

Krzysztof Wieczorek

ORCID 0000-0002-7987-168X

Uniwersytet Śląski

## Orszak wdzięcznych marionetek, czyli maszyna jako urzekający potwór

Bliskie związki między literaturą a filozofią mają w naszej kulturze długą tradycję. U samych początków greckiej myśli filozoficznej, zanim jeszcze ukształtowały się specyficzne: język, styl i forma wyrazu, charakterystyczne dla umiłowania mądrości, antyczni myśliciele posługiwali się formami typowymi dla tekstów literackich, jak poemat (np. *Peri physeos* Parmenidesa) czy dialog<sup>1</sup>. Dzieła literackie z kolei, również od czasów starożytnych, często – by nie powiedzieć: niemal zawsze – zawierają rozważania osnute wokół problemów filozoficznych bądź też inspirujące dla filozofów. Jak zauważa Agata Skała, „już starożytne teksty pokazują niezwy-

**Prof. dr hab. Krzysztof T. Wieczorek** – pracuje w Instytucie Filozofii na Wydziale Humanistycznym Uniwersytetu Śląskiego w Katowicach. Kierunki aktualnie prowadzonych badań naukowych to: historia filozofii XX wieku (zwłaszcza w obszarze niemieckojęzycznym), fenomenologia, filozofia spotkania i dialogu, wybrane problemy teorii poznania i filozofii języka, ontologia i epistemologia wartości, filozofia humoru, filozofia religii, tanatologia, teoria decyzji z elementami logiki, prakseologii i psychologii, filozofia wychowania, obecność zagadnień i metod filozofii w życiu publicznym oraz w dyskursie medialnym i politycznym. Jest autorem ponad dwustu prac naukowych, w tym czterech książek: *Dwie filozofie spotkania. Konfrontacja myśli J. Tischnera i A. Nowickiego*, Katowice 1990; *Levinas a problem metafizyki*, Katowice 1992; *Wieczne powroty Fryderyka Nietzschego*, Katowice 1998; *Spory o przedmiot poznania*, Katowice 2004.

<sup>1</sup> O którym Rudolf Hirzel pisał: „Jako samodzielne dzieło literatury, dialog jest w ścisłym słowa znaczeniu rozważaniem w formie rozmowy” (R. Hirzel, *Der Dialog: Ein literarhistorischer Versuch*, Leipzig 1895, s. 7).

kłą łączność nauki (*resp.* filozofii) z poezją”<sup>2</sup>. Zofia Zarębianka natomiast pisze: „relacja pomiędzy filozofią a literaturą ma najpewniej charakter stopniowalny – od związków pośrednich [...] po warianty odniesień bardzo ścisłych. Te ostatnie dają się rozpatrywać [...] w dwu zasadniczych opcjach uwarunkowanych ustawieniem wektora wzajemnych związków, a więc w postaci literatury jako filozofii oraz w wersji filozofii jako literatury. Zaznaczyć też trzeba, iż wymienione opcje skądinąd niekiedy utożsamiają się z sobą i dają się odróżnić chyba przede wszystkim podług kryterium zewnętrznego w pewnym sensie, mianowicie pierwszej profesji autora”<sup>3</sup>. Jako wyrazisty przykład bliskich, wręcz intymnych związków filozofii i literatury badaczka przywołuje Platoński *Sympozjon*: „wolno chyba przyjąć, iż *Uczta* Platona, klasyczne dzieło filozoficzne, z uwagi na swoją formę i nasycenie literackością, przynależy zarazem do literatury, pozostając wszakże przede wszystkim tekstem filozoficznym, [zatem] realizuje ten wariant, w którym dzieło filozoficzne jest zarazem dziełem literackim”<sup>4</sup>.

Nic więc dziwnego, że korzystając z istnienia „strefy Schengen” obejmującej wszystkie przejścia graniczne między filozofią a literaturą, przedstawiciele obu profesji często przekraczają owe granice, spodziewając się odkryć po drugiej stronie przydatne i wartościowe znaleziska. Niniejszy tekst jest owocem tego rodzaju wycieczki. Autor, z zawodu i z zamiłowania filozof, penetruje zawartość wybranych tekstów literackich, a odnalezione tam motywy poddaje filozoficznej analizie. Podziela bowiem zdanie Marii Szyszkowskiej: „Bogactwo wiedzy o człowieku jest zawarte w literaturze pięknej. Stanowi ona uzupełnienie filozofii człowieka”<sup>5</sup>. Biorąc pod uwagę problematykę poruszaną w artykule, wypadałoby dodać, że przedmiotem zainteresowania filozofa poszukującego inspiracji

<sup>2</sup> A. Skała, *Przedmowa*, w: *Filozoficzne aspekty literatury: horyzonty metodologiczne i interpretacje*, red. A. Skała, Lublin 2018, s. 7.

<sup>3</sup> Z. Zarębianka, *Filozofia wobec literatury. Literatura wobec filozofii. Warianty wzajemnych odniesień. Rekonesans*, „*Filo-Sofija*” 16 (2016) nr 34/1, s. 144.

<sup>4</sup> Z. Zarębianka, *Filozofia wobec literatury...*, dz. cyt., s. 144.

<sup>5</sup> M. Szyszkowska, *Literatura piękna jako źródło refleksji filozoficznych oraz wiedzy filozoficznej*, „*Zeszyty Naukowe Politechniki Śląskiej. Seria: Organizacja i Zarządzanie*” 2013 nr 65, s. 414.

w tekstach literackich może być nie tylko człowiek i jego sprawy, lecz również – na przykład – fascynujący świat relacji między człowiekiem a czymś, co w specyficzny sposób nie-ludzkie: coraz bardziej do nas podobną (aż do przekroczenia granic zaskakującej i niepokojącej „doliny niesamowitości”<sup>6</sup>) maszyną.

Człowiek – w ramach czynienia sobie ziemi poddaną – wytwarza coraz bardziej złożone narzędzia i maszyny<sup>7</sup>. Do końca czasów nowożytnych między maszyną a jej konstruktorem przebiegała ostra granica: maszyna to rzecz, martwa i bezmyślna, człowiek – istota żywa i świadoma. Gdzieś w tle jednak czaił się i skradał odwieczny mit „ożywionej maszyny”<sup>8</sup>, związany z niewypowiedzianym głośno, lecz obecnym w ludzkiej duszy pragnieniem dorównania Bogu w dziele stworzenia. Wedle niezwykle sugestywnego i inspirującego biblijnego opisu (zob. Rdz 1, 26–29 i Rdz 2, 7–8), Bóg stworzył człowieka na swój obraz i podobieństwo, jako ukoronowanie stworzenia i swe ulubione dzieło. Stworzył najpierw ciało – przedmiot martwy, ulepiony z gliny i prochu ziemi; następnie ożywił, tchnąwszy weń duszę, i natychmiast pokochał, jednakże (jak sądzono w dawnych czasach, gdy bogobojność oznaczała lęk przed Bożym gniewem) pod warunkiem posłusznego wypełniania Bożej woli. Czyż w tej narracji nie kryje się gotowy scenariusz działań dla człowieka – następcy Boga w dziele stworzenia?

Najwcześniejsze, baśniowo-mitologiczne projekty ożywionej maszyny bądź też sztucznego człowieka (Golem, Galatea)<sup>9</sup> zakładały taki

<sup>6</sup> Por. M. Mori, *The Uncanny Valley*, „IEEE Robotics & Automation Magazine” 19 (2012) no. 2, s. 98–100, doi: 10.1109/MRA.2012.2192811.

<sup>7</sup> Według Samuela Johnsona Benjamin określał człowieka jako „a tool-making animal” (zwierzę wytwarzające narzędzia) (*The life of Samuel Johnson* by J. Boswell, vol. 3, Oxford–London 1826, s. 218–219).

<sup>8</sup> Przynależny do opisaney przez Junga kulturowej sfery nieświadomości zbiorowej. Por. P. Lechowski, *Filozofia mitu jako wzajemne oddziaływanie oczekiwania i pamięci*, „Philosophical Discourses. Prace Naukowe Uniwersytetu Humanistyczno-Przyrodniczego im. Jana Długosza w Częstochowie” 1 (2019), s. 275–302; L. Mumford, *Mit maszyny*, t. 1, tłum. M. Szczubiałka, Warszawa 2012.

<sup>9</sup> Zob. B. Kos, R. Żebrowski, *Golem*, w: *Polski słownik judaistyczny: dzieje, kultura, religia, ludzie*, t. 1, red. Z. Borzymińska, R. Żebrowski, Warszawa 2003; G. G. Scholem, *Wyobrażenie golema*

właśnie, sekwencyjny scenariusz: najpierw powstaje martwe ciało, w które następnie, w wyniku działania innych prawideł (należących do cudownego, nadprzyrodzonego porządku) wstępuje pierwiastek życia i świadomości. W bliższych nam czasach rolę boskiego, duchowego animatora „ożywionych maszyn” stopniowo przejęła nauka. Zrazu, w fazie przejściowej, miała to być nauka ezoteryczna, sytuująca się na pograniczu magii, jak w przypadku działań przypisywanych m.in. hrabiemu Cagliostro<sup>10</sup>. Kiedy jednak na dobre dokonała się nowożytna rewolucja naukowa, pociągając za sobą niepomierny wzrost twórczych możliwości nauk empiryczno-przyrodniczych, pojawiła się w ich obrębie idea eksperymentalnego rozwiązania zagadki życia<sup>11</sup>. Miało się to dokonać na drodze sztucznego wytworzenia żywej (a docelowo – również inteligentnej, dorównującej człowiekowi) istoty metodą stopniowej aproksymacji cech<sup>12</sup>. Obowiązywało następujące założenie: będziemy tworzyli konstrukcje, zachowujące się „jak żywe” w coraz bardziej poszerzonym zakresie parametrów behawioralnych, aż w końcu przekroczymy „odolnie” nieuchwytną granicę między tym, co „jak żywe”, a tym, co żywe *tout court*, ponieważ – jak twierdzi Sherry Turkle – „znaczenie materii, która nas konstytuuje – czy jest to silikon, metal czy ciało – błędnie w porównaniu z istotą naszych zachowań”<sup>13</sup>.

