnej samotności.

Proces doskonalenia AI i jej wdrożeń w życie społeczne nie uwolni człowieka od konieczności nieustannego postępu rozumienia świata „obok nas” – pod wieloma aspektami istniejącego już na wysokim poziomie w ramach intelektualnych relacji ilościowych. Wciąż także istnieć będzie zagrożenie związane z ryzykiem egzystencjalnym. Może go spowodować błąd w funkcjonowaniu algorytmu, pomijanie tego ryzyka na etapie badań nad rozwojem AI, a także świadome „ocenzurowanie” technologii ze względów politycznych⁴⁸. Czy wówczas wystarczy odkładanie momentu rozejścia się osobliwości technologicznej i tożsamości

inteligencji], *Critical Policy Studies* 19, nr 1 (2024): 39-56, <https://doi.org/10.1080/19460171.2024.2315431>.

⁴⁷ Np. Cédric Villani, francuski matematyk i nauczyciel akademicki, laureat m.in. Medalu Fieldsa w 2010 i dyrektor Instytutu Henri-Poincaré (2009-2017). Omówienie jego wkładu w analizę zagadnień technologii, zob. Cédric VILLANI, „Le problème, ce n’est pas que l’IA nous remplace, c’est notre monde de solitude”, dostęp 21.11.2024, <https://www.la-croix.com/a-vif/cedric-villani-le-probleme-ce-n-est-pas-que-lia-nous-remplace-cest-notre-monde-de-solitude-20240416>. Podobne stanowisko reprezentują niektórzy badacze sztucznej inteligencji w Japonii. Np. badaczka Noriko Arai swój projekt z 2017 roku skoncentrowała na zagadnieniu „Czy robot może dostać się na uniwersytet?”. Zakończyła go stwierdzeniem: „Myślenie, że mój robot może wejść na uniwersytet, nie ma sensu. To nie jest człowiek. On nie jest matematykiem. Z definicji tylko ludzie mogą dostać się na uniwersytet”. Cyt. za: Cédric VILLANI, „Le problème, ce n’est pas que l’IA nous remplace, c’est notre monde de solitude”, dostęp 21.11.2024, <https://www.la-croix.com/a-vif/cedric-villani-le-probleme-ce-n-est-pas-que-lia-nous-remplace-cest-notre-monde-de-solitude-20240416>. Oryginalny wykład autorki na stronie: Noriko ARAI, „Can a robot pass a university entrance exam?”, dostęp 21.11.2024, https://www.ted.com/talks/noriko_arai_can_a_robot_pass_a_university_entrance_exam?subtitle=en.

⁴⁸ Nick BOSTROM, „*Existential Risks*”, *Journal of Evolution and Technology* 9 (2001):1-36. Zwraca on uwagę m.in. na obszary ryzyka, jakimi są przyszłościowe technologie antropogenetyczne (zaawansowane formy syntetycznej biologii, nanotechnologii i superinteligentnych maszyn). Nieodłącznym uwarunkowaniem w eliminacji różnych form ryzyka egzystencjalnego, jego zdaniem, może okazać się faktor ekonomiczny znacznie podwyższający koszty z badaniami dla potencjalnego uniknięcia tego ryzyka. O ryzyku zagrożeń technologicznych w kontekście AI pisał

humanistycznej z jej etycznym i dziejowo-ewolucyjnym fundamentem? Czy wypada zgodzić się – w imię jakich idei – na aksjologię, w której niestabilność wolności (jako najwyższej wartości) zostanie zastąpiona przez komfort dobrego samopoczucia? Czy wreszcie osiągnięcie poziomu osobliwości technologicznej w transhumanistycznym świecie ubezwłasnowolni ludzki podmiot dziejów do tego stopnia, że dokona on autodestrukcji w realnym procesie alienacji, którą filozofowie XX wieku przewidywali jako *malum imminens* w historycznych relacjach podmiotu i przedmiotu ludzkich działań?⁴⁹

3. PRAKSEOLOGICZNE UWARUNKOWANIA

Na przedłużeniu doświadczeń z wprowadzaniem nowych odkryć w przeszłości w niektórych dziedzinach dotyka granicy bezpieczeństwa gatunku *Homo sapiens*. Pierwszym przejawem konfliktu między odkryciami naukowymi, warunkującymi nowe technologie, a środowiskiem człowieka w pierwszej połowie XX wieku stało się kontrolowane rozszczepienie atomu i użycie tego odkrycia w celach militarnych. Polityczne ograniczenia rozwoju tego sektora technologii pozwoliły podjąć wyzwania naukowe w obszarze mikro-świata (cząstek elementarnych) i makro-świata (w eksploracji kosmosu). W efekcie znacznie poszerzono rozumienie środowiska człowieka i uświadomiono potrzebę odpowiedzialności za roztropne wdrażanie technologii. Spektakularne rozwiązania dotyczą m.in. kierowania się w XXI wieku w badaniach naukowych i wdrażania ich efektów różnymi prawami.

na początku XXI wieku m.in. James John BELL, „Technotopia and the Death of Nature: Clones, Supercomputers, and Robots”, *Clamor* 18 (2003):9-13.

