

Anna MAJ

EWOLUCJA INTELIGENCJI I JEJ BADANIE Oswajanie sztucznej inteligencji w perspektywie przemian kultury i komunikacji

Roboty wyposażone w sztuczną inteligencję, które w komunikacji z człowiekiem przejawiają autoironię i działania irracjonalne, postrzegane są przez ludzi jako bardziej ludzkie i budzą pozytywniejsze odczucia. Sugerując, że także (jak ludzie) popełniają błędy, realizują pewien algorytm, dzięki któremu w ludzkim odbiorze stają się bardziej do nas podobne. Taka symulowana niekonsekwencja i niedoskonałość byłoby zatem tymi elementami, które sprawiają, że technologia przestaje budzić lęk i wydaje się bliższa innym doświadczeniom komunikacyjnym.

OSWAJANIE SZTUCZNEJ INTELIGENCJI JAKO PROCES KULTUROWY

„Oswoić” znaczy: przemyśleć, udomowić, nauczyć się komunikować, wreszcie przywyknąć. Proces udomowienia sztucznej inteligencji już trwa. Wciąż uczymy się z nią komunikować, jednak zdążyliśmy do niej przywyknąć. Nie zauważamy jej, choć towarzyszy nam prawie wszędzie, subtelnie i „bezszwowo” wnikając w naszą codzienność, zgodnie z ideą ubicomp i calm technologies¹. Elon Musk, wynalazca, przedsiębiorca i propagator AI, zwraca uwagę, że musimy przemyśleć, w którym kierunku będzie się ona rozwijać². Zaproponowane przez niego wbudowanie w proces badań nad sztuczną inte-

¹ Na idei calm technologies i paradygmacie wszechobecnego przetwarzania danych (ubiquitous computing, ubicomp) opierają się współczesne urządzenia komunikacyjne i ich interfejsy: komputery, laptopy, smartfony, interaktywne tablice, urządzenia zarządzające inteligentnym domem (ang. smart home) i miastem (ang. smart city) oraz wszystkie urządzenia i sensory związane z Internetem Rzeczy (ang. Internet of Things). Wdrażane dziś koncepcje przeprojektowania interfejsów nowych technologii tak, by stały się one przyjazne dla użytkownika, powstały w laboratoriach Xerox PARC u progu lat dziewięćdziesiątych (por. M. W e i s e r, *The Computer for the 21st Century*, „Scientific American” 256(1991) nr 3, s. 94-104; M. W e i s e r, J.S. B r o w n, *Designing Calm Technology*, Xerox PARC, 21.12.1996, <https://people.csail.mit.edu/rudolph/Teaching/weiser.pdf>).

² Po lekturze *Superinteligencji* Nicka Bostroma w roku 2014 Musk, twórca PayPal, Tesli, SpaceX i innych innowacyjnych przedsięwzięć, wyraził przekonanie, że sztuczna inteligencja może być groźniejsza od bomby atomowej. Z jego inspiracji powstała wówczas inicjatywa Open AI oraz Future of Life Institute. Open AI opisuje swoje cele jako prowadzenie badań i wdrożeń w obszarze AI, której „misją jest zapewnienie, że ogólna sztuczna inteligencja przyniesie korzyści całej ludzkości” (Open AI, *About*, <https://openai.com/about/>). Por. N. B o s t r o m, *Superintelligence: Paths, Dangers, Strategies*, Oxford University Press, Oxford 2014 [wydanie polskie: *Superinteligencja. Scenariusze, strategie, zagrożenia*, tłum. D. Konowrocka-Sawa, Wydawnictwo Helion, Gliwice 2016]; E. M u s k, „Worth reading *Superintelligence* by Bostrom. We need to be super careful

ligencją etyczno-kulturowych hamulców w postaci inicjatywy Open AI nie musi jednak zakończyć się sukcesem. Bardziej prawdopodobne jest, że stanie się katalizatorem nieuchronnych zmian. Sukcesy chatbotów Open AI (GPT2, GPT3, a następnie ChatGPT3.5³) wskazują na to, że ów „etyczny projekt AI” stanie się kolejnym akceleratorem wielu gałęzi biznesu i dokona redefinicji kultury w obszarach tak istotnych, jak: edukacja (na wszystkich poziomach), przetwarzanie tekstu, dziennikarstwo, marketing, badania naukowe⁴, eksperytyza i szkolenia, programowanie, wyszukiwanie informacji i inne usługi związane z dostępem do wiedzy⁵. Oznacza to uwikłanie „projektu etycznego AI” w niezwykle zacięty wyścig o pierwszeństwo między sieciowymi gigantami, zwanymi „cytadelami danych” (Google, Microsoft, Amazon, Meta, Facebook), które nie bez powodu zwróciły się ku sztucznej inteligencji, zwłaszcza ku przetwarzaniu języka naturalnego za pomocą dużych modeli językowych (ang. Large Language Models)⁶. Przyjęcie konkretnych zasad etycznych w projektowaniu AI staje się zatem decyzją znaczącą dla kształtu cywilizacji przyszłości.

with AI. Potentially more dangerous than nukes”, Twitter, 3.08.2014, <https://twitter.com/elonmusk/status/495759307346952192>.

³ Pierwsza wersja API GPT3 została udostępniona do testów przez Open AI z pewnymi ograniczeniami 11 czerwca 2020, a następnie bez ograniczeń 18 listopada 2021. Na jego bazie stworzono następnie GPT-3.5 oraz ChatGPT-3.5, który udostępniono bez ograniczeń 30 listopada 2022, w ciągu pierwszego miesiąca przetestowało go z powodzeniem ponad milion użytkowników. T. C h a m o r r o - P r e m u z i c, *How ChatGPT Is Redefining Human Expertise: Or How To Be Smart When AI Is Smarter Than You*, „Forbes”, 12.01.2023, <http://bit.ly/3Ef5PUM>.

⁴ W czerwcu 2022 roku A.O. Thunström zadała GPT3 napisanie autotematycznego artykułu naukowego. Proces ten oraz jego konsekwencje w aspekcie psychologicznym i kulturowym został opisany przez badaczkę w „Scientific American”. Artykuł napisany przez AI został podpisany jego imieniem (GPT3 figuruje jako główny autor, a Thunström jako współautorka) i zgłoszony do recenzji w czasopiśmie naukowym, jest dostępny online w formie preprintu we francuskim archiwum HAL OA. Eksperyment Thunström przenosi rozważania nad autorstwem sztucznej inteligencji w świecie nauki na inny poziom, dotychczas wskazywano raczej na negatywne i pozytywne aspekty praktyk wspomaganie się GPT3 w pisaniu prac naukowych, lecz nie na możliwość wykazywania AI jako współtwórcy tekstów (badań). Por. G p t G e n e r a t i v e P r e t r a i n e d T r a n s f o r m e r, A. O. T h u n s t r ö m, S. S t e i n g r i m s s o n, *Can GPT-3 write an academic paper on itself, with minimal human input?*, HAL [Open Science Archive], <https://hal.science/hal-03701250>; A. O. T h u n s t r ö m, *We Asked GPT-3 to Write an Academic Paper about Itself – Then We Tried to Get It Published*, „Scientific American”, 30.06.2022, <http://bit.ly/3IDS4YM>.

⁵ S. M i n t z, *ChatGPT: Threat or Menace?*, „Inside Higher Ed”, 16 I 2023, <http://bit.ly/412ObwI>.

⁶ Metaforę „cytadeli danych” zaproponował Martin Warnke (por. M. W a r n k e, *Databases as Citadels in The Web 2.0*, w: *UnLike Us Reader. Social Media Monopolies and Their Alternatives*, red. G. Lovink, M. Rasch, Institute of Network Cultures, Amsterdam 2013, s. 86). Projekty wielkich korporacji IT w obszarze AI reprezentują między innymi: ChatGPT-3 (Open AI we współpracy z Microsoftem, 2022), LaMDA (Google, 2021), BlenderBot3 / OPT-175B (Facebook / Meta, 2022), Alexa (Amazon, 2018).

Warto przyjrzeć się nie tylko w s p ó ł c z e s n y m p r z e j a w o m „o s w a j a n i a i n t e l i g e n c j i”, ale też bardziej nieoczywistemu p r o c e s o w i e w o l u c j i i n t e l i g e n c j i (o d b i o l o g i c z n e j, p r z e z s z t u c z n ą, p o r o z s z e r z o n ą) i dotychczasowym interdyscyplinarnym badaniom nad sztuczną inteligencją, na przykład rozwojowi sposobów rozumienia terminu „inteligencja” w tym obszarze. Z uwagi na szerokość tego typu badań przedstawione zostaną tu jedynie wybrane koncepcje i eksperymenty, pochodzące z różnych obszarów wiedzy oraz z pogranicza nauki i sztuki nowych mediów. Wieloletnie i interdyscyplinarne badania nad sztuczną inteligencją (a także cyborgizacją, robotyzacją, poznaniem i percepcją) przyczyniły się do wypracowania koncepcji i rozwiązań technicznych, które wpływają pośrednio na c y b e r k u l t u r o w e p r z e m i a n y k o n c e p t u a l i z a c j i c i a ł a i u m y s ł u oraz na w s p ó ł c z e s n e p r a k t y k i k u l t u r o w e, zwłaszcza w zakresie zachowań komunikacyjnych między innymi w sferze relacji międzyludzkich czy transformacji znaczenia przestrzeni domowej i publicznej.

Wysokie wartościowanie w cyberkulturze kategorii nowości i rozwoju technologicznego oraz atrakcyjność nowych interfejsów i funkcjonalności urządzeń sprzyjają wytworzeniu się i p o p u l a r y z a c j i p r o t e c h n o l o g i c z n e g o d y s k u r s u, który zachęca potencjalnych użytkowników nowych technologii do zaangażowania i interakcji z pomocnymi, inteligentnymi aplikacjami, agentami konwersacyjnymi⁷ i robotami. Prowadzi to do humanizacji technologii, a jednocześnie do pozytywnego wartościowania maszyny per se i postrzegania cyborgizacji jako naturalnego i dobroczynnego procesu.

⁷ W naukowej i branżowej literaturze przedmiotu stosowane są wymiennie terminy: „agent konwersacyjny”, „inteligentny agent wirtualny”, „agent sztucznej inteligencji”, „konwersacyjna sztuczna inteligencja (AI)”, „chat bot”, „bot konwersacyjny”, „asystent komunikacyjny”, „asystent głosowy”, „cyfrowy asystent/ka”, „aplikacja sztucznej inteligencji” i tym podobne. Opisują one oprogramowanie działające w określonym celu i wykorzystujące sztuczną inteligencję, umożliwiające użytkownikowi naturalną komunikację za pośrednictwem konwersacji (najczęściej głosowej, ale także tekstowej). Ich istotą jest pozyskiwanie i przetwarzanie danych nie tylko z predefiniowanej i zamkniętej bazy danych, ale też na podstawie danych pozyskanych z sieci, wielkich zbiorów danych oraz wieloetapowego procesu „uczenia się” na podstawie interakcji z użytkownikiem (ang. machine learning), które zazwyczaj trwa wiele miesięcy. Przykładami takiego oprogramowania są między innymi: Apple Siri, Google Now, Google Assistant, Microsoft Cortana, Amazon Alexa. Przykład definicji z obszaru literatury branżowej por. Microsoft Learn, *Co to są agenci sztucznej inteligencji?*, 01.12.2022, <https://learn.microsoft.com/pl-pl/azure/cloud-adoption-framework/innovate/best-practices/conversational-ai>; z kolei wczesne naukowe opracowanie kwestii terminologii por. A. J a n a s, *Zastosowanie agentów konwersacyjnych w nauczaniu*, „Nierówności Społeczne a Wzrost Gospodarczy” 2011, nr 22, s. 281. Termin „agent konwersacyjny”, jakkolwiek wydaje się niezgrabny językowo, jest jednym z częściej stosowanych przez naukowców, projektantów i producentów oprogramowania.