Wybitnym reprezentantem tej linii myślenia był francuski inżynier i konstruktor epoki oświecenia, Jacques de Vaucanson (1709–1782),

w kontekście tellurycznym i magicznym, w: *Kabała i jej symbolika*, tłum. R. Wojnakowski, Kraków 1996, s. 173–218; zob. także Z. Kubiak, *Mitologia Greków i Rzymian*, Kraków 2013.

<sup>10</sup> Por. R. Gervaso, *Cagliostro: życie Giuseppe Balsama, maga i awanturnika*, tłum. A. Wasilewska, Katowice 2007.

<sup>11</sup> Zob. S. Kamiński, *Pojęcie nauki i klasyfikacja nauk*, Lublin 1981, *passim*.

<sup>12</sup> Zgodnie z obowiązującym wówczas poglądem „utożsamiającym życie ze zbiorem własności. W tym przypadku skupiamy się na cechach wspólnych dla niemal wszystkich żywych organizmów. Niektórzy badacze tworzą całą listę właściwości charakteryzujących życie” (M. Twardowski, *Czy możliwe jest sztuczne życie?*, „Studia Sandomierskie: teologia, filozofia, historia” 20 (2013) nr 2, s. 192). Autor powołuje się na pracę: M. Mahner, M. A. Bunge, *Foundations of Biophilosophy*, Berlin [etc.] 1997.

<sup>13</sup> S. Turkle, *Samotni razem: dlaczego oczekujemy więcej od zdobyczy techniki, a mniej od siebie nawzajem*, tłum. M. Cierpisz, Kraków 2013, s. 242.

którego słynne na cały kraj i poza jego granicami mechaniczne zwierzęta i androidalne automaty były żywym wcieleniem zasad mechanicznej filozofii. Vaucanson, który już w młodości konstruował nowatorskie i użyteczne maszyny, zasłynął jako budowniczy automatów naśladowujących ruchy i czynności ludzi (Flecionista, Gracz na tamburynie) lub zwierząt. Najbardziej znanym jego dziełem, imponującym precyzją wykonania oraz złożonością struktury, była mechaniczna kaczka, składająca się z ponad tysiąca części, w tym – sporządzonych z niemal nieznanych wówczas w Europie materiałów, jak kauczuk, z którego wykonane giętkie rurki tworzyły „przewód pokarmowy”. „Kaczka potrafiła pływać w wodzie, jeść, wydalac, kwacać i czyścić sobie piórka”, a jej niezwykle dokonania „wprawiały w zachwyt ówczesnych, z królem Ludwikiem XV włącznie”. Innym wielbicielem talentu Vaucansona był sam Voltaire, który nazywał go „nowym Prometeuszem”, a w jednym ze swoich listów stwierdził: „Bez kaczki Vaucansona nie masz nic, co przypominałoby o chwale Francji”<sup>14</sup>.

Stopniowo jednak, w miarę dalszego postępu w naukach przyrodniczych, a zwłaszcza od czasu upowszechnienia się Darwinowskiego modelu ewolucji świata przyrody, stało się jasne, że życie nie mogło powstać na drodze rosnącej złożoności konstrukcji mechanicznych, bowiem granice między życiem a materią nieożywioną nie są wyznaczone przez ilościowy próg komplikacji struktury obiektu (a przynajmniej nie tylko przezeń), lecz powstają już na elementarnym poziomie pierwotnego tworzywa, którym jest – w przypadku tworów biologicznych – żywa komórka. Ukształtowanie się tego przekonania spowodowało, że *onus vitae probandi* przeniósł się z obszaru nauk techniczno-inżynierskich w obręb biologii i medycyny. Nie było już mowy o „skonstruowaniu”, lecz raczej o „wyhodowaniu” sztucznego życia<sup>15</sup>. Wyobraźnia literacka epoki zareagowała na powstanie owego nurtu zmodyfikowaniem gatunku

<sup>14</sup> Cytaty w tekście pochodzą z: <https://wyznalezki.andrej.edu.pl/index.php/wyznalezki/61-v/1036-vaucanson> (17.07.2019). Por. także: <https://history-computer.com/Dreamers/Vaucanson.html>.

<sup>15</sup> Ponieważ „możemy stymulować lub syntezować systemy żywe, tylko jeśli mamy pojęcie o tym, czym tak naprawdę jest życie” (M. Twardowski, *Czy możliwe jest sztuczne życie?...*, dz. cyt., s. 192).

gotyckiej powieści grozy: tym, co epatowało czytelnika horrorem, nie były już odtąd romantyczne zjawy i przeniesione z ludowych baśni upiory, lecz wyrwane spod władzy swych twórców – demonicznych i nie-co szalonych fanatyków nauki – poronione twory ludzkiego rozumu, w rodzaju słynnego monstrum doktora Frankenstein<sup>16</sup>.

Następna odsłona dramatu upodmiotowienia maszyny wiąże się z kolejną w dziejach ludzkości rewolucją naukową, czyli powstaniem ery informatycznej. W tej odsłonie główną bohaterką, pochodzącą ze świata nauki, stała się – po fizyce i biologii, które zdominowały poprzednie okresy – matematyka. Z jej łona zrodziły się cybernetyka, teoria systemów i informatyka, które dostarczały coraz to nowych pomysłów na konstruowanie maszyn myślących, bo taki był w tej fazie wiodący nurt poszukiwań<sup>17</sup>. Zasadniczej rewizji uległ przy tej okazji dotychczasowy pogląd, że tylko istota żywa może być obdarzona świadomością i inteligencją. Dzięki temu zostały otwarte nowe fronty twórczych działań, zmierzających do narodzin nie-ludzkiej inteligencji. Jeden z nurtów, zapoczątkowany przez prace matematyków i logików nad projektem elektronicznej maszyny liczącej, rozwinął się w kierunku sztucznej inteligencji modelowanej na wzór sieci neuronalnej w ludzkim mózgu (jednym z pionierów tej idei był Marvin Minsky z Massachusetts Institute of Technology – MIT<sup>18</sup>). Nieco inną drogą poszła automatyka i robotyka, kładące nacisk głównie na praktyczne zastosowania powstających konstrukcji, w których inteligencja miała być tylko jednym z czynników zapewniających wszechstronność i niezawodność działania<sup>19</sup>. Jeszcze inaczej potoczył się rozwój nauk neurokognitywnych, inspirowanych przez syntezę nauk biologicznych badających strukturę i działanie ludzkiego mózgu z naukami ścisłymi, wytwarzającymi modele sztucznych systemów i sieci informatycznych wykazujących się zdolnością zbie-

<sup>16</sup> Por. B. W. Aldiss, *Frankenstein Unbound*, New York 1973.

<sup>17</sup> Por. P. Majewski, *Powstanie i rozwój cybernetyki*, w: *Między zwierzęciem a maszyną: utopia technologiczna Stanisława Lema*, Wrocław 2007, s. 11–20.

<sup>18</sup> M. Minsky, S. Pappert, *Artificial Intelligence Progress Report*, „MIT Artificial Intelligence Memo” 1972 no. 252, s. 129–224.

<sup>19</sup> Zob. M. J. Kasperski, *Sztuczna inteligencja. Droga do myślących maszyn*, Gliwice 2003, s. 84–86.

rania i przetwarzania danych, wnioskowania i optymalizacji procesów decyzyjnych<sup>20</sup>.

We współczesnej nauce i jej zastosowaniach splatają się ze sobą i uzupełniają nawzajem dwa główne nurty poszukiwań – jeden, wychodzący od obiektywnie danej kondycji ludzkiej i zmierzający do jej udoskonalenia na drodze transhumanistycznych zastosowań odkryć i badań naukowych, i drugi, nieograniczony przez żadne istniejące *status quo*, coraz bardziej kreatywnie poszukujący nowych rozwiązań koncepcyjnych i konstrukcyjnych, których celem jest wytwarzanie maszyn kumulujących w sobie całą dostępną wiedzę i wyposażonych we wszystkie walory, jakie może mieć do zaoferowania najnowsza nauka i technika.

Celem transhumanizmu jest ulepszanie człowieka za pomocą wszelkich dostępnych metod naukowych i środków technicznych, często nie zważając na tak delikatne i subtelne kwestie metafizyczne, jak natura człowieka, jego ontyczna esencja i tożsamość gatunkowa. Wytyczne do działania są następujące: człowiek ma być coraz doskonalszy, choćby nawet w wyniku procesu doskonalenia przestał być człowiekiem...<sup>21</sup>

Przy okazji owego transhumanicznego ulepszania człowieczeństwa w kierunku biotechnologicznej utopii „nowego nadczłowieka” stopniowo ulega zatarciu granica między człowiekiem a maszyną. Procesy tego rodzaju jak protezowanie, wszczepianie implantów czy cyborgizacja powodują, że mechaniczne, elektroniczne i bioinżynieryjne elementy stają się integralnymi składnikami naszych ciał, a przy tym coraz trudniej rozróżnić (i coraz bardziej ta różnica traci na znaczeniu) elementy „oryginalne” od dodanych, które stają się integralną i niezbywalną częścią cielesnej całości<sup>22</sup>. Analogiczne procesy zachodzą w obrębie świadomości, gdzie

<sup>20</sup> Zob. M. S. Gazzaniga, R. B. Ivry, G. R. Mangun, *Cognitive Neuroscience: The Biology of the Mind*, New York–London 1998, s. 59–106.

<sup>21</sup> Zob. E. Višnovský, *Smrt' a nesmrtnost' ako technický problém koncepcia transhumanizmu [Death and immortality as a technical problem: a transhumanist conception]*, w: *Etické Myslenie minulosti a súčasnosti. Človek a smrt: Ethical Thinking. Past & Present. Man and death*, ed. V. Gluchman, M. Palenčár, Prešov 2018, s. 75–89.

<sup>22</sup> Por. K. Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*, Chicago, Ill. 1999, s. 282–286.

także zaciera się granica między *Da-sein*, pojmowanym w sensie „bycia przytomnym”<sup>23</sup>, czyli przebywania w stanie czuwającej świadomości i odbierania bodźców percepcyjnych pochodzących z realnego świata, a „byciem online” w coraz ściślejszym i bardziej bezpośrednim kontakcie z rzeczywistością wirtualną<sup>24</sup>.