⁴⁹ Panoramę tego wątku filozoficznej analizy superinteligentnego, apersonalnego podmiotu poznania omawia m.in. Szymon WRÓBEL, „Historyczność albo alienacja. Od Georga Wilhelma Friedricha Hegla do Marka Siemka”, *Ethos* 30, nr 1/117 (2017):156-189, https://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.ojs-issn-0860-8024-year-2017-volume-30-issue-1_117_article-5219/c/5219-5051.pdf. O odniesieniach perspektywy ewolucyjnej do hipotetycznego, nowego etapu transhumanizmu, zob. m.in. publikacje Raymonda Kurzweila (ur. 1948, absolwent Martin Van Buren High School w 1965 roku oraz Massachusetts Institute of Technology w 1970 roku): Raymond KURZWEIL, *The Singularity Is Near* (New York: Penguin Group, 2005); Raymond KURZWEIL, *How to Create a Mind* (London: Penguin Books, 2013); Raymond KURZWEIL, *Transcend: Nine Steps to Living Well Forever* (New York: Harmony/Rodale/Convergent, 2010); Ray KURZWEIL, Terry GROSSMAN, *Fantastic Voyage: Live Long Enough to Live Forever* (New York: Publication Data New York, 2005); Raymond KURZWEIL, *The Age of Spiritual Machines* (London: Penguin Books, 2000); Raymond KURZWEIL, *The Age of Intelligent Machines* (Cambridge: MIT Press, 1992); Bill HIBBARD, „Super-intelligent machines”, *Siggraph Computer Graphics* 35, nr 1 (2001):13-15.

1) Uwzględnienie postępu w technologiach dotyczących procesów kwantowych, ważnych z punktu widzenia perspektyw komputerów kwantowych. Tutaj nadal pozostają ważne spostrzeżenia Nielsa Bohra, sformułowane w jego „kopenhaskiej interpretacji mechaniki kwantowej”. Według niego, na najniższym szczeblu tych procesów uwzględnić należy fakt, że cząstka elementarna jednocześnie istnieje we wszystkich stanach, będąc „superpozycją wszystkich możliwych stanów”. Znamienne jest w tym względzie stwierdzenie Bohra, że „gdy chodzi o atomy, możemy używać języka tylko jak w poezji”⁵⁰. To stwierdzenie jest w praktyce rozwinięciem zasady nieoznaczoności Wernera Heisenberga, zgodnie z którą dualizm korpuskularno-kwantowy zakłada nieodłączną niepewność a zarazem komplementarność w strukturze przyrody⁵¹.

2) Monitorowanie badań nad ciemną materią kosmosu, zdaniem jej odkrywcy, Jana Hendrika Oorta (1900–1992), odpowiedzialną za znaczącą część masy i energii wszechświata. Obecnie przyjmuje się duże znaczenie cząstek subatomowych w procesach warunkujących istnienie i działanie ciemnej materii.

3) Monitorowanie postępów w badaniach nad teorią identyczności mózgu-umysłu oraz prakseologicznych postulatów współczesnej kognitywistyki, w oparciu o sugestie Jean Piageta (1896–1980), Herberta Simona (1916–2001) i Noama Chomsky’ego (ur. 1928).

4. KONKLUZJA

Etyczne wyzwania wobec coraz szerszej implementacji współczesnych technologii, w tym AI, mogą zyskać pozytywne przesłanki m.in. w modelu „przyjaznym człowiekowi”. Pewną sugestią stwarza XX-wieczny teilhardyzm, godzący procesy ewolucji z aksjologią kręgu cywilizacji grecko-rzymskiej, dla której naczelną wartością pozostaje dobro człowieka. Filozoficzne i prakseologiczne przesłanki korzystania z nowych technologii w znacznej mierze określają też społeczne poczucie bezpieczeństwa i przewidywanych korzyści dla osiągnięcia dobra wspólnego cywilizacji przyszłości, jako efektu rzetelnych i zgodnych z etyką odkryć naukowych. Słusznie stwierdza Luciano Floridi, że etyka powinna stanowić centralny komponent refleksji nad technologią. Etyka – w przeciwieństwie do prawa – nie tylko określa, co jest dozwolone, ale

⁵⁰ Robert ARP, red., *1001 idei, które zmieniły nasz sposób myślenia* (Poznań: Wydawnictwo Publicat SA, 2017), 670.