UMYSŁ, CIAŁO, KOMUNIKACJA

Transhumanizm i cyborgizacja, zwłaszcza koncepcja „upgrade wania” ludzkiego ciała i umysłu, dawno już przeszły od futurologicznych wizji nie tylko do naukowych (w tym medycznych) praktyk, ale i do popkulturowych realizacji⁸. Na poziomie rozszerzania możliwości umysłu w ciągu ostatniej dekady dokonała się znacząca zmiana. Prawie wszyscy przechodnie na ulicach współczesnych miast i w przestrzeniach publicznych pochłonięci są nie rozmowami z innymi ludźmi, ale zapatrzeni w ekrany smartfonów, tabletek i laptopów; delegują swe umysły do cyberprzestrzeni, zatapiając się w wirtualnym kontakcie z nieobecnymi lub poszukując informacji w sieci. Doświadczeniem komunikacyjnym, pogłębionym przez pandemię i rozwój usług sieciowych, stała się rozmowa nie z drugim człowiekiem, a z chatbotem, czyli agentem konwersacyjnym opartym na sztucznej inteligencji, takim jak Siri czy Alexa. Wykorzystywany przez tego typu oprogramowanie interfejs głosowy (zwany „konwersacyjnym interfejsem użytkownika”, CUI) symuluje do pewnego stopnia naturalną komunikację międzyludzką dzięki przetwarzaniu i modelowaniu języka naturalnego. Na razie trudno ocenić, czy zachodzi tu realny upgrade, czyli zwiększenie możliwości intelektualnych człowieka, czy częściej downgrade, dewaluacja ludzkich zdolności kognitywnych, odpowiadająca ewolucyjnie wypracowanym potrzebom szybkiej gratyfikacji, jaką daje dostęp do niekończącego się strumienia informacji⁹. Kultura partycypacji

⁸ Zgodnie z deklaracją transhumanistyczną wypracowaną przez World Transhumanist Association (1998) na podejście transhumanistyczne składają się: idea transcendencji biologii oraz przyspieszenia rozwoju człowieka za pomocą wszelkich dostępnych narzędzi technologicznych (pragmatyzm), a także memetyczna propagacja idei transhumanizmu i protechnologicznego dyskursu, idea realizacji celów transhumanizmu we wszystkich obszarach życia społecznego, uznania dla różnorodności koncepcji i idea nieustannej ewolucji w zgodzie z rozwojem technologicznym transludzkości (zob. The World Transhumanist Association. H+pedia [wiki]: *Transhumanist Declaration* [hasło], https://hpluspedia.org/wiki/Transhumanist_Declaration. Kontynuacją aktywności stowarzyszenia jest działalność Humanity+). Organizacja w swoim serwisie deklaruje przywiązanie do idei etycznego wykorzystania technologii, ale też do promowania badań i wdrożeń technologii przyszłości wykraczających poza dzisiejszą „normalną” ludzką egzystencję, to jest przedłużanie życia i poszerzenie możliwości ciała i umysłu poprzez technologiczne i biotechnologiczne protezy, bioniczne rozszerzenia i terapie genowe (zob. Humanity+, <https://www.humanityplus.org/about>). Opisywaną w tym kontekście cyborgizację można natomiast zdefiniować jako „wspomaganie funkcji umysłowych i życiowych zdrowych ludzi” (P. Łaszczyca, *Człowiek i jego maszyny. Operatorzy i protezy*, „Filo-Sofija” 17(2017) nr 39(1), s. 58). O popkulturowych realizacjach pisze: E. Olzacka, *Od Homo sapiens do Homo immortalis. Idea nieśmiertelności w rosyjskich projektach filozoficznych i społeczno-politycznych*, w: *Technokultura: transhumanizm i sztuka cyfrowa*, red. D. Gałuszka i in., Wydawnictwo Libron, Kraków 2016, s. 47-62.

⁹ Badania na temat neurologicznych podstaw zachowań związanych z konsumpcją informacji, zarówno użytecznej, jak i bezużytecznej, wskazują, że informacja sama w sobie „postrzegana” jest przez mózg jako nagroda. Tłumaczy to zachowania związane z kompulsywnym korzystaniem

i konsumpcji symbolicznej, wirusowe rozsiewanie treści za pomocą sieci społecznościowych i narzędzia dające poczucie aktywności oraz relacji, sprzyjają transformacji umysłu człowieka niekoniecznie (lub nie zawsze) w kierunku jego doskonalenia¹⁰.

Przemiana dokonuje się także na poziomie konceptualizacji ciała¹¹. Liczne przeróbki, tak zwane „remake’i” ciała można postrzegać jako pierwszy stopień akceptacji procesu cyborgizacji¹². Ciało jest traktowane jako przedmiot

z mediów cyfrowych, w tym infoholizm (por. K. Kobayashi, M. Hsu, *Common Neural Code for Reward and Information Value*, „PNAS” 116(2019) nr 26, s. 13061-13066, <https://www.pnas.org/doi/full/10.1073/pnas.1820145116>).

¹⁰ Niewątpliwie, towarzyszące korzystaniu z mediów i z usieciowionej informacji, poczucie omnipotencji buduje jednak we współczesnym człowieku wrażenie dostępności wiedzy, która nie została zinternalizowana, a w konsekwencji – ograniczenie procesu uczenia się. Wiedza dostępna potencjalnie dzięki asyentom AI nie jest jednak wiedzą przetworzoną i poddaną refleksji w umyśle człowieka, jest to wiedza technicznie zapośredniczona, zdalna i peryferyczna.

¹¹ Popkulturowa wizja ciała to ciało wytatuowane, zmodyfikowane, urozmaicone – poprzez różnego rodzaju ozdoby, obrazy, implanty. Trend ten zapoczątkowały nie tylko ruchy kontrkulturowe, ale też sztuka nowych mediów, między innymi dzieła Stelarc’a (robotyczne nogi i ręce, egzoskielety, wreszcie implant ucha na przedramieniu). Artysta przeszedł drogę od maltretowania własnego ciała w trakcie performansów z podwieszaniem i nakłuwaniem go, przez bolesne eksperymenty robotyczne, podpinające jego ciało do robotycznych protez i sieci (wraz z oddaniem nad nim pełnej kontroli publiczności), po radykalny eksperyment, w koncepcji trwale modyfikujący jego ciało poprzez wszczepienie implantu ucha na rękę, czyli sztucznego organu komunikacyjnego, który miał pozwalać nasłuchiwać artyście treści z internetu. Ostatnie dzieło można uznać za nie do końca udane, gdyż implant nie zadziałał i trzeba było go usunąć – zbuntowało się bowiem ciało Artysty. Dokumentacja prac por. *Suspensions, Third Hand, Exoskeleton, Muscle Machine, Ear on Arm*, Stelarc, <http://stelarc.org>. Działalność takich artystów należy postrzegać jako świadomy dialog człowieka z technologią, stymulujący przesuwanie granic „wyobraźmalnego”. Eksperymenty te spotykały się zawsze z dużym oddźwiękiem zarówno wśród krytyków sztuki mediów, jak i szerokiej publiczności. Popkultura przetwarza takie bodźce docierające ze świata sztuki, odzierając je z sensów i znaczeń dyskursotwórczych. Pozostają znaki bez oznaczników, puste ornamenty. Zamiast poważnego dyskursu na temat przyszłego rozwoju człowieka, w świadomości społecznej pozostaje obraz ucha na rękę i towarzyszące mu zadziwienie oraz akceptacja biohybrydyzacji. Ciało naturalne (w znaczeniu niezmodyfikowane), podobnie jak niezmediatyzowany umysł, przestaje mieć rację bytu – a w każdym razie części społeczeństwa wydaje się nudne. Kultura pielęgnuje dziś model „DIY: do it yourself”. Kreatywność przenoszona jest na otoczenie, a tym samym również ciało staje się jej nośnikiem i łatwo dostępnym medium.

¹² Kolejnym jest afirmacja biohybrydyzacji, w której wyraźną pomocą staje się chirurgia plastyczna, niestroniąca od kreatywnych realizacji, pozwalających na „transgatunkowe” czy „intergalaktyczne” przeróbki twarzy i ciała klientów (na przykład zamianę w wybranego bohatera komiksu czy ulubionej serii science-fiction). W tym kontekście interesujące wydają się uwagi Anny Irsak na temat przemiany etyki w obrębie współczesnej medycyny, a może nawet wizji całej dyscypliny i przestrzegania jej zasadniczych celów (koncepcja kontraktualistyczna). Konstatacje te w znacznej mierze dotyczą także istoty i zasadniczych celów medycyny estetycznej, co uwiadacza się zwłaszcza w obrębie jej „artystycznych”, nie interwencyjnych (to jest niemających na celu pomocy medycznej) zastosowań podszytych pogonią za sławą i oryginalnością, a często ideologią transhumanistyczną (por. A. Irsak, *W obronie (tradycyjnej) medycyny*, „Ethos” 28(2015) nr 3(111), s. 176-188).

i środek przekazu¹³. Zakorzenie umysłu w technologii jest jednak znacznie bardziej powszechne i uzależniające, niesie też poważniejsze konsekwencje na przyszłość dla człowieka i kultury. Smartfony, wyposażone w agenta konwersacyjnego AI i potencjalnie nieskończoną rozszerzalną funkcjonalność (za pośrednictwem aplikacji i sklepów z aplikacjami), stają się prawdziwymi czynnikami zmiany kulturowej.

Istnieją już dowody na to, że małe dzieci uczące się mówić, posiadające w domu konwersacyjną sztuczną inteligencję, komunikują się inaczej, tworząc nowy rodzaj języka¹⁴. Dzieci adaptują się nieświadomie do sprawniej komunikacji z AI, dostosowując struktury językowe do wydawania poleceń „maszynie” i uzyskiwania od niej odpowiedzi. Struktury gramatyczne w takiej konwersacji różnią się od tradycyjnej (wyłącznie ludzkiej) socjalizacji, co więcej, zostają przeniesione na komunikację z rodziną. Agent konwersacyjny nie staje się może kolejnym domownikiem, ale z pewnością zmienia kontekst komunikacyjny, a także powoduje modyfikację pewnych zachowań komunikacyjnych. Obecność sztucznej inteligencji w środowisku domowym, takiej jak choćby Alexa (Amazon Echo), nie tylko zmienia nastrój wnętrza (na przykład włączając muzykę), ale też – mniej dostrzegalnie – redefiniuje relacje międzyludzkie w rodzinie¹⁵ i oswoja sztuczną inteligencję jako „naturalny” (bo wszechobecny i „udomowiony”) element bliskiego środowiska (domu). Może też – i od pewnego czasu już to czyni – rozwiązywać problemy najwyższej wagi: ratować domownikom życie, wzywając pogotowie ratunkowe¹⁶. Konsekwencją takich zależności komunikacyjnych i egzystencjalnych jest wzrost zaufania wobec nowej technologii i delegowanie na nią ludzkiej odpowiedzialności. Skoro AI ratuje życie i doradza w kwestiach zdrowotnych, rozwiązuje zadania domowe, zastępując w procesie edukacji dzieci pomoc rodzica¹⁷, dostarcza

¹³ Medycyna plastyczna w takich realizacjach wydaje się spełniać jedynie funkcję pozbawionego refleksji etycznej narzędzia rynkowego. Można rozumieć wspomniane modyfikacje ciała jako znak pojawienia się nowych tożsamości, jak chcą niektórzy badacze, zdaniem innych obrazują one kryzys duchowości i szacunku do ciała (zob. „Człowiek a/i/lub maszyna: Duchowość 3.0”, seminarium, Wydział Teologiczny Uniwersytetu Śląskiego w Katowicach, 13 kwietnia 2018; „Filo-Sofija” 17(2017) nr 39(1)).