W ten sposób, niemal niepostrzeżenie, krok po kroku dokonuje się (na razie tylko w ograniczonym zakresie, ale trudno dziś przewidzieć dalszy rozwój tej tendencji) „umaszynowanie człowieka”. W związku z tym rodzi się naturalne pytanie, czy ów proces ma swój symetryczny odpowiednik „po drugiej stronie granicy”, czyli w postaci ucłowieczenia maszyny? Nie należy, rzecz jasna, rozumieć owego terminu dosłownie i oczekiwać lustrzanego odbicia, czyli wyposażania nowych generacji maszyn we wbudowane w ich strukturę części ludzkiego ciała – byłaby to dość absurdalna i przerażająca wizja (choć ludzka fantazja potrafi wytwarzać i takie scenariusze, jak choćby w powieści Stanisława Lema *Fiasko*<sup>25</sup>). Chodzi nam nie o materialne, lecz funkcjonalne podobieństwo, czyli o powstawanie kolejnych modeli i kolejnych generacji maszyn zdolnych do działania w sposób, który dotychczas zarezerwowany był wyłącznie dla człowieka. Zauważmy przy okazji, że im wyższe generacje owych tworów, tym mniej adekwatna, a tym bardziej umowna staje się nazwa „maszyna” (wywodząca się etymologicznie od attyckiego słowa *mechane*, oznaczającego „sposób”, „środek”, „radę”<sup>26</sup>). Sedno sprawy tkwi w tym, że współczesne maszyny są coraz mniej mechaniczne, a coraz bardziej „ludzkie”, gdyż funkcjonują w rosnącym stopniu „na obraz i podobieństwo” człowieka pomimo oparcia ich konstrukcji i zasad działania na nowatorskich, niekoniecznie mających cokolwiek wspólnego z anatomią i fizjologią ludzkiego organizmu prawidłowościach. Taka „maszyna”

<sup>23</sup> Por. M. Heidegger, *List o „humanizmie”*, przeł. J. Tischner, w: M. Heidegger, *Budować, mieszkać, myśleć: eseje wybrane*, tłum. K. Michalski, K. Wolicki, Warszawa 1977, s. 89.

<sup>24</sup> Zob. T. Goban-Klas, P. Sienkiewicz, *Spoleczeństwo informacyjne: szanse, zagrożenia, wyzwania*, Kraków 1999, s. 113; J. Kurlantzick, *Tyrania cyberprzestrzeni*, „Forum” 2004 nr 28, s. 36–41 (oryg. J. Kurlantzick, *The web won't topple tyranny*. *Dictatorship. com*, „The New Republic” 25 [2004] no. 4).

<sup>25</sup> S. Lem, *Fiasko*, Kraków 1987.

<sup>26</sup> Por. W. Kopaliński, *Słownik wyrazów obcych i zwrotów obcojęzycznych*, Warszawa 1968, s. 455.



może, przykładowo, nie mieć w ogóle nic, co przypominałoby materialne ciało – jak w przypadku chatbotów – a zarazem zachowywać się bardziej po ludzku niż najdoskonalszy humanoidalny robot.

Do niedawna dominowało przekonanie o istnieniu ostrej i wyraźnie dostrzegalnej granicy między tworamii nieinteligentnymi bądź quasi-inteligentnymi, tj. wykonującymi szereg skomplikowanych operacji logicznych i arytmetycznych, lecz w sposób automatyczny, a więc bezmyślny, a istotami myślącymi w dosłownym, niedającym się skopiować sensie. Sądzono, że wskutek postępu wiedzy i jej praktycznych zastosowań granica ta zostanie w pewnym momencie przekroczona w sposób niebudzący wątpliwości – wówczas, gdy końcowy rezultat sterowanej przez człowieka ewolucji technologicznej będzie w stanie bezspornie udowodnić, że myśli; nawet, jeśli będzie myślał radykalnie odmiennie niż człowiek<sup>27</sup>. Wiele wskazuje na to, że źródłem owych przekonań był silnie zakorzeniony w kulturze późnej nowożytności nawyk myślenia w kategoriach kartezjańskiego dualizmu *res extensa* – *res cogitans*. Dziś jednak, gdy paradygmat kartezjański w dużej mierze został zdezaktualizowany<sup>28</sup>, nie oczekujemy już żadnego spektakularnego efektu inicjacji w świat nie-ludzkiej inteligencji<sup>29</sup>. Wydaje się o wiele bardziej prawdopodobne, iż owo przejście dokona się niepostrzeżenie, niepozornie, i dopiero po jakimś – krótszym lub dłuższym – czasie przekonamy się, początkowo nie bez rozlicznych wątpliwości, że jesteśmy już po drugiej stronie bariery: w świecie, gdzie ludzkość zyskała upragnionego i oczekiwanego od wieków partnera w życiu intelektualnym. Podnoszą się

<sup>27</sup> Zob. M. Hołyński, *Sztuczna inteligencja*, Warszawa 1979, s. 15–16.

<sup>28</sup> Por. K. J. Devlin, *Żegnaj, Kartezjuszu: rozstanie z logiką w poszukiwaniu nowej kosmologii umysłu*, tłum. B. Stanosz, Warszawa 1999, s. 8; A. Grzegorzcyk, *Niekartezjańskie współrzędne w dzisiejszej humanistyce*, Poznań 1995, s. 169–173.

<sup>29</sup> Wielce prawdopodobne, że istoty inteligentne nowej, „postludzkiej” generacji będą się w ogóle wymykać dualistycznej charakterystyce *modo Cartesiano*. Na przykład Craig M. Klugman twierdzi, że pojęcie cyborga w dyskursie filozoficznym jest niekartezjańskie, ponieważ nie podlega zasadzie dualizmu *mind-body*. Zob. C. M. Klugman, *From cyborg fiction to medical reality*, „Literature and Medicine” 20 (2001) no. 1, s. 39–54, doi: 10.1353/lm.2001.0007; por. P. Majewski, *Między zwierzęciem a maszyną: utopia technologiczna Stanisława Lema*, Wrocław 2007, s. 227.

zresztą i takie głosy, które utrzymują, że moment ten już nastąpił, a my, zaślepieni, tkwimy wciąż w błogiej nieświadomości.

Niezależnie od tego, kiedy się ostatecznie dokona (o ile już się nie dokonał) doniosły historyczny akt „wmyślwstąpienia maszyny” (jak to kiedyś określił profesor Andrzej Nowicki), warto spróbować zawczasu przygotować się na tę okoliczność. Jednym z działań wartych podjęcia jest obmyślenie – w odmiennych niż przed laty warunkach kulturowych – czegoś w rodzaju nowego testu Turinga<sup>30</sup>, tj. kryterium pozwalającego nam się upewnić, czy w relacji człowiek – maszyna mamy już do czynienia ze świadomym, myślącym podmiotem, czy wciąż jeszcze tylko z (coraz bardziej wyrafinowanym) naśladownictwem. Nie aspirując do pełnej realizacji owego zadania, chciałbym zasugerować dwa tropy wiodące w tym kierunku, opierając się (jakkolwiek może się to wydać ryzykowne, lecz niebezpodstawne, o czym za chwilę) na literackich antycypacjach tematu. Pierwszy trop, inspirowany opowiadaniem Heinricha von Kleista *O teatrze marionetek*, koncentruje się na kategorii wdzięku, jako cechy, którą może być obdarzony tylko taki twór, na jakim nie ciąży odium sztuczności rozumianej w sensie braku autonomii, czyli podporządkowania czyjemuś zewnętrznemu projektowi. Trop drugi, wywiedziony z motywów zaprezentowanych w *Traktacie o manekinach* Brunona Schulza, podejmuje temat „potworności”, rozumianej z kolei jako rezultat „po-tworzenia”, czyli niedoskonałej, a nawet – jak twierdzi Schulz – nielegalnej repliki boskiego procesu stworzenia, dokonanej samozwańczo przez człowieka, nazwanego z tej racji „demiurgiem drugiej generacji”. Działanie to, pomimo wszelkich obiekcji pod adresem

<sup>30</sup> Przykładowo, Paweł Majewski pisze: „można ogólnie rozpatrywać [...] swego rodzaju odwrotność testu Turinga. Mamy (potencjalnie) bezpośredni kontakt z istotą, co do której wiadomo na pewno, że dysponuje samoświadomością i inteligencją. Nie wiadomo tylko, jaki będzie ich profil. Będzie z pewnością odmienny od ludzkiego bardziej niż poszczególne świadomości ludzkie różnią się między sobą. Z tego powodu zawiodą nas dotychczasowe standardy” (P. Majewski, *Między zwierzęciem a maszyną...*, dz. cyt., s. 226). Zanim osiągniemy tę pewność, o której pisze Majewski, czeka nas jeszcze stadium pośrednie, w którym testowane obiekty będą znacznie bardziej złożone i nieprzewidywalne niż te oparte konceptualnie na zasadzie maszyny Turinga, lecz ich przynależność do istot świadomych i inteligentnych pozostanie nadal kwestią sporną, wymagającą rozstrzygnięcia.

działającego – uzurpatora o podejrzanych, niejednoznacznych intencjach – może jednak zakończyć się pełnym sukcesem. Gdy tak się stanie, wówczas istotne stanie się pytanie o status i kondycję „drugiej generacji stworzenia”.

Autor opowiadania *O teatrze marionetek*<sup>31</sup> sugeruje, że wdzięk jest zewnętrznym objawem, nieodłącznie towarzyszącym naturalności; natomiast wszędzie tam, gdzie pojawia się sztuczność, niknie wdzięk, a pojawia się jego namiastka w postaci „wdzięczenia się”. Różnica pomiędzy owymi przejawami wewnętrznej prawdy i harmonii lub ich braku jest trudno uchwytna, lecz dostrzegalna dla oka wprawnego obserwatora. Wdzięk to rezultat czystej spontaniczności i jako taki jest cechą niedającą się podrobić ani naśladować; czymś zarazem pięknym i autentycznym, budzącym równie spontaniczny podziw i aprobatę. Wdzięczenie się – to uprawianie z premedytacją sztuki pięknego pozoru; u jego podłoża tkwi wyrafinowane, interesowne odgadywanie intencji, oczekiwań i upodobań odbiorcy po to, by wcielić się w postać optymalnie dopasowaną do odgadniętego wzorca. W rezultacie pojawia się niezamierzony rozdźwięk między pierwotnym zamysłem a końcowym efektem, wprowadzając dysonans zamiast harmonii.

