

Małgorzata ABASSY

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

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

Celem artykułu jest udzielenie odpowiedzi na pytanie o rolę narracji – opowieści o sztucznej inteligencji, jakich słuchamy i jakie opowiadamy, dając w ten sposób wyraz naszym nadziejom i obawom – w oswojaniu sztucznej inteligencji. Zakładamy, że kreujemy w ten sposób rzeczywistość, najpierw za pomocą słów, a następnie – poprzez wybrane działania. Performatywna funkcja języka¹, chociaż marginalizowana w badaniach nad sztuczną inteligencją², pojawia się w powieści Neala Stephensona *Zamieć*³. Realizowana jest również, najpierw jako dialog wewnętrzny bohatera, a później w formie działania w tekście Iana McEwana *Maszyny takie jak ja*⁴. Rolę literatury jako jednego z wtórnych systemów modelujących kultury podkreślali semiotycy szkoły tartusko-moskiewskiej⁵. Teoria semiotyczna Jurija Łotmana stanowi rozszerzenie teorii o fatycznej funkcji języka, dostarczając materiału do refleksji nad funk-

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

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

³ Zob. N. Stephenson, *Zamieć*, tłum. J. Polak, Warszawa 2009.

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

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

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

Wybór materiału źródłowego został podyktowany czynnikiem czasowym – powstanie tekstów dzieli ponad dwadzieścia lat – oraz problemami, które, sygnalizowane w pierwszym z analizowanych tekstów, zyskały rozwinięcie w drugim. W *Zamieci* główny bohater Hiro otrzymuje ostrzeżenie: „– Jesteś hakerem. Musisz troszczyć się o swoje głębokie struktury. – Jakie struktury? – Neurolingwistyczne ścieżki w mózgu. Pamiętasz, jak uczyłeś się kodu dwójkowego? – Owszem. – Przecierałeś wówczas nową ścieżkę w mózgu. Głęboką strukturę. Gdy korzystasz z niej, twoje nerwy tworzą nowe wypustki [...] aksony dzielą się i wpychają między gangliony i neurofibryle [...] twój bioware ulega samodoskonaleniu [...] software staje się częścią hardware’u”⁶. Mowa tutaj o manipulacji zewnętrznej, jednak zniekształcanie wyobrażeń o rzeczywistości poprzez kwestionowanie jej podstawowych znaczeń może być również wynikiem dialogu wewnętrznego. Ostrzeżenie przed przekształcaniem ścieżek neuronalnych w mózgu, odpowiedzialnych za istnienie słów, którymi określamy rzeczywistość, staje się faktem, gdy główny bohater powieści *Maszyny takie jak ja* uznaje robota Adama za „przedstawiciela tego samego co ja gatunku”⁷. W powieści *Zamieć* problem języka można traktować jako wstęp do refleksji nad pierwotnym systemem modelującym i warunkami jego przekształcania bądź zniekształcania. Z kolei *Maszyny takie jak ja* to przykład elementu wtórnego systemu modelującego: zawiera on obraz humanoidalnej maszyny funkcjonującej w relacjach z ludźmi. Nie będąc rozprawą filozoficzną, proponuje odpowiedzi na istotne pytania o podmiotowość. Pojawiająca się w powieści kreacja postaci Alana Turinga – matematyka, kryptografa, pomysłodawcy eksperymentu mającego na celu zdefiniowanie sztucznej inteligencji⁸, stanowi dodatkowy czynnik wskazujący na płynność granic między fikcją literacką a rzeczywistością. Ostatecznie to postać wynalazcy udziela w powieści odpowiedzi na pytanie, czy artefakt sztucznej inteligencji – Adam – jest przedmiotem czy podmiotem.

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

⁶ Stephenson, dz. cyt., s. 121.

⁷ McEwan, dz. cyt., s. 104.