¹⁴ Por. A. Hiniker i in., *Can Conversational Agents Change the Way Children Talk to People?*, „Interaction Design and Children”, ACM, June 24, 2021, s. 338-349.

¹⁵ Por. E. Beneteau i in., *Parenting with Alexa: Exploring the Introduction of Smart Speakers on Family Dynamics*, w: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, ACM, New York, April 2020, s. 1-13.

¹⁶ Por. Z. Kleinman, *Apple's Siri Calls Ambulance for Baby*, „BBC News”, 07.06.2016, <https://www.bbc.com/news/technology-36471180>.

¹⁷ Zob. A. Agraval i in., *Prediction Machines. The Simple Economics of Artificial Intelligence*, Harvard Business School Publishing, Boston 2018.

wszelkich informacji – nie tylko staje się ekspertem, który pomaga decydować, ale też człowiek zaczyna odczuwać wobec niego wdzięczność i przestaje ufać własnym kompetencjom.

To zapośredniczenie następuje na wielu poziomach komunikacji, w pewnym stopniu niezauważalnie. W korzystaniu z asystenta AI czy opartej na AI aplikacji nie ma nic podejrzanego, oferują one komodyfikację i racjonalizację naszych działań, zwiększając efektywność podejmowanych czynności i dając poczucie optymalizacji zachowań i wyborów. *Sztuczna inteligencja, oferując automatyzację zachowań i wyborów, zastępuje coraz częściej człowieka w podejmowaniu drobnych, codziennych decyzji.* Podpowiada, jakie ubrania w naszej szafie lepiej do siebie pasują, co powinniśmy dziś zjeść i co kupić, jak dojechać najsprawniej do danej lokalizacji, z kim powinniśmy się spotykać, jakie wspomnienia celebrować, a także kiedy iść spać, by wyspać się najlepiej¹⁸. Algorytmizacja codzienności oraz zbieranie wzorców mikrobehawioralnych następują niepostrzeżenie, przekraczając kolejne granice prywatności.

Transhumaniści mówią o potrzebie integracji nośników cyfrowych danych z ciałem (za pośrednictwem chipów komunikacyjnych lub implantów domózgowych i nanotechnologii), lecz nie wydaje się to konieczne. Współczesny etap „kultury kciuka” i „szyi smartfonowej” oznacza wystarczającą integrację z technologią, optymalną z perspektywy przystosowania ewolucyjnego ciała i aparatu percepcyjnego. Używając inteligentnego smartfona, homo sapiens korzysta bowiem, jak dotąd, to jest w całym procesie ewolucji, z dostępnych w środowisku narzędzi (używał ich już przecież homo habilis, czyli człowiek (już) zręczny, choć (jeszcze) nie mądry, to jest nie sapiens). Współczesne narzędzia komunikacyjne dają poczucie kontaktu, dostęp do informacji, a jednocześnie pozwalają delegować umysł i dostosować aparat percepcyjny do nowych wyzwań usieciowionej technocodzienności.

Ten etap ewolucyjny zostanie przekroczony, gdy tylko ciało człowieka dostosuje się trwale do narzędzi cyfrowych, a cyborgizacja będzie codziennością. Stanie się to wraz z przekroczeniem granic oporu ciała i kultury. To jedna z możliwych perspektyw spojrzenia na przełomowy punkt nadejścia „osobliwości” (ang. singularity), projektowany przez Raya Kurzweila, wynalazcę, futurologa i promotora transhumanizmu¹⁹. Transgresja jest przecież wpisana w transhumanistyczny proces postępu technologii i ewolucji człowieka. Wy-

¹⁸ Przykłady aplikacji realizujących powyższe zadania dzięki AI: Acloset – AI Fashion Assistant, FoodVision: Ai Food Tracker, Google Maps, FindBee: Find My Family&Friends, iPhoto/iCloud, Smart Self-Control to Focus, Smart Self-Control to Wake Up et cetera (Apple Store).

¹⁹ Por. R. Kurzweil, *The Singularity is Near: When Humans Transcend Biology*, Duckworth, London 2009, s. 260. Zob. t e n ż e, *Nadchodzi osobliwość. Kiedy człowiek przekroczy granice biologii*, tłum. E. Chodkowska, A. Nowosielska, Kurhaus Publishing, Warszawa 2016.

czekiwany moment osobliwości to narodziny superinteligencji (i tym samym kres człowieka biologicznego)²⁰.

CYBORGIZACJA CZŁOWIEKA I HUMANIZACJA MASZYNY

Ted Nelson, genialny twórca koncepcji hipertekstu i filozof epoki cyfrowej, już w latach siedemdziesiątych przestrzegał w swym wpływowym manifestie *Computer Lib/Dream Machines* przed postrzeganiem komputerów i AI jako narzędzi myślących²¹. Badacz akcentował demiurgiczną rolę człowieka wobec technologii komputerowych. Transhumaniści dążą jednak do zmiany relacji między człowiekiem i technologią, widząc w cyborgizacji człowieka i humanizacji maszyny szansę. Zamiast patrzeć na transhumanizm jedynie z perspektywy filozoficznej, dostrzegając w nim propozycję na czasy kryzysu humanizmu²², można też postrzegać go z perspektywy kulturowej i ekonomicznej. To nowy rodzaj darwinizmu społecznego – darwinizm technologiczny, planowany dla przyszłych pokoleń przez współczesne. Dostęp do nowych technologii (biotechnologii, nanotechnologii, protetyki, augmentacji) będzie miał decydujący wpływ na pozycję jednostki w społeczeństwie: poprzez ciało i umysł augmentowane według woli i możliwości finansowych rodziców przyszłych dzieci²³. Sztuczna inteligencja będzie pełniła tu rolę podstawową i definiującą dla postspołeczeństw przyszłości, zarządzając nimi i optymalizując wszelkie wysiłki jednostek.

Współczesne podejście protechnologiczne, jednoczące wyraźnie świat nauki, polityki i globalnej gospodarki, w którym sztuczna inteligencja oraz roboty stanowią nadzieję ludzkości, zakorzenione jest w myśli transhumanistycznej. Marvin Lee Minsky, jeden z wiodących myślicieli transhumanizmu, pionier

²⁰ „Superinteligencja” rozumiana jest jako sztuczna inteligencja przekraczająca poziom ludzkiej inteligencji, co ma uniemożliwić konkurowanie człowieka z maszynami (por. I. J. G o o d, *Speculations Concerning the First Ultrainelligent Machine*, „Advances in Computers” 6(1965), s. 31-88).

²¹ Por. T. N e l s o n, *Computer Lib/Dream Machines*, <http://worrydream.com/refs/Nelson-ComputerLibDreamMachines1975.pdf>.

²² W zależności od ujęcia badaczy oznacza to: nowy humanizm (Carry Wolfe) lub antyhumanizm (posthumanizm), czyli albo odrodzenie myśli humanistycznej, lecz w nowym kontekście technologicznym (stawiającej człowieka i jego rozwój w centrum zainteresowania), albo zaprzeczenie wszystkim wartościom humanizmu (konstatacja końca epoki człowieka, akcentowanie destrukcyjnego dla planety i innych gatunków wpływu człowieka, zgoda na przejęcie świata przez nową formę, czyli maszyny). Por. M. G a r b o w s k i, *Transhumanizm: Geneza – założenia – krytyka*, „Ethos” 28(2015) nr 3(11), s. 34; E. B i ũ c z y k, *Inżynieria klimatu a inżynieria człowieka. Dyskursy na temat środowiska w epoce antropocenu*, „Ethos” 28(2015) nr 3(11), s. 156.

²³ Zob. F. F u k u y a m a, *Transhumanism. A Special Report*, „Foreign Policy”, 23.10.2009, <https://foreignpolicy.com/2009/10/23/transhumanism>.

sztucznej inteligencji i robotyki, był jednym z patronów instytucjonalizacji badań nad sztuczną inteligencją²⁴. To dzięki niemu oraz Johnowi McCarthy'emu powołano instytuty naukowe zajmujące się AI w Stanford i w MIT. Równolegle podobne badania rozpoczęły korporacje IBM, Rand czy Bell Labs²⁵. To w tych laboratoriach wypracowano nowe koncepcje inteligencji i ich przełomowe aplikacje w obszarze łączącym informatykę z robotyką oraz teorią poznania i percepcji. Pojęcie „inteligencji” nieodwracalnie zostało zredefiniowane, przestało być kategorią opisową zarezerwowaną dla działań człowieka. Wcześniejszego rozszerzenia zakresu pojęcia „inteligencji” na świat zwierząt dokonali cybernetycy wspólnie z etologami i badaczami systemów biologicznych. Dzisiejsze dyscypliny badań nad sztuczną inteligencją, robotyka, protetyka, informatyka i wiele innych korzystają z dokonań obydwu nurtów. Transhumanistyczne założenia o logice nieustannego rozwoju systemów oraz o prawie do pozabiologicznego przyspieszenia rozwoju człowieka upowszechniają się wraz z ich sukcesami²⁶.

Współcześnie powszechnie i bezdyskusyjnie zakłada się, że wszystkie kryzysy zniweluje jedynie sztuczna inteligencja, zgodnie z rozpropagowaną japońską koncepcją Społeczeństwa 5.0, według której wszelkie problemy społeczne mają być rozwiązywane przez nowe technologie²⁷. Oddelegowa-

²⁴ Na przykład Marvin Minsky jest twórcą pierwszej sieci neuronowej (SNARC), wynalazcą jednego z pierwszych skanerów wizualnych, pierwszych robotów – „żółwia” LOGO, robotycznej dłoni z czujnikami dotyku oraz wielu interfejsów sprzętowych i oprogramowania, uhonorowany został między innymi nagrodą Turinga. Zob. Brief Academic Biography of Marvin Minsky, MIT, <https://web.media.mit.edu/~minsky/minskybiog.html>.

²⁵ Por. H. R h e i n g o l d, *Narzędzia ułatwiające myślenie. Historia i przyszłość metod poszerzania możliwości umysłu*, tłum. J. Szporko, Wydawnictwa Naukowo-Techniczne, Warszawa 2003, s. 150. Niebagatelny wpływ na tę dyscyplinę mieli Claude Shannon, Herbert A. Simon czy Allen Newell. Istotną rolę w konceptualizacji obszaru badawczego odegrały zwłaszcza prekursorskie prace Alana Turinga oraz Shannona, zwracające uwagę na inne aspekty niż biologicznie zorientowana cybernetyka Norberta Wienera (por. A. M a j, *Przemiany wiedzy w cyberkulturze. Badania nad kulturą, komunikacją, wiedzą i mediami*, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2021, s. 335).