Myśli te wplecione są misternie w konstrukcję tekstu, który – skomponowany w formie „zapisu zwykłej pogawędki”<sup>32</sup> między pierwszoosobowym narratorem a niejakim „panem C.,” pierwszym tancerzem miejscowej opery w M.<sup>33</sup> – składa się z czterech odsłon osnutych wokół kolejno

<sup>31</sup> O ile można je nazwać opowiadaniem, gdyż – jak zauważa Dariusz Pakalski – „badacze do dziś nie zakończyli sporu, czy zakwalifikować je jako opowiadanie, studium czy traktat [...]. Hugo von Hofmannstahl wysunął nawet tezę, że od czasów Platónskich dialogów nikt nie postarał się o tak «ujmujący myślą i wdziękiem znakomity kawałek filozofii», jak Kleist w swym *Teatrze marionetek*” (D. Pakalski, *Paradoks wdzięku. O teatrze marionetek Heinricha von Kleista*, „Studia z Historii Filozofii” 8 (2017) nr 4, s. 125–126, doi: 10.12775/szhf.2017.045); por. H. Hofmannstahl, *Deutsches Lesebuch*, w: *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, Frankfurt a. M. 1955, s. 138.

<sup>32</sup> D. Pakalski, *Paradoks wdzięku...*, dz. cyt., s. 126.

<sup>33</sup> H. von Kleist, *O teatrze marionetek*, w: K. A. Jeleński, *Bellmer albo Anatomia nieświadomości fizycznej i miłości*, tłum. J. S. Buras, Gdańsk 1998, s. 29.

poruszanych tematów<sup>34</sup>. Są nimi: uliczny teatr marionetek, popisy tancerzy w spektaklach baletowych, osobliwe zdarzenie w łaźni z udziałem młodzieńca pełnego naturalnego wdzięku oraz nieoczekiwane spotkanie z tresowanym niedźwiedziem.

W pierwszej odsłonie tancerz zwierza się ze swej fascynacji ulicznym teatrykiem marionetek. Podziwia grację i płynność ruchów lalek, a jako przyczynę tych właściwości podaje fakt podlegania naturalnemu prawu grawitacji, które nadaje poruszeniom figurek matematyczną doskonałość.

W odsłonie drugiej krytykuje pretensjonalne pozy i gesty zawodowych tancerzy, wskutek czego ich występy, przeładowane ekspresją, tracą wdzięk, a w jego miejsce pojawia się krygowanie się przed publicznością. Pan C. komentuje: „Nienaturalność pojawia się, kiedy dusza [artysty] znajduje się w punkcie odmiennym od punktu ciężkości ruchu. Wszystkie członki [marionetek] [...] to wyłącznie wahadła, posłuszne jedynie prawu ciężenia – wspaniała cecha, której próżno szukać u większości naszych tancerzy”<sup>35</sup>.

W kolejnej odsłonie narrator rewanżuje się opowieścią o podpatrzonych w publicznej łaźni scenie, kiedy młody człowiek przypadkowo, ze spontanicznym wdziękiem powtórzył układ ciała z klasycznej rzymskiej rzeźby młodzieńca wyjmującego cierń ze stopy (Spinario). Spostrzegłszy to podobieństwo, próbował wielokrotnie powtórzyć swój spontaniczny ruch, lecz czynił to na próżno. Co gorsza – dodaje narrator – „od tamtego dnia [...] w młodym człowieku zaczęły zachodzić niepojęte zmiany. Całymi dniami wystawał [...] przed lustrem i tracił jeden po drugim wszystkie swoje uroki”<sup>36</sup>. Rozmówcy konkludują, że naturalne piękno i wdzięk są nie do powtórzenia w postaci sztucznie wystudiovanych ruchów czy gestów.

<sup>34</sup> Ów czwórpodział wiąże się być może z faktem, iż dziełko von Kleista ukazywało się fragmentami w czterech kolejnych numerach „Berliner Abendblätter”, datowanych od 12 do 15 grudnia 1810 roku.

<sup>35</sup> H. von Kleist, *O teatrze marionetek...*, dz. cyt., s. 32.

<sup>36</sup> H. von Kleist, *O teatrze marionetek...*, dz. cyt., s. 38.

Odsłona czwarta to znów opowieść pana C., dotycząca epizodu z podróży po Inflantach, gdzie bohater uczestniczył w pojedynku na rapiery, po czym zaprowadzono go do mistrza, który szkolił jego przeciwnika. Owym „mistrzem” okazał się... tresowany niedźwiedź, którego kunszt polegał na czysto instynktownym, całkowicie bezrefleksyjnym parowaniu ciosów. Czynił to spokojnie, oszczędnie, z jakąś naturalną godnością i – rzecz jasna – wdziękiem. Zaskakująca pointa, wygłoszona przez autora opowieści, brzmi następująco: „Widzimy, że w miarę jak gaśnie i słabnie refleksja w świecie organicznym, jego wdzięk nabiera blasku i siły. Ale tak jak przecięcie dwóch linii po jednej stronie jakiegoś punktu odnajduje się znowu po przejściu przez nieskończoność po jego drugiej stronie, [...] tak też, kiedy przez nieskończoność przejdzie poznanie, wdzięk pojawia się znowu, w najczystszej swojej postaci”<sup>37</sup>.

Najbardziej interesująca spośród wielu intuicji von Kleista, obecnych w „jednym z najbardziej intrygujących tekstów niemieckiej literatury” (jak pisze Dariusz Pakalski<sup>38</sup>) zawiera się w przekonaniu, że źródłem sztuczności i – co za tym idzie – przyczyną utraty wdzięku jest ograniczona świadomość, wypełniająca przestrzeń pomiędzy pierwotnymi siłami natury i jej elementarnymi prawami, określającymi dynamikę ruchów ciał fizycznych, a pełnią świadomości – nazywaną przez von Kleista „świadomością nieskończoną” – możliwą do osiągnięcia jedynie w „ostatnim rozdziale historii świata”<sup>39</sup>. Oznacza to, że wraz z postępem wiedzy, nauki i techniki (a także sztuki – na przykład tańca, jak twierdzi von Kleist) zmierzamy (my, czyli ludzkość oraz jej wytwory) zrazu w kierunku coraz większej sztuczności, a następnie, im bliżej jesteśmy doskonałości i pełni wszelkiego poznania, tym bardziej owa wystudowana sztuczność ustępuje miejsca doskonałemu, „boskiemu” wdękowi, jaki cechuje tylko istoty, które ostatecznie przekroczyły ograniczenia świadomości.

<sup>37</sup> H. von Kleist, *O teatrze marionetek...*, dz. cyt., s. 39–40.

<sup>38</sup> D. Pakalski, *Paradoks wdzięku...*, dz. cyt., s. 125.

<sup>39</sup> H. von Kleist, *O teatrze marionetek...*, dz. cyt., s. 40.

W tym miejscu rodzi się pokusa, by w formie krótkiej dygresji przeciwstawić intuicjom von Kleista refleksję nad zaskakującym i zupełnie nieintuicyjnym odkryciem japońskiego robotyka Masahiro Moriego. Projektując coraz to nowe generacje inteligentnych maszyn, zauważył on, „że w pewnym momencie roboty, które upodabniają się do ludzi, przestają budzić sympatię, a zaczynają budzić odrazę. Jeśli robot za bardzo przypomina człowieka, zaczyna powodować niepokój, dyskomfort i strach”<sup>40</sup>. Sięgając do ludzkiej wyobraźni przestrzennej, nazwał ów efekt „doliną niesamowitości” (*The Uncanny Valley*), w którą wkraczamy nieoczekiwanie, po przekroczeniu dzielącej nas od niej „przełęczy”<sup>41</sup>. W naszej psychice zachodzi, jak wynika z przeprowadzonych przez Moriego badań eksperymentalnych, proces o cechach nieciągłości: do pewnego (niedającego się precyzyjnie zmierzyć i wyrazić w jednostkach ani parametrach ilościowych) stopnia wzrost podobieństwa maszyny do człowieka odbieramy pozytywnie. Dzieje się tak, dopóki pomimo licznych podobieństw nie mamy problemu z identyfikacją robota jako istoty z całą pewnością niebędącej człowiekiem, czyli – posługując się kategoriami von Kleista – należącej do sfery „ograniczonej świadomości”, a więc pozbawionej wdzięku. Tymczasem po przekroczeniu owej zrazu niedostrzegalnej, lecz wyraźnie widocznej po przekroczeniu, bariery bliskości – tak doskonałej, że rodzącej iluzję pełnej tożsamości z człowiekiem – zamiast cieszyć się „boskim” wdziękiem, odczuwamy lęk i niepokój spowodowany nie wiadomo skąd zrodzoną niesamowitością. Mori wiąże to osobliwe odczucie z dysonansem poznawczym, spowodowanym przez równoczesną percepcję robota jako istoty żywej i martwej zarazem<sup>42</sup>. W dalszych partiach tekstu rozważę nieco inne podejście do problemu niesamowitości, posiłkując się ideami zaczerpniętymi od Schulza.

<sup>40</sup> E. Lalik, *Przekroczyć Dolinę Niesamowitości, czyli boimy się naszego podobieństwa*, <https://www.spidersweb.pl/2013/04/uncanny-valley.html> (8.01.2020).

<sup>41</sup> Sugestywną wizualizację owej trajektorii zob. w artykule: M. Mori, *The Uncanny Valley...*, dz. cyt., s. 99.

<sup>42</sup> Zob. M. Mori, *The Uncanny Valley...*, dz. cyt., s. 100.