⁵¹ ARP, *1001 idei, które zmieniły nasz sposób myślenia*, 676. Zob. paralelny do wizji Teilharda de Chardin model wszechświata w rozwoju autorstwa Detlev Van Ravenswaay/Science Photo Library, 677.

wskazuje, co jest moralnie pożądane dla dobra wspólnego. Floridi formułuje pojęcie infosfery, w której wszystkie byty informacyjne współistnieją w relacjach wzajemnej odpowiedzialności. W tym kontekście rozwój AI oraz świadomości sieciowej wymaga równowagi między innowacją a moralnym namysłem. Wprowadzając pojęcia *ethics shopping* i *ethics bluwashing*, opisuje zjawisko instrumentalizacji etyki w praktykach korporacyjnych. W pierwszym przypadku organizacje wybierają te zasady etyczne, które najlepiej wspierają ich interesy, w drugim – używają języka etyki do budowania wizerunku, nie podejmując rzeczywistych działań. Zjawiska te prowadzą do deflacji znaczenia etyki AI i utraty społecznego zaufania do inicjatyw technologicznych. L. Floridi zwraca uwagę, że etyka powinna być rozumiana nie jako zestaw deklaracji, lecz jako proces translacji zasad w praktykę. Przekład zasad etycznych na procedury zarządzania, audyt algorytmiczny i mechanizmy odpowiedzialności jest warunkiem koniecznym dla rozwoju „godnej zaufania” sztucznej inteligencji (ang. *trustworthy AI*). Bez tego istnieje ryzyko, że etyka stanie się – jak pisze L. Floridi – jedynie „retoryką moralnej legitymizacji innowacji”⁵². Z kolei Max Tegmark⁵³ zauważa, że ewolucja inteligencji, zarówno biologicznej, jak i sztucznej, powinna być postrzegana jako wspólna droga ku „Life 3.0”, etapowi, w którym świadomość i informacja stają się podstawowymi czynnikami istnienia. Z kolei Nick Bostrom⁵⁴ ostrzega przed tzw. scenariuszami katastroficznymi związanymi z rozwojem superinteligencji. Jednocześnie część badaczy, m.in. Luciano Floridi⁵⁵ oraz Max Tegmark⁵⁶, zwraca uwagę, że sama kategoria „superinteligencji” może być antropocentryczna, zakłada bowiem porównanie sztucznego umysłu z ludzkim jako jedynym wzorcem odniesienia. W tym sensie superinteligencja nie musi być „lepszą” od ludzkiej, lecz po prostu inaczej ukierunkowana poznawczo. Z perspektywy etyki informacji L. Floridiego⁵⁷, bardziej owocne jest rozumienie inteligencji nie jako hierarchii, lecz wielopoziomowej współzależności systemów poznawczych: ludzkich, maszynowych i środowiskowych. Oznacza to przesunięcie od antropocentrycznego modelu „kontroli” ku relacyjnemu modelowi współistnienia w infosferze.

Wspólnym wątkiem krytycznych analiz jest potrzeba opracowania etyki relacyjnej, która przekracza opozycję człowiek – maszyna i wpisuje rozwój

⁵² Luciano FLORIDI, *The Ethics of Information* (Oxford: Oxford University Press, 2013), 12.

⁵³ Max TEGMARK, *Life 3.0: Being human in the age of artificial intelligence* (New York: Vintage, 2018), 6.

⁵⁴ BOSTROM, *Superintelligence*, 40-43.

⁵⁵ FLORIDI, *The Ethics of Information*, 13.

⁵⁶ TEGMARK, *Life 3.0: Being human in the age of artificial intelligence*, 7.