²⁶ Pewną rolę w popularyzacji koncepcji sztucznej inteligencji w świadomości społecznej odegrały również powiązania Minsky'ego z Hollywood. Minsky oraz Irving J. Good byli doradcami Stanleya Kubricka w realizacji filmu *2001: Odyseja kosmiczna* (USA, 1968), który wpłynął na kolejne pokolenia twórców. Minsky miał znaczący wpływ na filmową wizję komputera HAL 9000, znacznie różniącą się od wersji z noweli Arthura C. Clarka (zob. D.G. S t o r k *Scientist on the Set: An Interview with Marvin Minsky*, w: *HAL's Legacy: 2001's Computer as Dream and Reality*, red. D.G. Stork, wstęp A.C. Clarke, MIT Press, Cambridge, London 1998). Kolejne sukcesy nowych technologii przyemiły jednak przez lata dystopijne wizje Kubricka i innych artystów, a technoop-tymistyczna polemika inżynierów i informatyków dominuje dziś nad bardziej krytycznym wobec technologii dyskursem humanistycznym.

²⁷ Por. J. G r u d o w s k a, D. Z i e l i Ń s k i, *Społeczeństwo 5.0: Refleksja krytyczna*, „Transformacje” 2(2022), s. 234. Wśród kryzysów rozwiązywalnych przez AI wymienia się w literaturze na przykład katastrofę klimatyczną, pandemię, epidemie, trudne operacje medyczne, problematyczne

nie odpowiedzialności i przekazanie jej systemom ekspertowym już jest codziennością w coraz większej ilości obszarów życia społecznego. Systemy opierające się na analizie big data stają się nieodzowne, zmieniając logikę współczesnego działania człowieka i pozbawiając go ciężaru decyzyjnego, który dotąd definiował kondycję ludzką. Maszyny i systemy informatyczne, a zatem narzędzia nowej generacji, stały się ośrodkiem zaufania jako reprezentanci porządku algorytmicznego, któremu od oświecenia powierzamy życie. Od koncepcji mózgu elektronowego, popularnej w połowie dwudziestego wieku, przeszliśmy daleką drogę do autonomicznych i samouczących się sieci neuronowych, których przetwarzania informacji nie jesteśmy w stanie w pełni zrozumieć.

Na ten kierunek rozwoju społeczeństwa informacyjnego i stopniową algorytmizację kultury zwracali od lat uwagę badacze z różnych dyscyplin²⁸. Dziś dyskusja filozoficzna i humanistyczna podążać musi za nowymi przełomami w naukach informatycznych, biotechnologii i robotyce, tym samym jest ona sprowadzana do podwójnej roli – przestrzegania i komentowania.

Podwójną rolę początkowo pełnili także refleksyjni eksperymentatorzy i wynalazcy, tacy jak Kevin Warwick, profesor cybernetyki na University of Reading, a następnie University of Coventry w Wielkiej Brytanii, „pierwszy cyborg”, czy Steve Mann, profesor inżynierii elektrycznej i komputerowej z University of Toronto, „ojciec technologii ubieralnych”. Naukowcy i artyści zarazem, którzy u progu dwudziestego pierwszego wieku podejmowali pierwsze próby cyborgiczne na sobie samych, stawiali intrygujące pytania o dalsze możliwości współegzystencji człowieka i maszyny oraz o jej zasady. Ryzykując własnym zdrowiem, przesuwali granice poznania. Mann i Warwick niezależnie pracowali nad możliwością wykorzystania nowych technologii w protetyce i w obszarze assistive technologies, czyli technologii wspierających osoby niepełnosprawne. Ich dokonania wpłynęły na współczesną medycynę. Mann, eksperymentując w obszarze substitucji i technicznego rozszerzania zmysłu wzroku, rozwinął koncepcję technologii ubieralnych (ang. wearables). Warwick poddał się ope-

diagnozy, zarządzanie przestrzenią miejską czy infrastrukturą krytyczną, sprawiedliwy osąd więźniów lub rozpoznanie podejrzanych, wytypowanie pracowników do zwolnienia.

²⁸ Por. T. M i c z k a, *Specifics of Algorithmization in Data Culture*, w: *ICSRS 2019: Conference Proceedings, International Conference on Recent Social Studies and Research*, 25-26 October 2019, CORIS Department Sapienza University, Rome 2019, s. 240-247; K. K r z y s z t o f e k, *Człowiek – społeczeństwo – technologie. Między humanizmem a transhumanizmem i posthumanizmem*, „Ethos” 28(2015) nr 3(111), s. 191-213; L.W. Z a c h e r, *Od społeczeństwa informacyjnego do społeczeństwa wiedzy (dylematy tranzycyjne: między informacją, wiedzą i wyobraźnią)*, w: *Społeczeństwo informacyjne. Wizja czy rzeczywistość?*, t. 1, red. L.H. Haber, Akademia Górniczo-Hutnicza, Kraków 2004, s. 103-112; T. B i e l a k, G. P t a s z e k, *Algorytmiczne doświadczanie kultury*, „Kultura Współczesna” 2019, nr 1(104), s. 10-14.

racjom wszczęcia kolejno dwóch chipów w swoje ciało, czyli dobrowolnej cyborgizacji, i przeprowadził prekursorskie badania w zakresie transferu sygnału na odległość, zdalnej kontroli robotów i systemów telekomunikacyjnych oraz zachowań ludzkiego ciała i umysłu podłączonego do komputera²⁹. Badacz od dawna przekonuje, że jego doświadczenia i sukcesy komunikacyjne nowych technologii skazują tych, którzy nie będą chcieli się zmodyfikować poprzez biochipowanie i protezowanie ciała na ewolucyjną klęskę. Porównuje komunikację cyborgów i ludzi ze współczesną komunikacją (odpowiednio) ludzi i krów³⁰.

Być może perspektywa szybszego rozwoju bardziej sprawnych procesów myślowych zwycięży, jednak na razie obserwujemy raczej drugą stronę tego procesu: doskonalenie się sztucznej inteligencji, ucłowieczanie robotów i rozwój systemów ekspertowych zarządzających biologicznie nierozszerzonymi ludźmi i społeczeństwami. Warwick dopuszcza jednak ewolucję podejścia społeczeństw do kwestii upgrade'u, sam bowiem przeszedł od lęku do fascynacji możliwościami cyborga³¹.

Idea cyborgizacji ciała i sztucznej inteligencji spotykają się w usieciowieniu i dostępie do wszechobecnych danych (ang. ubicomp). Cyborg jawi się jako sprawniejszy mentalnie, a zarazem lepiej przetwarzający dane w czasie rzeczywistym. Dziś systemy ekspertowe, agenty konwersacyjne i różne systemy oparte na sztucznej inteligencji karmione są wielkimi zbiorami danych i trenowane na określonych modelach po to, by same się uczyły. Nawet metaforyka kontaktu ze sztuczną inteligencją przypomina model rodzicielskiego kontaktu z małym dzieckiem. Daje to poczucie kontroli, ale jednocześnie mówi o kreatywnym aspekcie tego procesu. Kreatywność kojarzona bywa zazwyczaj pozytywnie, dzięki czemu taki dyskurs może mieć charakter techniczny (skupia się na działaniach), zarazem nie wywołując lęku. Jednak po sukcesach sztucznej inteligencji w grze w go, gdy Google DeepMind pokonał Lee Sedolę³², oraz po raporcie o chatbotach wymyślających własny język, by samodzielnie negocjować, można mówić o nowych obszarach budzących potencjalnie lęk, ale też refleksję nad zachodzącymi procesami humanizacji technologii, robotów i AI. Dyskusja nad sukcesami „twórczego” i „rozumnego” GPT-3.5 oraz „wrażliwego” i „podmiotowego” LaMDA jest kolejną odsłoną tego procesu: GPT-3.5 przechodzi test Turinga, zdobywając uznanie naukowców i dzienni-

²⁹ Zob. K. Warwick, *I, Cyborg*, University of Illinois Press, Urbana–Chicago 2004.

³⁰ Por. Infomia, *Cyborg Life: Kevin Warwick*, YouTube, 14.04.2008, http://www.youtube.com/watch?v=RB_17SY_ngl.

³¹ Por. Warwick, dz. cyt., s. 295.

³² Por. J. Bohannon i in., *Artificial Intelligence Beats Go Champ*, w: *From AI to Protein Folding: Our Breakthrough Runners-up*, „Science”, 22.12.2016.

karzy, a LaMDA przekonuje nadzorującego go inżyniera z Google, że posiada świadomość i należy uznać go za osobę³³.

Nelson, w przywołanym wcześniej założycielskim manifestie na temat relacji technologii komputerowych i człowieka, zdecydowanie r o z r ó ż n i a ł i n t e l i g e n c j ę c z ł o w i e k a i i n t e l i g e n c j ę m a s z y n y, podkreślał rolę intencji, wolnej woli i kreatywności³⁴. Dziś wszystkie te obszary, włącznie z empatią, emocjami, inteligencją, zdolnościami komunikacyjnymi i wolną twórczością, stanowią obszar badań i eksperymentów rozwijających zarówno roboty, jak i sztuczną inteligencję. Te – dotąd postrzegane jako t y p o w o l u d z k i e – c e c h y s t a ł y s i ę o b s z a r e m p o d l e g a j ą c y m p r o j e k t o w a n i u (a zatem są badane, modelowane, programowane i prototypowane). Ta zmiana wiąże się z e w o l u c j ą m y ś l e n i a o s a m e j i n t e l i g e n c j i i z e z n a c z n ą z m i a n ą r o z u m i e n i a j e j i s t o t y w s t o s u n k u d o p i e r w s z y c h p i s m c y b e r n e t y k i i k o n c e p c j i w i d z ą c y c h i n t e l i g e n c j ę j a k o s y m u l a c j ę p r o c e s ó w l o g i c z n y c h, n a p r z y k ł a d g r y w s z a c h y.

Zmiana myślenia widoczna jest w obszarze projektowania robotów-androidów. Japoński konstruktor-robotyk Hiroshi Ishiguro rozwija koncepcję „senzai-kan”, charakterystyczną dla ludzi świadomości „bycia-tu”, poczucia istnienia³⁵. Roboty Ishiguro, między innymi Geminoid (wierna kopia twórcy), budzą zarazem i fascynację, i lęk³⁶. Mimo wysokiego zaawansowania programistycznego i technologicznego, nie pokonały one wciąż „doliny osobliwości” (ang. uncanny valley), zdefiniowanego przez Masahiro Mori³⁷ problemu robotyki: bliskiego podobieństwa i jednocześnie dziwnej, odczuwanej jako nieprzyjemna, obcości robota (przypominającego zombie).