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

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

Tytułowe powieści dzieli ponad dwadzieścia lat: *Zamieć* (*Snow Crash*⁹) została napisana w roku 1992, podczas gdy druga ze wspomnianych pozycji, *Maszyny takie jak ja* (*Machines Like Me*¹⁰), ukazała się w 2019. Łączy je fakt, że stanowią namysł nad przyszłością, którą ludzkość tworzy, wykorzystując język jako narzędzie kreacji. Są projektowaniem nadziei i obaw. Ich autorzy, wychwytyjąc wątki codziennych narracji, stawiają pytania o granice światów i gatunków. W tle brzmi również niepostawione wprost pytanie o ludzką sprawczość, a ostatecznie o odpowiedzialność za świat, w którym przychodzi nam żyć. Przyspieszenie informacyjne sprawiło, że obecne pokolenia nie tworzą rzeczywistości dla swoich potomków, lecz dla siebie samych: nasze „dzisiaj” błyskawicznie staje się naszym „jutro”. Prognozowane przez Stephensona narodził się Metawersum – równoległej rzeczywistości wirtualnej współistniejącej z realną, stają się faktem wskutek zabiegów Maxa Zuckenberg, właściciela platformy społecznościowej Facebook¹¹. Czy literacka postać humanoidalnej maszyny – „Adama”, pojawiająca się na kartach powieści McEwana również zaistnieje w naszym życiu jako partner, przyjaciel i kochanek? Narracja, która rozpoczyna powieść, nie pozostawia co do tego wątpliwości: „Była to nadzieja zrodzona z religijnej tęsknoty, prawdziwy Święty Graal nauki. Nasze ambicje były niebosiężne i przyziemne – chodziło o potwierdzenie mitu stworzenia i o monstualny akt miłości własnej. Kiedy tylko stało się to możliwe, nie mieliśmy innego wyjścia: musieliśmy zrealizować nasze marzenia i ponieść wszystkie tego konsekwencje. Mówiąc górnolotnie, chcieliśmy uciec przed

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

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

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

własną śmiertelnością, stanowić czoło boskości albo nawet zastąpić ją idealną jaźnią¹².

„OSWAJANIE” I „SZTUCZNA INTELIGENCJA” SEMANTYKA POJĘĆ

W powieści *Mały książę* Antoine’a de Saint Exupery’ego między Księciem a lisem toczy się znamieny dialog o znaczeniu słowa „oswajac”: „Szukam przyjaciół. Co to znaczy «oswoić»? – To rzecz, o której niestety zapomniano – powiedział lis. – To znaczy «stworzyć więzi»¹³. „Oswajanie” to proces o znaczeniu egzystencjalnym. Nie tylko pozwala dostrzec wyjątkowość i podmiotowość: „Jeżeli jednak mnie oswoisz, będziemy sobie wzajemnie potrzebni. Będziesz dla mnie jedyny na świecie. I ja będę dla ciebie jedyny na świecie”¹⁴, ale także otwiera na nowe postrzeganie rzeczywistości: „Jeżeli jednak mnie oswoisz, moje życie będzie jakby opromienione słońcem”¹⁵. Czy taką definicję oswajania, jaka wyłania się z rozmowy księcia i lisa, możemy stosować do naszych relacji ze sztuczną inteligencją? Zakładałaby ona tworzenie więzi, a więc podmiotowość stron wchodzących w interakcję. Odruchową reakcją na wysuniętą hipotezę jest odrzucenie podmiotowości sztucznej inteligencji, nawet w jej humanoidalnej postaci¹⁶. Warto jednak zadać sobie pytanie, czy w głębi naszej świadomości nie odbywa się niezauważalne zacieranie granic i mieszanie pojęć¹⁷.

Terminu „sztuczna inteligencja” po raz pierwszy użył amerykański informatyk John McCarthy¹⁸ podczas konferencji w Darmouth w roku 1955. W definicji zawartej w *Encyclopedia Britannica* czytamy: „Sztuczna inteligencja to zdolność komputera cyfrowego lub robota sterowanego komputerowo do wykonywania zadań powszechnie kojarzonych z inteligentnymi istotami. Termin ten jest często stosowany do projektu rozwoju systemów obdarzo-

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

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

¹⁴ Tamże.

¹⁵ Tamże.

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

¹⁷ Zob. S. L e m, *Dzienniki gwiazdowe*, Wydawnictwo Literackie, Kraków 2019.