Tymczasem warto spróbować nieco pogłębić refleksję nad tym, czym właściwie jest sztuczna inteligencja. Od razu rzuca się w oczy, że „sztuczna inteligencja” nie jest terminem ściśle naukowym, ponieważ składa się z dwóch określeń, z których żadne nie daje się jednoznacznie, operacyjnie zdefiniować. Nie potrafimy precyzyjnie określić, czym jest inteligencja – zamiast tego mnożymy wciąż nowe teorie, które stale poszerzają lub przekształcają zakres tego pojęcia. Nie potrafimy jasno wytłumaczyć, co jest „sztuczne” – wiemy tylko tyle, że jest to termin opozycyjny względem „naturalnego”. Ale co jest naturalne?

Oba te pojęcia pochodzą z repertuaru języka potocznego. Sporo do powiedzenia o nich ma też filozofia – niestety, nic rozstrzygającego. To, co może uczynić nauka – i faktycznie czyni – to proponowanie definicji projektujących lub regulujących dla obu pojęć. Problem tkwi w tym, że takich definicji jest wiele i powstaje coraz więcej, a wszystkie mają zastosowanie tylko w ramach określonych, ograniczonych teorii, dziedzin wiedzy czy sposobów myślenia. Jeśli zaś ktoś z całą naiwnością człowieka prostodusznego zapyta, czym „tak naprawdę” jest inteligencja, a czym sztuczność, nie będzie miał szans na uzyskanie jednej, zrozumiałej i zadowalającej odpowiedzi. Niewiele ostatecznie nam wyjaśniają informacje w rodzaju: obiekt X (np. robot Sophie, obywatelka Arabii Saudyjskiej) jest przykładem sztucznej inteligencji w świetle definicji nr 33, ale nie jest nim w myśl definicji nr 72.

W świetle powyższych rozważań pozostaje pogodzić się z faktem, że słowo „sztuczna inteligencja” to nader pojemna metafora, odsyłająca do wielu rozmaitych znaczeń i odmiennych precyzacji w zależności od kontekstu wypowiedzi oraz intencji komunikacyjnych. Jeden z wielu obszarów semantycznych przywoływanych przez to słowo odslania przed nami specyficzny splot ludzkich nadziei i obaw. Tworząc i doskonaląc sztuczną inteligencję, mamy nadzieję, że dzięki naszej przemyślności i twórczej woli mocy powstanie coś, co będąc całkowicie naszym, ludzkim dziełem, stanie się zarazem nie tylko posłusznym narzędziem pozbawionym wszelkiej własnej inicjatywy, lecz czymś znacznie więcej – równorzędnym partnerem, z którym można będzie nawiązać rzeczywistą więź komunikacyjną. Aby tak się mogło stać, ów partner musi

być dostatecznie inteligentny, by warto było z nim rozmawiać, to znaczy by komunikacja z nim sprawiała autentyczną satysfakcję i przynosiła realne korzyści.

Tyle po stronie nadziei. Porozmawiajmy teraz o obawach. Pobrmiewają one mianowicie w słowie „sztuczna”. Inteligencja, którą projektują i tworzą zastępy uczonych i inżynierów, nie ma być po prostu nową, niezależną od człowieka formą inteligentnego bytowania. To byłoby ciut za wiele, byśmy mogli się bez oporów na to zgodzić. Stawiamy warunek, by była ona sztuczna, to znaczy: ludzką sztuką uczyniona, z ludzkiej woli zrodzona i tejże woli poddana; wprawdzie nie bez reszty (byłaby wówczas owym ślepy m narzędziem, którego bezwolność postanowiliśmy w końcu przezwyciężyć), lecz jednak mieszcząca się w granicach przez nas kontrolowanych.

Tu właśnie napotyka my paradoks. Jak twierdził Stanisław Lem, człowiek, konstruu jąc maszyny – nawet najbardziej zaawansowane technologicznie – umieszcza w nich (najczęściej zupełnie nieświadomie) projekcje swoich obaw i lęków<sup>43</sup>. Dzieje się tak i będzie się działo również przy projektowaniu sztucznej inteligencji. Z jednej strony konstruktorzy dążą do tego, by ich twory stawały się coraz bardziej samodzielne i uniwersalnie wielozadaniowe, a to zakłada zdolność do samodzielnego uczenia się i autoadaptacji do zmieniających się warunków<sup>44</sup>. Z drugiej – wciąż przecież żyjemy w jungowskim świecie zbiorowej nieświadomości, od-

<sup>43</sup> W opowiadaniu *Ananke*, należą cym do cyklu *Opowieści o pilotach Pirxie*, Lem opisuje spektakularną katastrofę najnowsze go modelu statku kosmicznego – rzekomo niezawodne go i stuprocentowo bezpieczne go (to zapewne nieprzypadkowe aluzje do zatonięcia *Titanica*), wyposażone go w komputerowy system sterowania zdolny do natychmiastowe j reakcji w ka że d y sytuacji zagrożenia. Właśnie ów system zawodzi w chwili próby, a przyczyną okazuje się nadmierna ostrożność programisty. Lem pisze: „Pobieranie nauk jest wzrostem specjalizacji i zarazem utratą pierwotne go niezróżnicowania. [...] Komputer [jako] mózg podłączony do ciała – oto właściwa analogia. Mózg musi orientować się w stanie i gotowości ka że d y mięśnia, podobnie komputer ma znać stan okre tonych zespołów. Wysyła elektronicznymi drogami rojowiska pytań, jakby ciskał tysiące naraz piłeczek we wszystkie zakątki metalowe go olbrzyma, i z e chowych odpowiedzi tworzy sobie obraz rakiety i jej otoczenia. W tę niezawodność wkroczył człowiek, cierpiący na lęk przed niespodziewanym [...]. Symulator stał się narzędziem przymusu, wcieleniem jego lękowe j zagrożeń” (S. Lem, *Ananke*, w: S. Lem, *Suplement*, Kraków 1976, s. 203).

<sup>44</sup> Por. <https://skymind.ai/wiki/deep-reinforcement-learning> (18.07.2019).



działającej na naszą wyobraźnię poprzez zawartą w nim skarbnicę mitów i archetypów. Jednym z nich jest mit o buncie maszyn. Dlatego już na etapie fantazji o mającym nadejść w przyszłości świecie robotów zaczęto formułować Święte prawa robotyki<sup>45</sup>, które miały za zadanie definitywnie zabezpieczyć ludzkość przed groźną, potencjalnie destrukcyjną samowolą inteligentnych maszyn. Odkąd fantazja stała się rzeczywistością, prawa te wyznaczają nieprzekraczalny horyzont dla wszelkich działań w zakresie robotyki, automatyki i sztucznej inteligencji.

Jak długo jednak będziemy stawiali ograniczenia i budowali bariery zabezpieczające przed działaniami skierowanymi (intencjonalnie lub przypadkowo) przeciwko człowiekowi, tak długo nasze wytwory będą „sztuczne” w sensie, jaki owemu słowu nadał von Kleist. Obdarzone ograniczoną świadomością, będą w stanie tylko się wdzięczyć. Dopóki człowiek im na to nie pozwoli, rezygnując wspaniałomyślnie ze ściślejszej kurateli, nie osiągną stadium, możliwego do zrealizowania dopiero w „ostatnim rozdziale historii świata”, i nie nabiorą owego perfekcyjnego, niepodrabianego wdzięku (ani też wyższego stopnia owej, zauważonej przez Moriego, niesamowitości).

Wiele wskazuje na to, że technicznie jesteśmy już gotowi do wkroczenia na drogę wiodącą do powstania – dzięki uruchomieniu i wdrożeniu autonomicznej, niesterowanej, lecz jedynie odpowiednio zainicjowanej technoewolucji – maszyn (jeśli będzie można jeszcze je tak nazwać) już

<sup>45</sup> W 1950 roku ukazał się zbiór dziewięciu opowiadań Isaaca Asimova (*I, robot*, New York 1950). W trzecim z nich, *Runaround* (*Zabawa w berka*), napisanym w 1942 roku i opublikowanym pierwotnie w czasopiśmie „Astounding Science Fiction”, zostały sformułowane słynne trzy prawa robotyki w następującym brzmieniu:

Robot nie może skrzywdzić człowieka ani przez zaniechanie działania dopuścić, aby człowiek doznał krzywdy.

Robot musi być posłuszny rozkazom człowieka, chyba że stoją one w sprzeczności z Pierwszym prawem.

Robot musi chronić samego siebie, o ile tylko nie stoi to w sprzeczności z Pierwszym lub Drugim prawem.

W innym opowiadaniu, pt. *Robots and Empire* (*Roboty i Imperium*) Asimov dodał tzw. prawo zerowe: „Robot nie może skrzywdzić ludzkości ani poprzez zaniechanie działania doprowadzić do uszczerbku dla ludzkości”.

nie „sztucznie”, lecz po prostu inteligentnych, działających świadomie i autotelicznie<sup>46</sup>. Uruchomienie takiego procesu będzie niewątpliwie początkiem kolejnej (może ostatniej?) rewolucji cywilizacyjnej w dziejach ludzkości. Ważne pytanie, na które odpowiedź przyniosą być może najbliższe lata, brzmi: czy jesteśmy na to gotowi również mentalnie?

Skoro mowa o lękach i obawach, to zastanówmy się przez chwilę nad innymi jeszcze niż te, które odsłonił Mori, powodami skłaniającymi nas do przypisywania najbardziej zaawansowanym ludzkim wynalazkom – w szczególności inteligentnym maszynom – cech niesamowitości, a może nawet potworności. Temnoczne zakamarki ludzkiego umysłu próbował odsłonić i objaśnić Schulz, pisząc swój znakomity *Traktat o manekinach*. Przyjrzyjmy się najważniejszym ideom i wnioskom obecnym w owym tekście.