Problem stanowi nie tylko dostępna technologia, ale przede wszystkim ludzka percepcja i oczekiwania poznawcze. Zwiększanie perfekcji odwzorowania ludzkiego ciała i doskonalenie sztucznej inteligencji nie przyniosły dotąd rozwiązania. Paradoksalnie, roboty wyposażone w sztuczną inteligencję, które w k o m u n i k a c j i z c z ł o w i e k i e m p r z e j a w i a j ą a u t o i r o n i ę i d z i a ł a n i a i r r a c j o n a l n e, postrzegane są przez ludzi jako bardziej ludzkie i budzą pozytywniejsze odczucia³⁸. Sugerując, że

³³ Por. B. L e m o i n e, *Is LaMDA Sentient? – an Interview*, „Medium”, 11.06.2022, <https://cajundiscordian.medium.com/is-lambda-sentient-an-interview-ea64d916d917>; M i n t z, dz. cyt.

³⁴ Por. N e l s o n, dz. cyt.

³⁵ Mówi o tym Ishiguro w wywiadzie w filmie dokumentalnym *Mechanical Love* (Finlandia, 2007, reż P. Ambo).

³⁶ Zob. H. I s h i g u r o, *Symbiotic Human-Robot Interaction Project*, <https://www.jst.go.jp/erato/ishiguro/en/index.html>.

³⁷ Zob. M. M o r i, *The Uncanny Valley* (1970), „IEEE Spectrum”, 12.06.2012, <https://spectrum.ieee.org/the-uncanny-valley>.

³⁸ Zob. M. L. T r a e g e r i n., *Vulnerable Robots Positively Shape Human Conversational Dynamics in a Human–Robot Team*, „PNAS” 117(2020) nr 12, s. 6370–6375.

także (jak ludzie) popełniają błędy, realizują pewien algorytm, dzięki któremu w ludzkim odbiorze stają się bardziej do nas podobne. Nie są nielogiczne, a jednak mogą budować takie wrażenia w interaktorach. Taka symulowana (udawana) niekonsekwencja i niedoskonałość byłyby zatem tymi elementami, które sprawiają, że technologia przestaje budzić lęk i wydaje się bliższa innym doświadczeniom komunikacyjnym (naturalnie zakorzenionym w interakcjach z ludźmi, pełnymi niekonsekwencji). Jest to początek *p r o g r a m o w a n i a m a n i p u l a c j i l u d z k ą p e r c e p c j ą* w kontakcie z AI.

W powiązaniu z obserwacjami różnych zespołów badawczych w zakresie *k r e a t y w n o ś c i k o m u n i k a c y j n e j s z t u c z n e j i n t e l i g e n c j i* ukazuje to interesujący nowy aspekt „oswajania” AI. Boty wytwarzają już własne języki o charakterze tekstowym lub matematyczno-wizualnym³⁹. Ludziom wydają się one nielogiczne i niekomunikatywne, lecz prowadzą do udanych negocjacji między botami, a zatem spełniają zadanie środka komunikacji⁴⁰. Historia rozwoju dziedziny sztucznej inteligencji pełna jest zaskakujących zwrotów, jak zauważają twórcy Open AI. Trudno zatem przewidzieć, w którym kierunku będzie zmierzać rozwój tych technologii⁴¹.

Obecnie zaczynamy już rozważać sztukę robotów, sztukę AI, samoreplikację robotów, samodoskonalenie AI, uczenie się, emocje robotów i AI, świadomość, niedługo zacniemy zapewne mówić o ich duchowości. Proces ten można uznać nie tylko za zmianę w obrębie rozumienia samej inteligencji, ale też za *h u m a n i z a c j ę r o b o t ó w i s z t u c z n e j i n t e l i g e n c j i*. Postępuje ona wraz z powszechną akceptacją transhumanistycznej negacji wyjątkowości człowieka i transcendencji, biologizacją kultury i aktywności duchowej, umysłowej, a także aksjologii i emocjonalności człowieka w dyskursie naukowym, wraz z popularyzacją neoewolucjonizmu i memetyki⁴²,

³⁹ Zob. M. L e w i s i i n., *Deal or No Deal? End-to-End Learning for Negotiation Dialogues*, <https://s3.amazonaws.com/end-to-end-negotiator/end-to-end-negotiator.pdf>.

⁴⁰ Zob. tamże.

⁴¹ Zob. G. B r o c k m a n, I. S u t s k e v e r, O p e n A I, *Introducing Open AI*, <https://openai.com/blog/introducing-openai/>.

⁴² Neoewolucjonizm jest prądem na nowo analizującym koncepcję ewolucji, stawiającym tezę, że koncepcja K. Darwina nie jest błędna, choć bywała błędnie interpretowana. Memetyka z kolei dokonuje powiązania pomiędzy ewolucją genetyczną a transmisją kulturową i rozwojem cywilizacji oraz techniki, poszukując najmniejszej cząstki informacji, która jest przenoszona różnymi drogami we wszystkich tych procesach (por. S.J. G o u l d, *Niewczesny pogrzeb Darwina*, tłum. N. Kancewicz-Hoffman, Państwowy Instytut Wydawniczy, Warszawa 1991 [neoewolucjonistyczne podejście do ewolucji i rozwoju człowieka w świetle badań antropologicznych, paleontologicznych i archeologicznych końca XX wieku]; R. D a w k i n s, *Samolubny gen*, tłum. M. Skoneczny, Prószyński i S-ka, Warszawa 1996 [memetyczna koncepcja rozwoju organizmów na Ziemi na podstawie ewolucji genetycznej i kulturowej (memetycznej), propozycja terminu „mem”]; S. B l a c k m o r e, *Maszyna memowa*, tłum. N. Radomski, Rebis, Poznań 2002 [rozszerzenie koncepcji memetycznej o ewolucję technologiczną, propozycja wyróżnienia „temów”].

a w medycynie postrzeganiem człowieka jako bytu hormonalnego, czyli biologicznego, lecz algorytmicznego, zależnego od procesów chemicznych. Jest to kilka prądów myślowych, wśród których wyróżnia się zwrot biologiczny⁴³, transhumanizm oraz algorytmizacja poznania⁴⁴. Ich wspólnym mianownikiem jest mechanistyczne spojrzenie na człowieka oraz antyhumanizm. Powrót do biologiczności i odwrót od duchowości (a także kultury, filozofii, psychologii) można postrzegać jako *d e h u m a n i z a c j ę c z ł o w i e k a*, konieczną do „oswojenia” AI i powstania *humanity*⁺.

EWOLUCJA INTELIGENCJI: AI TO NIE KONIEC

Epoka „eksplozji inteligencji”, przewidywana w latach sześćdziesiątych przez Irvinga J. Gooda, matematyka i kryptologa, współpracownika Alana Turinga w Bletchley Park, zajmującego się teorią prawdopodobieństwa i statystyką Bayesowską, niewątpliwie albo już nadeszła, albo jest blisko. Należy oczekiwać zatem pojawienia się *s u p e r i n t e l i g e n c j i*. Zdaniem Gooda nie powinniśmy nigdy dopuścić do tego momentu, gdyż będzie to koniec inteligencji ludzkiej⁴⁵. Do tej dystopijnej myśli wracano wielokrotnie w momentach przełomu technologicznego. Uczynił to między innymi Vernor Vinge u progu lat dziewięćdziesiątych i technologii globalnej, cywilnej sieci czy Nick Bostrom w 2014 roku w kontekście rozwoju AI. Bostrom zauważył, że pionierzy sztucznej inteligencji nie podejrzewali, że ich praca mogłaby powodować jakiegokolwiek zagrożenie dla ludzkości⁴⁶. Stan techniki poprzednich

⁴³ Chodzi tu o interdyscyplinarne badania, łączące różne dziedziny i tradycje. Badacze mówią tu o zwrocie biologicznym z uwagi na poszukiwania podstaw procesów i zjawisk społecznych, kulturowych czy technicznych w procesach i zjawiskach o charakterze biologicznym, takich jak ewolucja, ekosystem, prawa przyrody, systemy biologiczne. Wyróżnia się tu koncepcja konsilencji, czyli łączenia poszukiwań z różnych obszarów, które dają wspólne uogólnienia, zaproponowana przez W. Whewella w 1840 roku, a przypomniana na współczesny użytek łączenia nauk przez E.O. Wilsona w jego książce z 1998, będącej jednocześnie udaną realizacją badań w duchu konsilencji w zakresie socjobiologii (zob. E.O. W i l s o n: *Konsilencja. Jedność wiedzy*, tłum. J. Mikos, Wydawnictwo Zys i S-ka, Poznań 2002). Por. *Biological Turn. Idee biologii w humanistyce współczesnej*, red. D. Wężowicz-Ziółkowska, E. Wiczorkowska, Wydawnictwo Uniwersytetu Śląskiego, Katowice 2016.

⁴⁴ Nurt ten związany jest z problematyką big data i paradygmatem data-driven science, czyli nauki opartej na danych. Chodzi tu o przejście paradygmatów od nauki: 1) opartej na obserwacji i doświadczeniach empirycznych, 2) przez opartą na teorii i modelowaniu, 3) następnie na technikach obliczeniowych i symulowaniu procesów, aż po 4) opartą na analizie big data. Paradygmat ten obejmuje nie tylko nauki ścisłe, ale też przyrodnicze, społeczne, a nawet w coraz większym stopniu humanistyczne (zob. *The Fourth Paradigm: Data-intensive Scientific Discovery*, red. T. Hey, S. Tansley, K. Tolle, Microsoft Research, Redmont, Washington 2009).

⁴⁵ G o o d, dz. cyt., s. 31-88.

⁴⁶ B o s t r o m, dz. cyt., s. 4n.

dziesięcioleci nie pozwalał na takie rozważania poza dystopiami nurtu science-fiction. Poza tym rozwój następował stopniowo, dając stosunkowo wąskie możliwości do nadużycia nowej technologii. Musk, podobnie jak Bostrom czy Ajay Agrawal twierdzą, że obecnie sytuacja wygląda inaczej: zbliżamy się do wynalezienia ogólnej sztucznej inteligencji⁴⁷.

Ważną kwestią pozostaje jednak w y o b r a ż n i a n a u k o w c ó w r e p r e z e n t u j ą c y c h n a u k i ś c i ś l ę, skupionych na liczbach, algorytmach, rozwiązaniach kolejnych wąskich problemów, którzy nie patrzą często poza horyzont tego, co aktualnie możliwe⁴⁸. Naukowcy zajmujący się sztuczną inteligencją funkcjonują w ramach określonego paradygmatu nauki. Dzięki naukom informatycznym, obok wcześniej funkcjonujących paradygmatów nauk ścisłych (empirycznego, teoretycznego i komputacyjnego), wypracowano paradygmat datalogiczny (ang. data-driven, datalogic), oparty na współpracy, usieciowieniu i analizie wielkich zbiorów danych (ang. big data)⁴⁹. Od wczesnego eksperymentowania, tworzenia modeli i symulacji naukowcy w wielu dziedzinach powszechnie przeszli do „karmienia algorytmów” określonymi zbiorami danych, co przyspiesza rozwiązywanie problemów naukowych i stawianie prognoz opartych na danych (a więc wiarygodnych). Nastąpiło to w efekcie realnego przyspieszenia rozwoju oprogramowania opartego na sieciach neuronowych. Stopniowo przemiany te doprowadziły do algorytmizacji kultury i życia społecznego, nieodwracalnego procesu wykraczającego poza obszar nauki.