⁵⁷ FLORIDI, *The Ethics of Information*, 13.

technologii w kontekst ekologiczny oraz społeczny. M. Tegmark⁵⁸ wskazuje, że przyszłość życia inteligentnego („Life 3.0”) zależy od umiejętności współistnienia różnych form świadomości w ramach jednej infosfery. W tym ujęciu celem filozofii AI nie jest zastąpienie człowieka, lecz rozszerzenie pola świadomości, które obejmuje współdziałanie, współodpowiedzialność i wspólnotę poznawczą. Tylko taka etyka, zakorzeniona w krytycznej refleksji nad relacją technologii, człowieka i świata, może przeciwdziałać uproszczeniom ideologicznym transhumanizmu i zapewnić zrównoważony rozwój świadomości globalnej.

BIBLIOGRAFIA

- ARAI, Noriko. „Can a robot pass a university entrance exam?”. Dostęp 21.11.2024. https://www.ted.com/talks/noriko_arai_can_a_robot_pass_a_university_entrance_exam?subtitle=en.
- ARP, Robert. Red. „1001 idei, które zmieniły nasz sposób myślenia”. Poznań: Wydawnictwo Publicat SA, 2017.
- BARRETT, Anthony M., Seth D. BAUM. „A model of pathways to artificial superintelligence catastrophe for risk and decision analysis”. *Journal of Experimental & Theoretical Artificial Intelligence* 29, nr 2 (2017):397-414.
- BARTNIK, Czesław. *Problem historii uniwersalnej w teilhardyzmie*. Lublin: Wydawnictwo Katolickiego Uniwersytetu Lubelskiego, 1972.
- BARTNIK, Czesław. *Teilhardowska wizja dziejów*. Lublin: Wydawnictwo Katolickiego Uniwersytetu Lubelskiego, 1975.
- BELL, James John. „Technotopia and the Death of Nature: Clones, Supercomputers, and Robots”. *Clamor* 18 (2003):9-13.
- BERKOVITS Eliezer, David J. ARMOR. *Maximizing Intelligence*. New Jersey: Transaction Publishers, 2003.
- „Bill Gates przedstawił 5 zagrożeń związanych ze sztuczną inteligencją”. Dostęp 11.01.2025. <https://forsal.pl/lifestyle/technologie/artykuly/8781146,bill-gates-przedstawil-5-zagrozen-zwiazanych-ze-sztuczna-inteligencja.html>.
- BOSTROM, Nick. „Existential Risks”. *Journal of Evolution and Technology*” (2013).
- BOSTROM, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014.
- BRAY, Dennis. „Brain versus Machine”. W: *Singularity Hypothesis. A Scientific and Philosophical Assessment*, red. Amnon H. Eden, et al., 269-280. Berlin–Haidelberg: Springer, 2012.
- CARBONELL, Juan S. „Régulation de l’IA: Le problème n’est pas éthique, ni technologique, il est politique”. Dostęp 21.11.2024. <https://www.la-croix.com/a-vif/regulation-de-l-ia-le-probleme-nest-pas-ethique-ni-technologique-il-est-politique-20240521>.

⁵⁸ TEGMARK, *Life 3.0: Being human in the age of artificial intelligence*, 7.