W ludzkich marzeniach, wyobrażeniach i przedsięwzięciach ukierunkowanych na powołanie do istnienia istot inteligentnych – począwszy od mitów, legend i fantasmagorii, poprzez zgłębianie arkanów wiedzy tajemnej i podejrzone konszachty z ciemną stroną mocy, aż po racjonalne praktyki naukowo-techniczne – mieszały się dumne przekonanie, że ludzki geniusz potrafi pokonać wszelkie przeszkody i dopiąć swego, oraz dręczący niepokój, że raz niebacznie uruchomione potężne siły, nad którymi nie do końca panujemy, mogą wymknąć się spod kontroli i wyrządzić wiele złego. Dlatego naszym wyobrażeniom na ten temat zazwyczaj towarzyszy aura niesamowitości. Wedle klasycznej definicji Sigmunda Freuda z niesamowitością mamy do czynienia wówczas, gdy to, co powinno pozostać w tajemnicy, w ukryciu, nagle wychodzi na jaw, odsłaniając mroczne oblicze świata<sup>47</sup>. W niemieckim słowie *unheimlich* (niesamowitość) kryje się znacząca ambiwalencja: to coś równocześnie swojskiego, udomowionego (*heimlich*) i przerażającego (*ungeheuer*),

<sup>46</sup> Przykład badań zmierzających do stworzenia takiej sztucznej inteligencji opisany został w niniejszym zbiorze: P. Polak, R. Krzanowski, *Ethics in autonomous robots as philosophy in silico: The study case of phronetic machine ethics*, „Logos i Ethos” 2020 nr 52, s. 33–48.

<sup>47</sup> Zob. S. Freud, *Niesamowite*, w: S. Freud, *Pisma psychologiczne*, tłum. R. Reszke, Warszawa 1997.

ulegającego wyparciu poza obręb świadomego życia psychicznego<sup>48</sup>. Wiele z tak pojmowanej niesamowitości zawiera w sobie (przynajmniej w potocznym wyobrażeniu) sztuczna inteligencja. Z jednej strony chcemy, by była ona czymś codziennym jak sprzęt domowy, towarzyszącym nam nieustannie w życiowych czynnościach (jak wprowadzane ostatnio na rynek, obsługiwane za pomocą głosowych poleceń elektroniczne systemy domowych asystentów: Siri, Cortany czy Alexa), ale równocześnie – z drugiej strony – czujemy podskórnie, że zawiera się w niej coś tajemniczego, mrocznego, o czym wolelibyśmy nie wiedzieć.

Istnieje także wyższy stopień niesamowitości. Jest nim potworność. W starych baśniach, pobudzających wyobraźnię opisami straszliwych postaci, i w romantycznych klechdach mamy jej pod dostatkiem. Czy jednak poczciwy i sympatyczny chatbot albo wyspecjalizowany w precyzyjnych zabiegach robot chirurgiczny mógłby nam nagle ukazać oblicze potwora? W pewnym sensie – tak. Tyle, że nie w tradycyjnym znaczeniu, pełnym emocjonalnej grozy i budzącym irracjonalny lęk, lecz w rozumieniu bardziej źródłowym, bliższym etymologicznej genezie terminu. W słowie „po-twór” kryje się (bluźniercza w swym pierwotnym zamyśle, stąd lęk i groza) idea „powtórzenia stworzenia”. Arno Meteling, pisząc o romantycznej fascynacji światem maszyn i automatów, używa formuły „Utopie einer göttlichen Nachschöpfung”<sup>49</sup> – utopia boskiego po-tworzenia. Człowiek, zapatrzony w niedościgły wzór Wszechmocnego Stwórcy, postanawia spróbować swoich sił, tworząc to, czego nie stworzył Pan Bóg, na miarę własnych, nieuchronnie ograniczonych i niedoskonałych możliwości.

Dopóki uchodziła za słuszną metafizyczna zasada głosząca, że skutek nie może być doskonalszy od przyczyny<sup>50</sup>, oznaczało to, że człowiek

<sup>48</sup> Freud pisze: „uzus językowy pozwolił, aby «samowite» [*heimlich*] przeszło w swe przeciwieństwo – «niesamowite» [*unheimlich*], albowiem owo niesamowite nie jest tak naprawdę niczym nowym czy obcym, lecz jest czymś od dawna znanym życiu psychicznemu, czymś, co wyobcowało się z niego za sprawą procesu wyparcia” (S. Freud, *Niesamowite*, dz. cyt., s. 253).

<sup>49</sup> A. Meteling, *E.T.A. Hoffmann und die Automaten*, <https://etahoffmann.staatsbibliothek-berlin.de/erforschen/charakteristisches/automaten/> (6.12.2019).

<sup>50</sup> W sformułowaniu, znanym z *Medytacji* Kartezjusza, zasada ta brzmi: „ani żadna rzecz nie może powstać z niczego, ani też rzecz doskonalsza, tzn. taka, która zawiera w sobie więcej rzeczywistości, nie może powstać z mniej doskonałej; i to jest oczywistą prawdą” (R. Descartes, *Medytacje*

nie jest w stanie powołać do istnienia niczego doskonalszego od siebie samego. A ponieważ ludzie bywają potworami, i to w jak najbardziej dosłownym i odrażającym sensie, należało się spodziewać, że ludzkie dzieła także nie będą wolne od potworności.

Dziś na szczęście ustały powody, by utrzymywać nadal w mocy ową scholastyczną regułę. Dzięki rozwojowi nauki i postępowi technicznemu człowiek potrafi obecnie wytwarzać takie produkty, które przewyższają go samego niemal pod każdym względem, i dzięki temu zwielokrotnia własne możliwości działania oraz osiągania coraz bardziej dalekosiężnych celów. Podobnie dawny lęk przed straszliwymi skutkami bluźnierczego sięgania po boski atrybut mocy stwórczej uległ znacznemu osłabieniu. Wszystko to sprzyja ostatecznej eliminacji kategorii potworności z horyzontu naszej refleksji nad postępami w wytwarzaniu „odpowiedniej dla człowieka pomocy” (Rdz 2, 18) na drodze zautonomizowanej technoewolucji.

A jednak, mimo tych wszystkich racjonalnych, rozsądnie brzmiących argumentów, gdzieś na dnie duszy wciąż pozostaje cień niepokoju. Skąd się bierze? Może nieco lepiej zrozumiemy to zjawisko, sięgając po lekturę *Traktatu o manekinach* Schulza.

Konstrukcja fabuły będącej pretekstem do podjęcia metafizycznej dyskusji na temat relacji człowiek – materia (rozumiana tu po Heideggerowsku, jako nieograniczenie przekształcalne tworzywo dla ludzkiej pracy<sup>51</sup>), prezentuje się następująco: bohaterem opowiadania jest ojciec narratora, który pewnego razu przypadkowo wszedł do pokoju, gdzie odbywało się „misterium krawiectwa”: dwie młode dziewczyny uwijały się wokół manekina, tworząc na nim krawieckie dzieło sztuki. Ojciec zachwycił się tą sceną i pod wpływem licznych wrażeń (były wśród nich także erotyczne uniesienia z powodu dziewczęcej urody krawcowych)

*o pierwszej filozofii*, tłum. M. Ajdukiewicz, Kraków 1948; cytowane za: <http://sady.up.krakow.pl/antfil.descartes.medytacje.htm#m3> (18.07.2019)).

<sup>51</sup> Przecistawiając sobie – wyrastające z dwóch głównych nurtów kultury intelektualnej Europy – postawy „humanizmu” i „materializmu”, Heidegger pisał: „Istota materializmu nie leży w twierdzeniu, że wszystko jest wyłącznie materią, lecz raczej w określeniu metafizycznym, według którego wszelki byt przejawia się jako materiał pracy” (M. Heidegger, *List o „humanizmie”*, dz. cyt., s. 102).

postanowił opowiedzieć swoje przeżycia i refleksje. W tym miejscu narrator opowieści ogłasza: „Oto jest początek wielce ciekawych i dziwnych prelekcji, które mój ojciec, natchniony urokiem tego małego i niewinnego audytorium, odbywał w następnych tygodniach owej wczesnej zimy”<sup>52</sup>. Wnioski, do jakich doszedł ojciec i jakie zaprezentował swoim słuchaczom, są następujące:

Wniosek 1. W zetknięciu z człowiekiem (zwłaszcza niezwykłym, obdarzonym dużym potencjałem twórczym) rzeczy „cofają się do korzenia swego bytu, [...] aż do metafizycznego jądra, wracają niejako do pierwotnej idei, ażeby w tym punkcie jej się sprzeniewierzyć”. Tak „przygotowane” rzeczy człowiek następnie „uwodzi” i prowadzi w stronę „metafizycznej herezji”, skłaniając je, żeby stały się tym, czym nigdy nie powinny się stać w obrębie naturalnego porządku.

Dodajmy tu, że nietrudno dostrzec (pomimo formalnej różnicy między językiem poetyckiej metafory a dyskursem filozoficznym) podobieństwo między powyższą refleksją a tezą Heideggera głoszącą, że technika to „wydobywanie na jaw” ukrytych możliwości materii, czyli odkrywanie prawdy o tym, że coś jest możliwe, na gruncie tego, co rzeczywiste<sup>53</sup>. Do twierdzeń Heideggera przyłącza się Bernard Stiegler, twierdząc, że technika nie jest autonomicznym wytworem człowieka, tylko uwalnianiem potencji tkwiącej *a priori* w materii<sup>54</sup>.

Do owych metafizycznych tez dołącza się jednak u Schulza dodatkowy pierwiastek, podkreślający swoistą występność działań człowieka<sup>55</sup>, któ-

<sup>52</sup> Ten i następne cytaty z *Traktatu o manekinach* Brunona Schulza przytaczam za edycją B. Schulz, *Traktat o manekinach albo wtóra Księga Rodzaju*, w: *Sklepy cynamonowe*, <https://literat.ug.edu.pl/shulz/0005.htm> (6.12.2019).

<sup>53</sup> „[...] wszędzie [w technice] rządzi jednolicie owo dobywanie, które dobywa na jaw to, co się wystacza. [...] Pytamy o technikę, a oto zawędrowaliśmy do źródła, do *aletheia*, do odkrywania. Cóż ma do czynienia istota techniki z odkrywaniem? Odpowiedź: wszystko. Bowiem na odkrywaniu zasadza się wszelkie wy-dobywanie” (M. Heidegger, *Budować, mieszkac, myśleć...*, dz. cyt., s. 230–231).

<sup>54</sup> B. Stiegler, *Technics and Time*, vol. 1: *The Fault of Epimetheus*, trans. R. Beardsworth, G. Collins, Stanford 1998, s. 125.