Skoro określa się współczesność jako „eksplozję inteligencji” i nadejście „epoki superinteligencji”, w a r t o d o o k r e ś l i ć r o z u m i e n i e i n t e l i g e n c j i. Nie jest to bez znaczenia, przez lata bowiem znaczenie terminu uległo ewolucji. Podejście na początku rozwoju teorii informacji zakładało jako inteligentne takie zachowanie, które wiąże się z wykorzystaniem wnioskowania logicznego⁵⁰. Już teksty Turinga i postawione w nich problemy pozwalają jednak na co najmniej podwójne podejście do proce-

⁴⁷ Istnieją dwa sposoby rozumienia sztucznej inteligencji, co powoduje pewne nieporozumienia. Najczęściej mamy do czynienia ze słabą SI (ang. narrow, weak, applied AI), gdy mowa o konkretnych rodzajach oprogramowania, ale to silna SI (ang. strong, general, full AI) jest przedmiotem obaw ludzkości. Słaba SI (wąska, stosowana) jest wyspecjalizowana w wąskim zakresie obliczeniowym – to na niej skupiali się przez lata badacze. Silna SI (ogólna, pełna) to inteligencja maszynowa rozumiana szeroko: samodzielna, rozwinięta, samoświadoma. To ona może być porównywana z myśleniem ludzkim, jest pewnym horyzontem poznawczym dla naukowców i twórców science-fiction (zob. Agrawal i in., *Prediction Machines*).

⁴⁸ Zob. G. Bell, T. Hey, A. Szalay, *Beyond the Data Deluge*. „Science” 323(2009) nr 5919, <https://science.sciencemag.org/content/323/5919/1297>.

⁴⁹ Por. przyp. 44.

⁵⁰ Por. *The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life: Plus The Secrets of Enigma*, red. B.J. Copeland, Oxford University Press, Clarendon Press, Oxford 2004, s. 431.

su myślenia: (1) w orientacji związanej z teorią silnej sztucznej inteligencji poprzez myślenie rozumie się procesy logiczne i operacje obliczeniowe, a celem jest powielenie działania umysłu (symulacja); (2) w orientacji związanej z teorią słabej sztucznej inteligencji poprzez myślenie rozumie się natomiast skomplikowany i wielowarstwowy proces, w którym biorą udział nie tylko logika, ale i percepcja (multisensoryczność), przetwarzanie bodźców i emocje, zatem celem modelowania jest tworzenie symulacji, które można traktować jako podstawę do weryfikowania pewnych hipotez na temat samej natury procesów myślowych, ale nie jako pełnoprawny model umysłu. W zależności od podejścia maszynom można przypisać zdolność myślenia (strong AI⁵¹) lub nie (weak AI). Turing reprezentował pierwsze podejście, ale rozważał też drugie (tu pojawiają się słynne „truskawki Turinga” w jego „grze imitacyjnej”)⁵¹. W wykładzie w London Mathematical Society w 1947 roku matematyk zauważył jednak, że „zachowanie inteligentne” to takie, w którym udział bierze też pewna nieprzewidywalność, aleatoryczność czy randomowość⁵². A zatem: nie tyle logika i rutyna, ile odejście od nich – świadczą o inteligentnym, czyli twórczym podejściu do danego problemu. Badacz zwrócił też uwagę na to, że stan obserwatora i jego odczucia (czyli interpretacja przewidywalności bądź nieprzewidywalności danego zachowania interlokutora) mają znaczenie dla oceny zachowania jako inteligentnego lub nie⁵³. Turing wyraził również przekonanie, iż z czasem maszyna „będzie się sama uczyć na bazie doświadczenia”⁵⁴ – można dostrzec tu wizję uczenia maszynowego.

Za tymi konceptualizacjami podążyli kontynuatorzy badań Turinga. Minsky zwrócił uwagę na wielotorowość myślenia ludzkiego: m y ś l e n i e abstrakcyjne, konceptualne, zadawanie pytań i uczenie się⁵⁵. Zajmował się także tym, jak uczą się dzieci, by móc ekstrapolować tę wiedzę w obszar inżynierii i robotyki. Wiele uwagi poświęcił s a m o ś w i a d o m o ś c i o r a z e m o c j o m jako ważnym elementom pozwalającym „myśleć po ludzku”. Jego definicja „inteligencji” zakłada „umiejętność rozwiązywania trudnych problemów” z łatwością i efektywnością⁵⁶. Inteligencją

⁵¹ Tamże, s. 463.

⁵² Tamże, s. 375.

⁵³ Por. tamże, s. 431.

⁵⁴ Tamże, s. 429.

⁵⁵ Por. M.L. M i n s k y, *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*, Simon & Schuster, New York – London – Toronto – Sydney 2006, s. 33, 281.

⁵⁶ Por. t e n ż e, *The Society of Mind*, Simon & Schuster, New York – London – Toronto – Sydney 1985, s. 71.

wyróżnia się zatem człowiek, ale nie posiadają jej zwierzęta czy superorganizmy. Może cechować się nią jednak maszyna. Człowieka natomiast badacz-transhumanista uważa za „maszynę emocjonalną”. Zwraca uwagę na problem z definiowaniem inteligencji, którą utożsamia z niezrozumiałymi procesami mentalnymi⁵⁷.

Dzięki pewnej wieloznaczności oraz zmianie rozumienia pojęcia „inteligencji” ten obszar badawczy pozostaje interesujący, co przyczynia się do dalszych poszukiwań i aplikacji teorii (nie tylko w obszarze AI, ale też w neurologii, neurokognitywistyce i psychologii). Wielu badaczy mówi zresztą o ewolucji inteligencji *per se* (a nie: rozumienia jej znaczeń) od dawna – zwłaszcza ewolucjoniści, neoewolucjoniści i memetycy. Odkrywamy dziś nowe wątki, mające istotne znaczenie dla rozumienia prehistorii człowieka, jego relacji z otoczeniem oraz rozwoju komunikacji (badania etologów, paleogenetyków, antropologów)⁵⁸. Nie chodzi tu jednak tylko o przejście od inteligencji biologicznej do sztucznej (a w momencie wyzwolenia spod kontroli człowieka do superinteligencji), a o zrozumienie pierwszej fazy rozwoju – ewolucji inteligencji biologicznej. Kolejnym stadium jest inteligencja sztuczna i rozważane w obrębie informatyki oraz nauk technicznych znaczenia, o czym była już mowa. Trzeci etap, przez niektórych badaczy uważany za przejściowy, przez innych za docelowy, to *inteligencja hybrydyczna* (rozszerzona). To w niej, czyli w połączeniu inteligencji biologicznej – ludzkiej i sztucznej – maszynowej, Edward Feigenbaum, twórca teorii systemów ekspertowych, widzi rozwiązanie niedostatków człowieka i maszyny⁵⁹. Inteligencja rozszerzona (ang. *augmented intelligence*) może stanowić najlepsze rozwiązanie, zapewniające bezpieczeństwo interesom człowieka, a jednocześnie efektywność większą niż inteligencja czysto biologiczna⁶⁰.

OSWAJANIE SZTUCZNEJ INTELIGENCJI – OSWAJANIE KOGO?

Badacze sytuujący się w obrębie podejścia konsiliencyjnego, zwrotu biologicznego, neoewolucjonizmu, memetyki oraz data-driven science wy-

⁵⁷ Tamże.

⁵⁸ Por. M. Tomasełło, *Historia naturalna ludzkiego myślenia*, tłum. B. Kucharzyk, R. Ociepa, Copernicus Center Press, Kraków 2015, s. 64.

⁵⁹ Zob. E.A. Feigenbaum, hasło „Expert Systems: Principles and Practice”, w: *The Encyclopedia of Computer Science and Engineering*, 1992, <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.34.9207&rep=rep1&type=pdf>; t e n ż e, *Knowledge Engineering: The Applied Side of Artificial Intelligence*, <https://stacks.stanford.edu/file/druid:qy055zd8682/qy055zd8682.pdf>.

⁶⁰ Zob. S. Gourelly, *Big Data and the Rise of Augmented Intelligence*, TEDxAuckland, YouTube, 5.12.2012, https://www.youtube.com/watch?v=mKZCa_cjbfq.

suwają interesujące hipotezy i dokonują nieoczekiwanych odkryć, które stawiają w nowym świetle ewolucję człowieka, zarówno tę dotychczasową, jak i przyszłą. Inspirujące dla zrozumienia procesu osvajania sztucznej inteligencji są między innymi badania prowadzone w Duke University przez antropologów Michaela Tomasella czy Briana Hare'go nad zachowaniami zwierząt, ewolucją człowieka, rozwojem komunikacji grupowej i koncepcją samoudomowienia⁶¹. Dla procesu osvajania sztucznej inteligencji koncepcja samoudomowienia niesie interesujące konsekwencje. Tłumaczyć może nie tylko psychologiczne mechanizmy związane z lękami wobec nowych technologii (reprezentującej w tym przypadku „inny” gatunek), procesy i dyskursy humanizacji maszyn i technicyzacji człowieka, ale i zaobserwowaną niedawno tendencyjność sztucznej inteligencji. Tę ostatnią tłumaczy się „biasami” danych, którymi została ona „nakarmiona”⁶², jednak w szerszym ujęciu jest to równocześnie reprezentacja konsekwencji naszej ewolucji w kierunku samoudomowienia. Dane pozyskiwane z sieci społecznościowych czy jakichkolwiek innych źródeł (w tym wypadku tekstów kultury) ukazują właśnie ślady wypracowanych w toku ewolucji zachowań obronnych, uprzedzających potencjalne zagrożenie wspólnoty przez „obcych”. W kontekście trenowania sztucznej inteligencji istotne jest uświadomienie

⁶¹ Zakłada ona, że w toku ewolucji człowiek nie tyle oswoił dzikie zwierzęta, ile raczej oswoiły się one same w wyniku specyficznej autoselekcji osobników. Zwierzęta podążające za produkującymi znaczne ilości odpadków grupami ludzkimi były lepiej odżywione i rozmnazały się lepiej niż osobniki zupełnie dzikie, bardziej lękliwe lub agresywne wobec człowieka. W rezultacie sukces ewolucyjny zapewniły sobie osobniki przyzwyczajające się stopniowo do obecności człowieka, czyli coraz bardziej udomowione. Rozwijały się one we własnym kręgu, nie łącząc się z bardziej lęklivymi i agresywnymi, wzmacniając tym samym w procesie selekcji geny, które odpowiadały za przyjazne zachowania wobec człowieka, wpłynęły z czasem na wygląd zwierząt oraz wywołały trwałe zmiany behawioralne związane z polepszeniem komunikacji międzygatunkowej. Badacze znajdują liczne dowody potwierdzające koncepcję samoudomowienia. Hare zauważa, że może ona tłumaczyć także pewne zmiany w zachowaniach i wyglądzie homo sapiens, który przeszedł podobny proces, jednak w obrębie własnego gatunku (rywalizując z innymi rodzajami homo). Badacze zauważają jednak, że przyjacielskie nastawienie wobec innych osobników danego gatunku nie działa wobec wszystkich ludzi, a jedynie w pewnym obszarze wspólnoty rozpoznawanej jako własna (wspólnoty kulturowej, cywilizacyjnej, etnicznej i tym podobne). Wzrastająca przyjacielskość wobec „swoich” daje bowiem w rezultacie wzrastającą wrogość wobec „obcych”, wynikającą z potrzeby obrony wspólnoty w razie niebezpieczeństwa (Hare tym przystosowaniem ewolucyjnym tłumaczy też konflikty etniczne i wojny). Zob. B. Hare, V. Woods, *Przetrwają najzycliwsi. Jak ewolucja wyjaśnia istotę człowieczeństwa?*, tłum. K. Kalinowski, Copernicus Center Press, Kraków 2022. Zob. też: Tomase llo, dz. cyt.