- CHARDIN, Teilhard Pierre de. *Moja wizja świata*, tłum. Mieczysław Tazbir. Warszawa: Instytut Wydawniczy, 1987.
- CHARDIN, Teilhard Pierre de. *The phenomenon of man*. New York: Harper & Row, 1959.
- CHARDIN, Teilhard Pierre de. *Zarys wszechświata personalistycznego*, tłum. Konrad Waloszczyk. Warszawa: Instytut Wydawniczy Pax, 1984.
- CHOJNOWSKI, Maciej. *Etyka sztucznej inteligencji*. Warszawa: Centrum Etyki Technologii Instytutu Humanites, 2002. Dostęp 05.10.2025. https://ethicstech.eu/wp-content/uploads/2022/12/ESIW-v2.0_FINAL.pdf.
- CLARKE, Arthur C. *Profiles of the Future*. London: Indigo, 2000.
- Conseil de l'Union Européenne. „Législation sur l'intelligence artificielle (IA): le Conseil donne son feu vert définitif aux premières règles mondiales en matière d'IA”. Dostęp 23.05.2024. <https://www.google.com/url?sa=t&source=web&rct=j&opi=89978449&url=https://www.consilium.europa.eu/fr/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/&ved=2ahUKEwjeqMjOu-WKAX-VDAtsEHaYtA0YQFnoECBUQAQ&usq=AOvVaw1lh99CCmPV3utzHCawTQla>.
- DUCH, Włodzisław. „Ludzkość+ za rogiem? Prof. Duch o transhumanizmie: sztuczna inteligencja zmienia wszystko”. Dostęp 24.02.2025. <https://www.tvp.info/65670974/sztuczna-inteligencja-zmienia-wszystko>.
- FERRANDO, Francesca. *Philosophical posthumanism*. London: Bloomsbury Academic, 2019.
- FLORIDI, Luciano. „Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical”. *Philosophy & Technology* 32 (2019):185-193. <https://doi.org/10.1007/s13347-019-00354-x>.
- FLORIDI, Luciano. *The Ethics of Information*. Oxford: Oxford University Press, 2013.
- GOOD, Irving. *Good Thinking the Foundation of Probability and its Applications*. Minneapolis: University of Minesota Press, 1983.
- GRUCHOŁA, Małgorzata. „Interreligijna współpraca etyków wokół zagadnień sztucznej inteligencji”. W: *Sztuczna inteligencja – rzeczywiste wyzwania. Nowe technologie w komunikacji społecznej*, red. Katarzyna Dąg, 97-122. Tarnów: Wydawnictwo Biblos, 2024.
- HIBBARD, Bill. „Super-intelligent machines”, *Siggraph Computer Graphics* 35, nr 1 (2001):13-15. https://www.ted.com/talks/noriko_arai_can_a_robot_pass_a_university_entrance_exam?subtitle=en.
- HUEBNER, Jonathan. „A Possible Declining Trend for Worldwide Innovation”. *Technological Forecasting & Social Change* 72 (2005):980-986. <https://www.sciencedirect.com/science/article/abs/pii/S0040162505000235>.
- Intelligence artificielle: l'Union européenne pose des limites inédites*. Dostęp 23.05.2024. <https://www.la-croix.com/international/intelligence-artificielle-l-union-europeenne-pose-des-limites-inedites-20240520>.

- Intelligence artificielle: pourquoi les lauréats du Nobel de physique sonnent l'alarme*. Dostęp 20.12.2024. <https://www.la-croix.com/international/intelligence-artificielle-pourquoi-les-lauréats-du-nobel-de-physique-sonnent-l-alar-me-20241009>.
- KARAS, Marcin. *Historiozofia Teilharda de Chardin wobec tradycyjnej myśli chrześcijańskiej*. Kraków: Księgarnia Akademicka, 2012.
- KURZWEIL, Raymond i Terry GROSSMAN. *Fantastic Voyage: Live Long Enough to Live Forever*. New York: Publication Data, 2005.
- KURZWEIL, Raymond. *How to Create a Mind*. London: Penguin Books, 2013.
- KURZWEIL, Raymond. *The Age of Intelligent Machines*. Cambridge: MIT Press, 1992.
- KURZWEIL, Raymond. *The Age of Spiritual Machines*. London: Penguin Books, 2000.
- KURZWEIL, Raymond. *The Singularity is Near*. New York: Viking, Penguin Group, 2005.
- KURZWEIL, Raymond. *Transcend: Nine Steps to Living Well Forever*. New York: Harmony/Rodale/Convergent, 2010.
- LUBAS-BARTOSZYŃSKA, Regina. „Zrehabilitowany Pierre Teilhard de Chardin”. Dostęp 04.12.2024. <https://kongreskk.pl/2023/09/30/zrehabilitowany-pierre-teilhard-de-chardin/>.
- MORE, Max, Natasha VITA-MORE. Red. *The transhumanist reader: Classical and Contemporary essays on the science, technology, and philosophy of the human future*. Malden–Oxford: John Wiley & Sons Ltd. Inc., 2013.
- SALLY, Ade. „A big step forward in a project that aims for thinking chips”. Dostęp 07.12.2024. <https://spectrum.ieee.org/ibm-unveils-a-new-brain-simulator>.
- STAMBOLIEV, Eugenia, Christiaens TIM. „How empty is Trustworthy AI? A discourse analysis of the Ethics Guidelines of Trustworthy AI”. *Critical Policy Studies* 19, nr 1 (2024):39-56. <https://doi.org/10.1080/19460171.2024.2315431>.
- TEGMARK, Max. *Life 3.0: Being human in the age of artificial intelligence*. New York: Vintage, 2018.
- TZIMAS, Themistoklis. „The Ontology of AI”. W: *Legal and Ethical Challenges of Artificial Intelligence from an International Law Perspective*. red. Themistoklis Tzimas, 33-67. Cham: Springer International Publishing, 2021. https://books.google.pl/books?hl=pl&lr=&id=twM7EAAAQ-BAJ&oi=fnd&pg=PP7&dq=Tzimas+The+Ontology&ots=RfTOAx-K37&sig=iG-TakhwOre5fSEx7Urk1RfBpyU&redir_esc=y#v=onepage&q=TzimasTheOntology&f=false.
- ULAM, Stanisław. „Tribute to John von Neumann”. *Bulletin of the American Mathematical Society* 64, nr 2-3 (1958):1-49.
- URBAN, Tim. „The AI Revolution: The Road to Superintelligence”. Dostęp 15.11.2024. <https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html>.
- VINGE, Vernor. „*The Coming Technological Singularity: How to Survive in the Post-Human Era*”. *Referat na sympozjum NASA Lewis Center 30.03.1993*. Dostęp 24.11.2024. <https://edoras.sdsu.edu/~vinge/misc/singularity.html>.
- VINGE, Vernor. „The coming technological singularity: How to survive in the post-human era”. *Science fiction criticism: An anthology of essential writings* 81 (1993):352-363.