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

nych procesami intelektualnymi charakterystycznymi dla człowieka, takimi jak zdolność rozumowania, odkrywania znaczenia, uogólniania lub uczenia się z przeszłych doświadczeń¹⁹. Na oficjalnej stronie rządu Rzeczypospolitej Polskiej widnieje następująca definicja: „Sztuczną inteligencję (ang. Artificial Intelligence – AI) próbuje się definiować jako dziedzinę wiedzy obejmującą m.in. sieci neuronowe, robotykę i tworzenie modeli zachowań inteligentnych oraz programów komputerowych symulujących te zachowania, włączając w to również uczenie maszynowe (ang. machine learning), głębokie uczenie (ang. deep learning) oraz uczenie wzmocnione (ang. reinforcement learning)”²⁰. Wskazując na zakres zjawisk, definicje nie precyzują ich istoty. Człowiek jednak rozpoznaje pewne podobieństwo procesów zachodzących w maszynach do tych, które odbywają się w obrębie jego własnych sieci neuronowych: myślenia, wnioskowania, podejmowania decyzji, a nawet odwzorowywania reakcji w sferze zarezerwowanej, jak mogłoby się wydawać, wyłącznie dla istot ludzkich²¹. Sztuczna inteligencja jest pojęciem słabo rozpoznanym w kulturze masowej. Przesłania je postać cyborga oraz generowane przez nią znaczenia. Te ostatnie stanowią wypadkową oglądanych filmów, zasłyszanych opowieści oraz zanurzenia w media społecznościowe, gdzie zacierają się granice między tym, co fizyczne, a tym, co wirtualne, tym, co ludzkie, a tym, co nie-ludzkie; unieważnione zostają czas i przestrzeń. Potocznie sztuczna inteligencja kojarzy się z autonomicznymi maszynami, mogącymi wyrwać się spod ludzkiej kontroli. Szybki rozwój nowych technologii, wypieranie z rynku pracy ludzi na rzecz mniej zawodnych i tańszych w eksploatacji maszyn, spuścizna kultury masowej, zawierająca w sobie literackie i filmowe obrazy buntu maszyn, oraz brak zgodnej odpowiedzi ekspertów na pytanie, czy powinniśmy obawiać się rozwoju sztucznej inteligencji, czy raczej pokładać w niej nadzieję²², skłania do przyjęcia jeszcze jednej, nieobejmującej aspektu egzystencjalnego, definicji osvajania – jako „przyzwyczaić do kogoś lub do czegoś albo zapoznać z czymś; przyzwyczaić dzikie zwierzęta do przebywania wśród ludzi lub do służenia ludziom”²³.

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

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

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

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

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

Postacią, wyłaniającą się z tekstów kultury masowej, jest cyborg – połączenie maszyny i człowieka, obiektu nauki o systemach – cybernetyki. Twórcą tego ostatniego pojęcia, wywodzącego się od greckiego *kybernetikos*, jest Norbert Wiener, który w połowie dwudziestego wieku sięgnął do wczesnych hellenistycznych teorii dotyczących zarządzania i połączył je z ideami pojawiającymi się w dziedzinie informatyki, tworząc w ten sposób nowe interdyscyplinarne podejście do systemów²⁴. Cybernetyka koncentrowała się na przepływach informacji między systemami, które obejmowały komponenty organiczne i obliczeniowe. Znaczenie podejścia Wienera trudno przecenić. Jak zauważyła N. Katherine Hayles, twórczyni teorii posthumanizmu: „Ze wszystkich implikacji, jakie niesie ze sobą cybernetyka pierwszej fali, być może żadna nie była bardziej niepokojąca i potencjalnie rewolucyjna niż ta, że granice ludzkiego podmiotu są konstruowane, a nie dane. Postrzegając kontrolę, komunikację i informację jako zintegrowany system, cybernetyka radykalnie zmieniła sposób, w jaki pojmowano granice”²⁵. Wiener zdawał sobie sprawę z konsekwencji owych implikacji, stwierdzając, że kolejnym problemem, z jakim ludzkość będzie się mierzyć będzie relacja między człowiekiem a maszyną²⁶. Choć w wczesnych ujęciach systemowych nie było nacisku na ontologiczne połączenie tego, co organiczne, i tego, co sztuczne, Wiener zdawał sobie sprawę, że ludzka ciekawość i skłonność do przesuwania granic mogą doprowadzić do zacierania się granic między „organicznym” a „mechanicznym”.

„ZAMIEĆ” JAKO ZAMĘT UMYŚŁU

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

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

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

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

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

²⁷ S t e p h e n s o n, *Zamięć*, s. 44.

²⁸ Tamże, s. 43.

²⁹ Tamże, s. 73.

³⁰ Tamże, s. 179.

³¹ Tamże, s. 188.

³² Tamże.

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

GRANICE CZŁOWIECZEŃSTWA „MASZYNY TAKIE JAK JA”

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

³³ Tamże, s. 261.

³⁴ Tamże, s. 379.

³⁵ Tamże, s. 333.

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

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

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

³⁷ Tamże, s. 10.

³⁸ Tamże, s. 11.

³⁹ Tamże, s. 19.

⁴⁰ Tamże, s. 20.

⁴¹ Tamże, s. 37.

⁴² Tamże.