⁶² Por. A. Abid, M. Farooqi, J. Zou, *Large Language Models Associate Muslims with Violence*, „Nature Machine Intelligence” 3(2021), s. 461-463, <https://www.nature.com/articles/s42256-021-00359-2>; A. Abid, M. Farooqi, J. Zou, *Persistent Anti-Muslim Bias in Large Language Models*, arXiv:2101.05783v2 [cs.CL], January 18, 2021, <https://arxiv.org/pdf/2101.05783.pdf>.

sobie tendencyjności AI oraz takie jej przygotowanie, by w fazie aplikacji nie stymulowała zachowań dehumanizujących. W dalszej perspektywie może mieć to znaczenie dla przetrwania człowieka. Skoro AI może dehumanizować część ludzi, może dehumanizować ich jako gatunek, zyskawszy świadomość swej odrębności. Niektórzy badacze AI już donoszą, że przekraczamy ten próg⁶³.

Interesujących tropów dostarcza też memetyka, którą z koncepcją samoudomowienia łączy podejście do technologii, obie postrzegają ją bowiem jako narzędzie przyspieszające przemiany ewolucyjne. Memetyka, zwłaszcza w ujęciu Susan Blackmore⁶⁴, wychodzi z założenia, że technologia (a konkretnie jej najmniejsze elementy, czyli temy) wywołuje nowy rodzaj selekcji, analogicznego do ewolucji kulturowej (memów), a jeszcze wcześniej biologicznej (genów). Koncepcja samoudomowienia wyrasta natomiast z neoewolucjonizmu i konsiliencyjnego podejścia do nauk. Technologia, o jakiej tu mowa, to prehistoryczne narzędzia i zmiana organizacji społecznej i kultury. Badacze próbując zweryfikować hipotezę o samoudomowieniu człowieka, zwracają uwagę na fakt, że na uzyskanie przewagi nad innymi gatunkami wpływ miały pogłębiane dzięki selekcji zdolności komunikacyjne, które zapewniały lepszą, czyli zarówno dłuższą, jak i szerszą socjalizację, a ta z kolei dawała przewagę technologiczną i dostęp do wiedzy, która w większej populacji podlegała cyrkulacji oraz akumulacji i transmisji międzypokoleniowej. Obydwa te procesy przyspieszyły postęp technologiczny, stając się jego kołem zamachowym. Memetyka patrzy na ten proces z perspektywy przekazu informacji, jednak zasadniczo zgadza się z tymi konstatacjami.

Takie podejście do ewolucji inteligencji zakłada zatem pewną ciągłość między inteligencją biologiczną (człowieka) a jego wytworami (kulturą, technologią). Sztuczna inteligencja jest kolejnym wcieleniem inteligencji ludzkiej, niekoniecznie opozycją. To nie tyle metafora, co uzewnętrzniony wynik procesu przetwarzania przez homo sapiens bodźców ze środowiska, z którymi mamy do czynienia od tysiącleci. Sztuczna inteligencja oznacza możliwość eksterioryzacji tych procesów i uwolnienie ich od jednostkowej woli człowieka, jednak nie tylko jest dziełem człowieka, ale też trenowana jest na dużych zbiorach danych antropologicznych, co oznacza jej bliski związek z inteligencją biologiczną, lecz w skali całych społeczności. Warto jednak pamiętać, że choć ma ona swoje błędy, biasy i halucynacje, ich źródłem nie są kody, a antropogeniczne dane.

⁶³ Emily M. Bender z University of Washington zauważa, że antropomorfizacja sztucznej inteligencji wynika z samej metaforyki pojęć stosowanych w odniesieniu do procesu uczenia się AI (por. N. Tiku, *The Google Engineer Who Thinks the Company's AI Has Come to Life*, „The Washington Post”, 11.06.2022, <http://bit.ly/3YHwGk0>; Lemoine, dz. cyt.).

⁶⁴ Zob. S. Blackmore, *The Mem Machine*, Oxford University Press, Oxford 1999 [Wydanie polskie: taż: *Maszyna memowa*, przedm. R. Dawkins, tłum. N. Radomski, Poznań Rebis, 2002].

Lęki Nelsona czy Gooda zamienione zostały dziś w konkretne działania na rzecz rozwijania nowych technologii w sposób, który pozwoli opanować negatywne emocje wobec maszyn. Mnożą się inicjatywy działań na rzecz „dobrej AI”, postępuje też edukacja w zakresie programowania sztucznej inteligencji. Maszyny dziś to oprogramowanie z przyjaznym dla użytkownika interfejsem, na przykład użyteczne aplikacje i systemy informatyczne, roboty-zabawki, które pomagają w komunikacji przedszkolakom, osobom chorym czy starszym. W takim kontekście trudno myśleć o zagrożeniu (zapomnieliśmy już o wizjach z końca poprzedniego stulecia, takich jak filmowy Terminator), gdyż maszyny i oprogramowanie oparte na AI niosą pomoc. Pewne procesy są nieuniknione, jak rozwój cyfrowej pamięci czy współlistnienie starzejących się społeczeństw z robotami pełniącymi usługi pielęgnacyjne⁶⁵. Pokonanie lęku przed technologią staje się koniecznym etapem rozwoju społecznego.

Niezależnie od tego, że póki co nowe języki chatbotów wydają nam się zabawne i irracjonalne (na przykład „fejsbotlijski” stworzony przez chatboty Facebooka⁶⁶) – niosą one jednak pewną radykalną zmianę. Nie jesteśmy już jedyni w myśleniu kreatywnym. Następnym razem, patrząc na balet w wykonaniu robota KUKA i ludzkiego tancerza⁶⁷, dostrzec możemy zatem nie tylko piękno i współpracę, ale przede wszystkim kruchość człowieka wobec ramienia robota na co dzień składającego samochody w fabrykach. Kolejne czatujące roboty nie będą nas dziwić ani bawić, podobnie jak następne zwycięstwo sztucznej inteligencji w starożytnych (ludzkich) grach⁶⁸. Sprawniejsze przetwarzanie informacji jest ewolucyjnie bardziej efektywne, a to nie musi oznaczać ludzkiego sukcesu (w obszarze intelektualnym, duchowym czy egzystencjalnym). Koncepcja memetyczna zakłada, że celem genów, memów czy temów jest przetrwanie, niezależnie od „maszyn przetrwania” (a są nimi odpowiednio: ludzkie ciała, umysły, wytwory).

Koncepcja samoudomowienia daje natomiast nadzieję, że proces oswajania sztucznej inteligencji, podobnie jak procesy oswajania innych gatunków zwierząt i samoudomowienia homo sapiens, będzie postępował w kierunku wzmocnienia życzliwości oraz przyjaźliwości. Niesie także obserwacje związane z koniecznością opanowania agresji, lęków i tych emocji, które mogą wpłynąć na dehumanizację człowieka jako gatunku lub części ludzkich społeczności. Tak jak nauczyliśmy się komunikować ze zwierzętami i ludźmi, musimy nauczyć się komunikować z nowymi formami inteligencji, czyli oprogramowaniem AI, robotem, cyborgiem i innymi hybrydycznymi formami,

⁶⁵ Zob. *Mechanical Love*.

⁶⁶ Zob. Lewis in., dz. cyt.

⁶⁷ Zob. *A Duet of Human and Robot*, Ars Electronica, Linz 2017.

⁶⁸ Por. D. Silver in., *Mastering the Game of Go with Deep Neural Networks and Tree Search*, „Nature” 2016, nr 529, s. 484-489.

które w przyszłości mogą się pojawić. Oswajanie zachodzi dwukierunkowo, działa na obydwie strony procesu komunikacji – z m i e n i a i c z ł o w i e k a, i m a s z y n ę. Już dziś rozmowy ze sztuczną inteligencją pozwalają jej twórcom na wprowadzanie ulepszeń do kodów i ograniczeń w zakresie tendencyjności. Podobnie jak przemiany umysłu i ciała człowieka pod wpływem technologii zmieniły nasz gatunek, odmieni go także komunikacja ze sztuczną inteligencją. Być może wchodzimy właśnie nie tyle w epokę sztucznej czy hybrydycznej inteligencji, co w k o l e j n y e t a p s a m o u d o m o w i e n i a. Tym razem c z ł o w i e k b ę d z i e o s w a j a ł s z t u c z n ą i n t e l i g e n c j ę, a o n a c z ł o w i e k a. Może to oznaczać postępujące zbliżenie oraz symbiozę, podobną do tej, jaka zachodzi między człowiekiem a udomowionymi gatunkami zwierząt. Beneficjentami tego procesu są wszyscy, zarówno ludzie, jak i zwierzęta. Nie ma zatem powodów, by przypuszczać, że i tym razem będzie inaczej – wystarczy, by sztuczna inteligencja była nam życzliwa.

BIBLIOGRAFIA / BIBLIOGRAPHY

- Abid, Abubakar, Maheen Farooqi, and James Zou. "Large Language Models Associate Muslims with Violence." *Nature Machine Intelligence*, no. 3 (2021): 461–63. <https://www.nature.com/articles/s42256-021-00359-2>.
- Abid, Abubakar, Maheen Farooqi, and James Zou. *Persistent Anti-Muslim Bias in Large Language Models*, arXiv:2101.05783v2 [cs.CL]. January 18, 2021. <https://arxiv.org/pdf/2101.05783.pdf>.
- Agrawal, Ajay, et al. *Prediction Machines: The Simple Economics of Artificial Intelligence*. Boston: Harvard Business School Publishing, 2018.
- Bell, Gordon, Tony Hey, and Alex Szalay. "Beyond the Data Deluge." *Science*, 323 (5919) (2009). <https://science.sciencemag.org/content/323/5919/1297>.
- Beneteau, Erin, et al. "Parenting with Alexa: Exploring the Introduction of Smart Speakers on Family Dynamics." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. New York: ACM, April 2020: 1–13. <https://doi.org/10.1145/3313831.3376344>.
- Bielak, Tomasz, and Grzegorz Ptaszek. "Algorytmiczne doświadczanie kultury." *Kultura Współczesna*, no. 1 (104) (2019): 10–14.
- Bińczyk, Ewa. "Inżynieria klimatu a inżynieria człowieka. Dyskursy na temat środowiska w epoce antropocenu." *Ethos* 28, no. 3 (111) (2015): 153–75.
- Biological Turn. Idee biologii w humanistyce współczesnej*. Edited by Dobrosława Wężowicz-Ziółkowska and Emilia Wieczorkowska. Katowice: Wydawnictwo Uniwersytetu Śląskiego, 2016.
- Blackmore, Susan. *Maszyna memowa*. Translated by Norbert Radomski. Poznań: Rebis, 2002.