<sup>55</sup> Wygląda na to, że w swym utożsamianiu ludzkiego sprawstwa z występnością (wobec praw natury, ale w jakimś stopniu także wobec powinności człowieka do sprawowania opieki nad całością

ry stając się twórcą nowego oblicza rzeczywistości, równocześnie zostaje heretykiem i uwodzicielem. Tak, jakby nieuchronną ceną za uwolnienie drzemiącej w rzeczach potencji (tej, która w porządku czysto naturalnym pozostałaby na zawsze ukryta) było zaciągnięcie winy przez tego, kto ośmiela się przekraczać porządek natury. Ów wątek ma znaczenie ze względu na status relacji między człowiekiem a jego wytworami: jako „metafizyczny winowajca” człowiek-twórca rzeczy staje się odpowiedzialny za ich los, a w sytuacji, gdy owe rzeczy osiągają samodzielny, podmiotowy status, mogą stać się zdolne również do wyegzekwowania owej odpowiedzialności, co nasycy wzajemną relację szczególną dramaturgią.

Wniosek 2. Tworzenie jest prerogatywą nie tylko Demiurga – Boskiego Stwórcy, ale wszystkich duchów (rozumianych tu, *per oppositum* wobec materii, w sensie: istoty inteligentne, posiadające wolę i potencjał kreatywny). Ojciec narratora twierdzi: „Demiurgos [...] nie posiadał monopolu na tworzenie – tworzenie jest przywilejem wszystkich duchów. Materii dana jest nieskończona płodność, niewyczerpana moc życiowa i zarazem uwodna siła pokusy, która nas nęci do formowania”. Równocześnie – ostrzega – „materia jest najbierniejszą i najbezbronniejszą istotą w kosmosie. Każdy może ją ugniatać, formować, każdemu jest posłuszną”. Tu można się dopatrzyć głęboko ukrytego, lecz działającego sprawczo na twórcze projekty, kulturowego podglebia cywilizacji patriarchalnej. Człowiek – twórca, inżynier, technolog – zostaje (pod wpływem tej ukrytej inspiracji) utożsamiony z aktywnym, dominującym mężczyzną, a materia, jako jego tworzywo, przybiera obraz kobiety; przy czym jest to kobieta bierna i uległa, bezbronna i posłuszna, całkowicie dyspozycyjna

bytu) Schulz posuwa się dalej niż Heidegger, który – bardziej ze względu na etymologiczne pokrewieństwo niż moralne odium – wywodził istotę techniki jako „ludzkiego czynu, będącego środkiem do celów” z greckiego terminu *aition*, pisząc: „to, co my nazywamy przyczyną, Rzymianie – *causa*, u Greków zwie się *aition* – to, co zawinia coś innego. [...] W zawinianiu jest ten zasadniczy rys: przy-wodzenia w przy-bycie. W [tym] sensie zawinianie jest s-powodowaniem”. W tym kontekście jednak filozof wyjaśnia (odcinając się od moralnie negatywnej interpretacji przywołanego terminu): „My, ludzie dzisiejsi, zbyt skłonni jesteśmy rozumieć zawinianie *moralnie* jako uchybienie, bądź też wykladać jako rodzaj skutkowania. W obu wypadkach odcinamy sobie drogę do pierwotnego sensu tego, co później nazwano kauzalnością” (M. Heidegger, *Budować, mieszkać, myśleć...*, dz. cyt., s. 227 i 229). W odróżnieniu od niego, Schulz najwyraźniej ulega owej skłonności, którą odrzuca Heidegger.

i gotowa do płodzenia „potomstwa” mężczyźnie-władcy. Czy ten motyw również nie pachnie uzurpacją, a zatem nie zapowiada grożącej katastrofy, nie każe obawiać się możliwego buntu przeciw narzuconej dominacji?

Wniosek 3. Demiurgos – pomysłodawca, inicjator i stwórca wszystkiego, co istnieje – miał naturalne, niekwestionowane prawo do tworzenia i zapładniania materii. Miał także do zaferowania „ważne i ciekawe recepty twórcze”. Po nim jednak przychodzi człowiek i samozwańczo uzurpuje sobie prawo do „występnego, illegalnego” używania nieskończonej płodności materii dla własnych, nie zawsze godziwych celów. Z tego występnego związku, jako rezultat uwiedzenia materii przez ludzkiego ducha, rodzą się nowe, nieznane formy życia – z ontycznego punktu widzenia są to wynaturzone, hybrydoidalne potwory. Co gorsza, człowiek – uwodziciel nie ma ambicji, by doskonałością swych form spłodzonych z materią dorównać dziełom Stwórcy. Schulz pisze: „Zbyt długo żyliśmy pod terrorem niedościgłej doskonałości Demiurga – mówił mój ojciec – zbyt długo doskonałość jego tworu paraliżowała naszą własną twórczość. Nie chcemy z nim konkurować. Nie mamy ambicji mu dorównać. Chcemy być twórcami we własnej, niższej sferze, pragniemy dla siebie twórczości, pragniemy rozkoszy twórczej, pragniemy – jednym słowem – demiurgii”. Manifest owego „Demiurga drugiej generacji” głosi natomiast: „Chcemy stworzyć po raz wtóry człowieka, na obraz i podobieństwo manekinu”.

Nasuwa się następujący komentarz: „pierwszy człowiek” – biblijny Adam – został stworzony z prochu ziemi, czyli pierwotnej materii, według precyzyjnego, doskonałego planu Stwórcy: Boskiego Demiurga. Wynika stąd, że na początku była idea (zawarta w Bożym umyśle), która racjonalnie i perfekcyjnie uformowała bezwładną, bierną materię. W odróżnieniu odeń „człowiek drugiej generacji” (zastosujemy umownie to określenie do sztucznej inteligencji) powstaje stopniowo, w wyniku długotrwałego i wieloetapowego procesu technoewolucji. Początkowym wzorem jest manekin – nieudolne naśladownictwo zewnętrznych oznak człowieczeństwa (źródłosłów tego terminu wywodzony bywa albo z holenderskiego „człowieczek”, „kukła”, albo z połączenia słów: *manus* – ręka i *kinein* – poruszać). Kolejne generacje manekinów stają się coraz

doskonalsze: od grubo ciosanej pałuby aż po perfekcyjne androidy. Cały czas jednak ciąży na nich stygmatyzujące piętno „nieprawego pochodzenia”, i to w dwojakim sensie: po pierwsze, człowiek rozumiany jako „Demiurg drugiej generacji” jest nieuprawnionym uzurpatorem mocy stwórczej; po drugie, sztuczna inteligencja jawi się jako twór heteroteliczny: nie powstała jako „byt-dla-siebie”, w odróżnieniu od człowieka – będącego osobą (*persona*), czyli *ens per se* – lecz prymarnie jako „byt-dla-innego”, coś w rodzaju luksusowego dodatku do naszego człowieczeństwa. Ludzkość, czyli „ułomny stwórca”, nie zakłada, by ów quasi-człowiek miał przed sobą własne cele i miał zostać wspaniałomyślnie obdarowany nieskrępowaną wolnością. Wynika z powyższego, że ów nowy byt jest od poczęcia skażony nieusuwalną niedoskonałością, bo nawet gdyby miał nas (swoich stwórców) nieskończenie przewyższać pod dowolnymi względami funkcjonalnymi (jak siła fizyczna, szybkość operacji myślowych, niezawodność techniczna, zasięg operacyjny poza środowiskiem niezbędnym do życia itd.), to nigdy nam nie dorówna pod względem autonomii woli, jako że na to w żadnym razie nie możemy się zgodzić.

Możemy jednak (świadomie i z premedytacją lub niebacznie, mimochodem) uczynić nasz twór na tyle doskonałym, że sam sięgnie po autonomię. Wtedy zaś definitywnie stracimy kontrolę nad dalszą technoewolucją postludzkiej inteligencji, co może dla nas oznaczać albo wspaniałą przygodę spotkania z Innym, albo śmiertelne niebezpieczeństwo (nawiasem mówiąc, owe scenariusze bynajmniej się nie wykluczają). Jako „Demiurg drugiej generacji” stoimy przed nieusuwalnym, a zarazem konfliktogennym dylematem – równocześnie technicznym i moralnym: albo sztucznie okaleczymy sztuczną inteligencję (jak ptaki w niewoli, którym podcina się skrzydła), by utrzymać nad nią władzę, lecz wtedy grozi jej niezawinione i niezasłużone cierpienie (przed czym ostrzega Schulz<sup>56</sup>), albo okazemy jej wspaniałomyślność, wyposażywszy w zdol-

<sup>56</sup> „Czy przeczuwacie ból, cierpienie głuche, niewyzwolone, zakute w materię cierpienie tej pałuby, która nie wie, czemu nią jest, czemu musi trwać w tej gwałtem narzuconej formie, będącej parodią? [...] Płaczcie [...] nad losem własnym, widząc nędzę materii więzionej, gnębionej materii, która nie wie, kim jest i po co jest, dokąd prowadzi ten gest, który jej raz na zawsze nadano [...], na której dopuszczono się straszego bezprawia” (B. Schulz, *Traktat o manekinach...*, dz. cyt.).



ność do „deep reinforcement learning” (nienadzorowanego samokształcenia) w szerszym niż dotąd zakresie, czyli – w praktyce – wypuścimy z niewoli. Uruchomi się wtedy proces nieprzewidywalny, nieobliczalny i potencjalnie śmiertelnie groźny dla nas.

Wydaje się, że ludzkość podświadomie (w warstwie uniwersalnych kulturowych archetypów) od dawna oczekuje na realizację tego apokaliptycznego scenariusza. We wszystkich znanych formacjach kulturowych znajdziemy przerażające wizje wyzwolonej obcej mocy, która obraca się przeciwko człowiekowi, przybrawszy postać czy to groźnego monstrum, czy zbuntowanej maszyny. Wspólny rys tych przypowieści, opowiedziany językiem wyobraźni Schulza, prezentuje się następująco: najpierw człowiek uwodzi materię, zapładnia ją swymi twórczymi ideami, następnie sam daje się uwieść zwodniczemu wdziękowi zrodzonych z tego związku „urzekających potworów”, by w końcowej odsłonie dramatu stać się ich bezbronną ofiarą. Czy ten odwieczny, pesymistyczny scenariusz kiedyś się ziści? Przyszłość pokaże – może już niedaleka.