- VINGE, Vernor. *Deepness in the Sky*. New York: Tor Books Editor New York, 1999.
- VINGE, Vernor. *Marooned in Realtime*. New York: Bluejay Books/St. Martin's Press, 1986.
- VINGE, Vernor. *Otchłań w niebie*. Warszawa: Wydawnictwo Prószyński i S-ka, 2003.
- VINGE, Vernor. *True Names and the Cyberspace Frontier*. New York: Tor Publishing, 2001.
- VIDAL, Clément. „What is the noosphere? Planetary superorganism, major evolutionary transition and emergence”. *Systems Research and Behavioral Science* 41, nr 4 (2024): 614-622.
- VILLANI, Cédric. „Le problème, ce n'est pas que l'IA nous remplace, c'est notre monde de solitude”. Dostęp 21.11.2024. <https://www.la-croix.com/a-vif/cedric-villani-le-probleme-ce-n-est-pas-que-lia-nous-remplace-cest-notre-monde-de-solitude-20240416>.
- WRÓBEL, Szymon. „Historyczność albo alienacja. Od Georga Wilhelma Friedricha Hegla do Marka Siemka”. *Ethos* 30, nr 1 (2017):156-189.
- YUDKOWSKY, Eliezer. „Friednly Artificial Intelligence”. W: *Singularity Hypothesis. A Scientific and Philosophical Assessment*, red. Amnon H. Eden, et al., 181-195. Berlin–Haidelberg: Springer, 2012.
- ZWART, Hub. *Continental philosophy of technoscience*. Cham: Springer International Publishing, 2021.

SPOŁECZNE WYZWANIA OSOBLIWOŚCI TECHNOLOGICZNEJ.
KOMPLEKSYFIKACJA RZECZYWISTOŚCI HUMANO-WIRTUALNEJ

STRESZCZENIE

Współczesne technologie rozwijają się głównie pod wpływem motywacji poznawczych, ekonomicznych, kulturowych. Wyzwaniem społecznym staje się bilans korzyści i ewentualnych zagrożeń dla człowieka. Stąd potrzeba uwzględnienia kryterium etycznego w procesie ich odkrywania i wykorzystania. Głównie dotyczy to zrównoważonego rozwoju przy zachowaniu cywilizacyjnych paradygmatów aksjologicznych, ukształtowanych na gruncie cywilizacji grecko-rzymskiej. Oznacza to pozytywny stosunek do postępu technologicznego w imię zachowania prymatu antropologicznego. Jednym z filozoficznych kontekstów na tej drodze może być XX-wieczny teilardyzm.

Słowa kluczowe: antropocentryzm; etyka postępu cywilizacyjnego; osobliwość technologiczna; Pierre Teilhard de Chardin; sztuczna inteligencja; technologie współczesne; teilardyzm

SOCIAL CHALLENGES OF TECHNOLOGICAL SINGULARITY.
THE COMPLEXIFICATION OF THE HUMANO-VIRTUAL REALITY

SUMMARY

Modern technologies develop primarily under the influence of cognitive, economic, and cultural motivations. A key social challenge is balancing the benefits and potential threats to humanity. Therefore, the ethical criterion must be considered in the process of discovering and utilizing these technologies. This primarily concerns sustainable development while maintaining the axiological paradigms of civilization, shaped within the Greco-Roman tradition. This implies a positive attitude towards technological progress in the name of preserving the anthropological primacy. One of the philosophical contexts relevant to this approach may be 20th-century Teilhardism.

Keywords: anthropocentrism; artificial intelligence; ethics of civilizational progress; modern technologies; Pierre Teilhard de Chardin; technological singularity; Teilhardism