znosił to niedobrze. [...] Byłem pierwszym mężczyzną, któremu przypisał rogi artefakt. [...] Sam Alan Turing często mówił w młodości, że kiedy maszyny od człowieka nie da się odróżnić po zachowaniu, będziemy musieli przypisać jej człowieczeństwo, dlatego gdy nocne powietrze rozdarł nagle ekstatyczny krzyk Mirandy [...] potraktowałem Adama jak przedstawiciela tego samego co ja gatunku, ze wszystkimi wynikającymi stąd przywilejami i zobowiązaniami. Znienawidziłem go⁴³. Obawie, graniczącej z pewnością, że mężczyźni staną się zbędni⁴⁴, towarzyszy narastające przekonanie, że Adam jest prawdziwszą i lepszą wersją człowieka. Narracja obfituje w odwołania do Turinga, historia relacji człowieka i maszyny zostaje zaś podsumowana tyradą sławnego wynalazcy, będącą odpowiedzią na zniszczenie Adama: „Mam nadzieję – podjął chłodno – że pewnego dnia to, co pan zrobił Adamowi, zostanie uznane za poważaną zbrodnię [...]. Próbował pan zniszczyć życie. On był czującą istotą. Miał swoją jaźń. To, jak powstała, dzięki biologicznym neuronom, mikroprocesorom czy sieciom neuronowym, nie ma znaczenia⁴⁵. Uznając podmiotowość maszyny, należy zakwestionować słowa narratora: „Ja go kupiłem i ja miałem prawo go zniszczyć⁴⁶. Powieść, traktowana jako element wtórnego systemu modelującego, zawiera opowieść o współistnieniu ludzi i maszyn jako pełnoprawnych członków przestrzeni również aksjologicznej.

*

Obydwie powieści wybrane jako materiał źródłowy do analizy mechanizmów oswojania sztucznej inteligencji zawierają w sobie przesłanki do odczytania zarówno obaw, jak i nadziei pokładanych w rozwoju technologicznym. Poprzez ich werbalizację na kartach książki, a następnie wizualizację w wyobraźni czytelnika rodzi się kolejna historia, dając impuls do zaistnienia nowej rzeczywistości. Wagę narracji jako narzędzia kreacji podkreśliła Hayles: „Nosimy ze sobą nasze przeszłe historie jak ślimak skorupę⁴⁷. Wtórjuje jej Donna Haraway: „Mówiąc metaforycznie, jesteśmy jak krab pustelnik inkrustowany pąklami. Postrzegam siebie i wszystkich innych jako zmieniających muszle w miarę, gdy rośniemy. Ale każda muszla, którą podnosimy, ma swoją historię i z pewnością jej nie wybiera-

⁴³ Tamże, s. 101-104.

⁴⁴ Por. tamże, s. 103.

⁴⁵ Tamże, s. 346.

⁴⁶ Tamże, s. 318.

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

my⁴⁸. Nie tylko człowiek dziedziczy historie, które wpływają na przyswajanie i interpretację przezeń narracji płynących z otaczającego go świata. Sztuczna inteligencja w swoich wielorakich manifestacjach również obciążona jest treściami wcześniejszych opowieści: „Tak jak cyborg jest dzieckiem militarystyki i «wielkiej nauki», tak ja jestem dzieckiem katolicyzmu i «zimnej wojny»⁴⁹ – przypomina Haraway. Snucie historii nie tylko jest aktem kreacji nowych rzeczywistości, ale także może stać się aktem autokreacji. W dążeniu do przekroczenia tego, co postrzegane jest nie jako nierozpoznany potencjał, lecz jako ograniczenie, które należy przewyciężyć, stwarzamy na nowo siebie samych: „Ziściło się to ostatecznie w jesieni dwudziestego wieku: postawiono wtedy pierwszy krok na drodze ku spełnieniu się dawnego marzenia, zaczęła się długa lekcja, w trakcie której przekonał się, że niezależnie do tego, jak bardzo jesteśmy skomplikowani, jak bardzo zawodni i trudni do opisanego nawet w najprostszych działaniach i trybach istnienia, można nas skopiować i ulepszyć⁵⁰. W wybranych tekstach źródłowych funkcja języka jako narzędzia kreacji przejawia się na różne sposoby. W *Zamieci* zostajemy ostrzeżeni przed wykorzystaniem naszych głębokich, pierwotnych struktur mózgu – tych obszarów, w których rodzi się świadomość i kształtują najwcześniejsze połączenia neuronalne – przez hakerów, podmioty żądne władzy⁵¹. W *Maszyny takie jak ja* informacja o znaczeniu podana jest nie wprost, jako wewnętrzny monolog głównego bohatera, nabywcy „Adama”, próbującego rozpoznać sztuczny byt poprzez szukanie odniesień do siebie samego. Granice zacierają się: świadomość zostaje pochłonięta przez „zamięć”, „Adam” zaś zostaje potraktowany jak człowiek w kontekście aktu seksualnego – jednej z najbardziej intymnych form oddania się drugiemu człowiekowi.

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

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

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

⁵⁰ M c E w a n, dz. cyt., s. 9.

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

BIBLIOGRAFIA / BIBLIOGRAPHY

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

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

ABSTRAKT / ABSTRACT

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

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

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

Słowa kluczowe: sztuczna inteligencja, podmiotowość, język, narracje

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

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

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

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

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

Keywords: artificial intelligence, subjectivity, language, narratives

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

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

Google Scholar: 5HDH-8MAAAAJ

ORCID: 0000-0002-9294-6156