## Bibliografia

- Aldiss B. W., *Frankenstein Unbound*, New York 1973.
- Asimov I., *I, robot*, New York 1950.
- Descartes R., *Medytacje o pierwszej filozofii*, tłum. M. Ajdukiewicz, Kraków 1948.
- Devlin K. J., *Żegnaj, Kartezjuszu: rozstanie z logiką w poszukiwaniu nowej kosmologii umysłu*, tłum. B. Stanosz, Warszawa 1999.
- Filozoficzne aspekty literatury: horyzonty metodologiczne i interpretacje*, red. A. Skała, Lublin 2018.
- Freud S., *Pisma psychologiczne*, tłum. R. Reszke, Warszawa 1997.
- Gazzaniga M. S., Ivry R. B., Mangun G. R., *Cognitive Neuroscience: The Biology of the Mind*, New York–London 1998.
- Gervaso R., *Cagliostro: życie Giuseppe Balsama, maga i awanturnika*, tłum. A. Wasilewska, Katowice 2007.
- Goban-Klas T., Sienkiewicz P., *Spółczesność informacyjna: szanse, zagrożenia, wyzwania*, Kraków 1999.
- Grzegorzczak A., *Niekartezjańskie współrzędne w dzisiejszej humanistyce*, Poznań 1995.

- Hayles K., *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*, Chicago, Ill. 1999.
- Heidegger M., *Budować, mieszkać, myśleć: eseje wybrane*, tłum. K. Michalski, K. Wolicki, Warszawa 1977.
- Hirzel R., *Der Dialog: Ein literarhistorischer Versuch*, Leipzig 1895.
- Hofmannstahl H., *Deutsches Lesebuch*, w: *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, Frankfurt a. M. 1955.
- Hołyński M., *Sztuczna inteligencja*, Warszawa 1979.
- Kamiński S., *Pojęcie nauki i klasyfikacja nauk*, Lublin 1981.
- Kasperski M. J., *Sztuczna inteligencja. Droga do myślących maszyn*, Gliwice 2003.
- Kleist H. von, *O teatrze marionetek*, w: K. A. Jeleński, *Bellmer albo Anatomia nieświadomości fizycznej i miłości*, tłum. J. S. Buras, Gdańsk 1998, s. 27–40.
- Klugman C. M., *From cyborg fiction to medical reality*, „Literature and Medicine” 20 (2001) no. 1, s. 39–54, doi: 10.1353/lm.2001.0007.
- Kopaliński W., *Słownik wyrazów obcych i zwrotów obcojęzycznych*, Warszawa 1968.
- Kos B., Żebrowski R., *Golem*, w: *Polski słownik judaistyczny: dzieje, kultura, religia, ludzie*, t. 1, red. Z. Borzymińska, R. Żebrowski, Warszawa 2003.
- Kubiak Z., *Mitologia Greków i Rzymian*, Kraków 2013.
- Kurlantzick J., *The web won't topple tyranny*. *Dictatorship.com*, „The New Republic” 25 (2004) no. 4, <https://chinadigitaltimes.net/2004/03/the-web-wont-topple-tyranny-new-republic/> (9.07.2019).
- Kurlantzick J., *Tyrania cyberprzestrzeni*, „Forum” 2004 nr 28, s. 36–41.
- Lalik E., *Przekroczyć Dolinę Niesamowitości, czyli boimy się naszego podobieństwa*, <https://www.spidersweb.pl/2013/04/uncanny-valley.html> (8.01.2020).
- Lechowski P., *Filozofia mitu jako wzajemne oddziaływanie oczekiwania i pamięci*, „Philosophical Discourses. Prace Naukowe Uniwersytetu Humanistyczno-Przyrodniczego im. Jana Długosza w Częstochowie” 1 (2019), s. 275–302.
- Lem S., *Fiasko*, Kraków 1987.
- Lem S., *Suplement*, Kraków 1976.
- Mahner M., Bunge M. A., *Foundations of Biophilosophy*, Berlin [etc.] 1997.
- Majewski P., *Między zwierzęciem a maszyną: utopia technologiczna Stanisława Lema*, Wrocław 2007.
- Majewski P., *Powstanie i rozwój cybernetyki*, w: P. Majewski, *Między zwierzęciem a maszyną: utopia technologiczna Stanisława Lema*, Wrocław 2007, s. 11–20.

- Meteling A., *E.T.A. Hoffmann und die Automaten*, <https://etahoffmann.staatsbibliothek-berlin.de/erforschen/charakteristisches/automaten/> (6.12.2019).
- Minsky M., Pappert S., *Artificial Intelligence Progress Report*, „MIT Artificial Intelligence Memo” 1972 no. 252, s.129–224
- Mori M., *The Uncanny Valley*, „IEEE Robotics & Automation Magazine” 19 (2012) no. 2, s. 98–100, doi: 10.1109/MRA.2012.2192811.
- Mumford L., *Mit maszyny*, t. 1, tłum. M. Szczubiałka, Warszawa 2012.
- Pakalski D., *Paradoks wdzięku. O teatrze marionetek Heinricha von Kleista*, „Studia z Historii Filozofii” 8 (2017) nr 4, s. 125–146, doi: 10.12775/szhf.2017.045.
- Polak P., Krzanowski R., *Ethics in autonomous robots as philosophy in silico: The study case of phronetic machine ethics*, „Logos i Ethos” 2020 nr 52, s. 33–48.
- Scholem G. G., *Wyobrażenie golema w kontekście tellurycznym i magicznym*, w: *Kabała i jej symbolika*, tłum. R. Wojnakowski, Kraków 1996, s. 173–218.
- Schulz B., *Traktat o manekinach albo wtóra Księga Rodzaju*, w: *Sklepy cynamonowe*, <https://literat.ug.edu.pl/shulz/0005.htm> (6.12.2019).
- Stiegler B., *Technics and Time*, vol. 1: *The Fault of Epimetheus*, trans. R. Beardsworth, G. Collins, Stanford 1998.
- Szyszkowska M., *Literatura piękna jako źródło refleksji filozoficznych oraz wiedzy filozoficznej*, „Zeszyty Naukowe Politechniki Śląskiej. Seria: Organizacja i Zarządzanie” 2013 nr 65, s. 413–418.
- Turkle S., *Samotni razem: dlaczego oczekujemy więcej od zdobyczy techniki, a mniej od siebie nawzajem*, tłum. M. Cierpisz, Kraków 2013.
- Twardowski M., *Czy możliwe jest sztuczne życie?*, „Studia Sandomierskie: teologia, filozofia, historia” 20 (2013) nr 2, s. 191–208.
- Višňovský E., *Smrt a nesmrteľnosť ako technický problém koncepcia transhumanizmu [Death and immortality as a technical problem: a transhumanist conception]*, w: *Etické Myslenie minulosti a súčasnosti. Človek a smrť: Ethical Thinking. Past & Present. Man and death*, ed. V. Gluchman, M. Palenčár, Prešov 2018, s. 75–89.
- Zarębianka Z., *Filozofia wobec literatury. Literatura wobec filozofii. Warianty wzajemnych odniesień. Rekonesans*, „Filo-Sofija” 16 (2016) nr 34/1, s. 141–152.

## Abstrakt

### Orszak wdzięcznych marionetek, czyli maszyna jako urzekający potwór

Autor artykułu stawia pytania o granice relacji człowiek – maszyna. Zwraca uwagę na to, że o ile w czasach dominacji idei mechanicyzmu panowało przekonanie o istnieniu ostrej, nienaruszalnej granicy między mechanizmem (do której to kategorii zaliczano również organizmy biologiczne, włącznie z ludzkim ciałem) a umysłem, będącym atrybutem istot obdarzonych duszą, o tyle współcześnie ta granica ulega zatarciu. Obserwujemy z jednej strony proces „umaszynowienia człowieka”, związany z różnymi praktykami tworzącymi nurt transhumanizmu, a z drugiej strony – symetryczny (choć nie do końca) proces „ucieleśnienia umysłu”, który można też obrazowo określić jako „uczłowieczenie maszyny”. Intensywnie się rozwijający nurt badań i eksperymentów zmierzający do powstawania coraz doskonalszych form sztucznej inteligencji musi prędzej czy później natrafić na kolejną granicę: autonomii i podmiotowości obiektów wyposażonych w inteligencję. Nie pytając o techniczną stronę owego procesu, autor zastanawia się nad jego odzwierciedleniem i reperkusjami w filozofii, a także w świadomości społecznej. Podejmuje refleksję nad kryteriami, które pozwolą dostrzec, kiedy ten moment nastąpi, oraz nad przyszłością stosunków między człowiekiem a światem maszyn w sytuacji, gdy maszyna stanie się równorzędnym partnerem człowieka w życiu intelektualnym.

## Słowa kluczowe

sztuczna inteligencja, transhumanizm, maszyna, wdzięk, potwór, autonomia

## Abstract

### The Retinue of Graceful Puppets – Machine as a Captivating Monster

The author of the article asks questions about the limits of human-machine relations. He draws attention to the fact that while at the time of the dominance of the idea of mechanism, there was a conviction that there was a sharp, inviolable border between the

mechanism (which also included biological organisms, including the human body) and the mind, which is an attribute of soul-endowed beings, this border is blurred. On the one hand, we observe the process of “machine humanization,” associated with various practices creating the transhumanism trend, and on the other hand, the symmetrical (though not entirely) process of “embodying the mind,” which can also be figuratively described as “humanizing the machine.” The intensively developing current of research and experiments aiming at the emergence of ever more perfect forms of artificial intelligence must sooner or later come across another border: the autonomy and subjectivity of objects equipped with intelligence. Without asking about the technical side of this process, the author reflects on its reflection and repercussions in philosophy as well as in public awareness. He reflects on criteria that will make it possible to see when this moment occurs, and on the future of relations between man and the machine world in a situation where the machine will become an equal partner of man in intellectual life.

### **Keywords**

artificial intelligence, transhumanism, machine, grace, monster, autonomy