- Bohannon, John, et al. "Artificial Intelligence ups its game." In From AI to protein folding: Our Breakthrough runners-up, „Science”, 22.12.2016. <https://www.sciencemag.org/news/2016/12/ai-protein-folding-our-breakthrough-runners>.
- Bostrom, Nick. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press, 2014.
- . *Superinteligencja: Scenariusze, strategie, zagrożenia*. Translated by Dorota Konowrocka-Sawa. Gliwice: Wydawnictwo Helion, 2016.
- Brief Academic Biography of Marvin Minsky, MIT. <https://web.media.mit.edu/~minsky/minskybiog.html>.
- Brockman, Greg, Ilya Sutskever, and Open AI. *Introducing Open AI*. <https://openai.com/blog/introducing-openai/>.
- Chamorro-Premuzic, Tomas. "How ChatGPT Is Redefining Human Expertise: Or How To Be Smart When AI Is Smarter Than You." *Forbes*, January 12, 2023. <http://bit.ly/3Ef5PUM>.
- Dawkins, Richard. *Samolubny gen*. Translated by Marek Skoneczny. Warszawa: Prószyński i S-ka, 1996.
- A Duet of Human and Robot*. Ars Electronica, Linz 2017. YouTube. <https://www.youtube.com/watch?v=7moBSpAEkD4>.
- Ethos* 28, no. 3 (111) (2015): *Transhumanizm*. <https://czasopisma.kul.pl/index.php/ethos/issue/view/220>.
- Feigenbaum, Edward A. *Knowledge Engineering: The Applied Side of Artificial Intelligence* (1980). <https://stacks.stanford.edu/file/druid:qy055zd8682/qy055zd8682.pdf>.
- Filo-Sofija* 17, no. 39 (1) (2017). <http://www.filo-sofija.pl/index.php/czasopismo/issue/view/47/showToc>.
- Fukuyama, Francis. "Transhumanism: A Special Report." *Foreign Policy*, October 23, 2009. <https://foreignpolicy.com/2009/10/23/transhumanism/>.
- Garbowski, Marcin. "Transhumanizm: Geneza – założenia – krytyka." *Ethos* 28, no. 3 (111) (2015): 23–41.
- Good, Irving J. "Speculations Concerning the First Ultraintelligent Machine." *Advances in Computers* 6 (1965): 31–88. <https://www.sciencedirect.com/science/article/pii/S0065245808604180>.
- Gould, Stephen J. *Niewczesny pogrzeb Darwina*. Translated by Nina Kancewicz-Hoffman. Warszawa: Państwowy Instytut Wydawniczy, 1991.
- Gourley, Sean. *Big Data and the Rise of Augmented Intelligence*. TEDxAuckland, December 5, 2012. YouTube. https://www.youtube.com/watch?v=mKZCa_ejbfq.
- Gpt Generative Pretrained Transformer, Almira Osmanovic Thunström, and Steinn Steingrímsson. "Can GPT-3 write an academic paper on itself, with minimal human input?" HAL: Open Science Archive. <https://hal.science/hal-03701250>.
- Grudowska, Joanna, and Dominik Zieliński. "Społeczeństwo 5.0: Refleksja krytyczna." *Transformacje*, no. 2 (113) (2022): 232–55.
- Hare, Brian, and Vanessa Woods. *Przetwarzają najzyczliwsi: Jak ewolucja wyjaśnia istotę człowieczeństwa?* Translated by Kasper Kalinowski. Kraków: Copernicus Center Press, 2022.

- Hiniker, Alexis et al. "Can Conversational Agents Change the Way Children Talk to People?" In *Interaction Design and Children*, ACM, June 24, 2021: 338–49. <https://dl.acm.org/doi/10.1145/3459990.3460695>.
- Humanity+. <https://www.humanityplus.org/about>.
- H+pedia [wiki], s.v. "Transhumanist Declaration." https://hpluspedia.org/wiki/Transhumanist_Declaration.
- Infonomia. *Cyborg Life: Kevin Warwick*. YouTube. April 14, 2008. http://www.youtube.com/watch?v=RB_17SY_ngI.
- Irsak, Anna. "W obronie (tradycyjnej) medycyny." *Ethos* 28, no. 3 (111) (2015): 176–88.
- Ishiguro, Hiroshi. Symbiotic Human-Robot Interaction Project. <https://www.jst.go.jp/erato/ishiguro/en/index.html>.
- Janas, Agnieszka. *Zastosowanie agentów konwersacyjnych w nauczaniu*. „Nierówności Społeczne a Wzrost Gospodarczy”, no. 22 (2011): 281–87.
- Kleinman, Zoe. "Apple's Siri Calls Ambulance for Baby." *BBC News*, June 7, 2016. <https://www.bbc.com/news/technology-36471180>.
- Kobayashi, Kenji, and Ming Hsu. "Common Neural Code for Reward and Information Value." *PNAS* 116, no. 26 (2019): 13061–13066. <https://doi.org/10.1073/pnas.1820145116>.
- Krzysztofek, Kazimierz. "Człowiek – społeczeństwo – technologie: Między humanizmem a transhumanizmem i posthumanizmem." *Ethos* 28, no. 3 (111) (2015): 191–213.
- Kurzweil, Raymond. *The Singularity is Near: When Humans Transcend Biology*. London: Duckworth, 2009 [wydanie polskie: *Nadchodzi osobliwość. Kiedy człowiek przekroczy granice biologii*. Translated by Eliza Chodkowska and Anna Nowosielska. Warszawa: Kurhaus Publishing, 2016].
- Lemoine, Blake. "Is LaMDA Sentient? – an Interview." *Medium*, June 11, 2022. <https://cajundiscordian.medium.com/is-lambda-sentient-an-interview-ea64d916d917>.
- Lewis, Mike, et al. *Deal or No Deal? End-to-End Learning for Negotiation Dialogues*. <https://s3.amazonaws.com/end-to-end-negotiator/end-to-end-negotiator.pdf>.
- Łaszczycza, Piotr. "Człowiek i jego maszyny: Operatorzy i protezy." *Filo-Sofija* 17, no. 39 (1) (2017): 49–64.
- Maj, Anna. *Przemiany wiedzy w cyberkulturze: Badania nad kulturą, komunikacją, wiedzą i mediami*. Katowice: Wydawnictwo Uniwersytetu Śląskiego, 2021.
- Mechanical Love*. Directed by Phie Ambo. Finland 2007.
- Microsoft Learn. *Co to są agenci sztucznej inteligencji?* December 1, 2022. <https://learn.microsoft.com/pl-pl/azure/cloud-adoption-framework/innovate/best-practices/conversational-ai>.
- Miczka, Tadeusz. "Specifics of Algorithmization in Data Culture." Paper presented at the ICSRS 2019: Conference Proceedings, International Conference on Recent Social Studies and Research, CORIS Department Sapienza University, Rome, 25-26 October 2019.

- Minsky, Marvin Lee. *The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind*. New York, London, Toronto, and Sydney: Simon & Schuster, 2006.
- . *The Society of Mind*. New York, London, Toronto, and Sydney: Simon & Schuster, 1985.
- Mintz, Steven. “ChatGPT: Threat or Menace?” *Inside Higher Ed*, January 16, 2023. <https://bit.ly/412ObwI>.
- Mori, Masahiro. *The Uncanny Valley* (1970). Translated by Karl F. MacDorman and Norri Kageki. *IEEE Spectrum*, June 12, 2012. <https://spectrum.ieee.org/the-uncanny-valley>.
- Musk, Elon. Twitter. August 3, 2014. <https://twitter.com/elonmusk/status/495759307346952192>.
- Nelson, Ted. *Computer Lib/Dream Machines*. <http://worrydream.com/refs/Nelson-ComputerLibDreamMachines1975.pdf>.
- Olzacka, Elżbieta. “Od Homo sapiens do Homo immortalis: Idea nieśmiertelności w rosyjskich projektach filozoficznych i społeczno-politycznych.” In *Technokultura: transhumanizm i sztuka cyfrowa*. Edited by Damian Gałuszka et al. Kraków: Wydawnictwo Libron, 2016.
- Open AI. <https://openai.com/about/>.
- Rheingold, Howard. *Narzędzia ułatwiające myślenie: Historia i przyszłość metod poszerzania możliwości umysłu*. Translated by Jacek B. Szporcko. Warszawa: Wydawnictwa Naukowo-Techniczne PWN-WNT, 2003.
- Silver, David, et al. “Mastering the Game of Go with Deep Neural Networks and Tree Search.” *Nature*, no. 529 (2016): 484–89.
- Stelarc. <http://stelarc.org>.
- Stork, David G. “Scientist on the Set: An Interview with Marvin Minsky.” In *HAL’s Legacy: 2001’s Computer as Dream and Reality*. Edited by David G. Stork. Foreword by Arthur C. Clarke. Cambridge and London: MIT Press, 1998. <http://mitpress.mit.edu/e-books/Hal/chap2/two1.html>.
- The Essential Turing: Seminal Writings in Computing, Logic, Philosophy, Artificial Intelligence, and Artificial Life; Plus The Secrets of Enigma*. Edited by B. Jack Copeland, Oxford: Oxford University Press, Clarendon Press, 2004.
- The Fourth Paradigm: Data-intensive Scientific Discovery*. Edited by Tony Hey, Stewart Tansley, and Kristin Tolle. Redmont and Washington: Microsoft Research, 2009.
- The Encyclopedia of Computer Science and Engineering*, s.v. “Expert Systems: Principles and Practice” (by Edward A. Feigenbaum). 1992. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.34.9207&rep=rep1&type=pdf>.
- Thunström, Almira, Osmanovic. “We Asked GPT-3 to Write an Academic Paper about Itself – Then We Tried to Get It Published.” *Scientific American*, June 30, 2022. <http://bit.ly/3IDS4YM>.
- Tiku, Nitasha. “The Google Engineer Who Thinks the Company’s AI Has Come to Life.” *The Washington Post*, June 11, 2022. <http://bit.ly/3YHwGk0>.

- Tomasello, Michael. *Historia naturalna ludzkiego myślenia*. Translated by Bartłomiej Kucharzyk and Rafał Ociepa. Kraków: Copernicus Center Press, 2015.
- Traeger, Margaret L. et al. "Vulnerable Robots Positively Shape Human Conversational Dynamics in a Human–Robot Team." *PNAS* 117, no. 12 (2020): 6370-6375. <https://www.pnas.org/content/117/12/6370>.
- Warnke, Martin. "Databases as Citadels in The Web 2.0." In *Un Like Us Reader: Social Media Monopolies and Their Alternatives*. Edited by Geert Lovink and Miriam Rasch. Amsterdam: Institute of Network Cultures, 2013.
- Warwick, Kevin. *I, Cyborg*. Urbana and Chicago: University of Illinois Press, 2004.
- Weiser, Mark, and John Seely Brown. *Designing Calm Technology*. Xerox PARC, December 21, 1996. <https://people.csail.mit.edu/rudolph/Teaching/weiser.pdf>.
- Weiser, Mark. "The Computer for the 21st Century." *Scientific American* 265, no. 3 (1991): 94–104.
- Wilson, Edward O. *Konsiliencja: Jedność wiedzy*. Translated by Jarosław Mikos. Poznań: Wydawnictwo Zysk i S-ka, 2002.
- Zacher, Lech W. "Od społeczeństwa informacyjnego do społeczeństwa wiedzy (dylematy tranzycyjne: między informacją, wiedzą i wyobraźnią)." In *Społeczeństwo informacyjne: Wizja czy rzeczywistość?* Vol. 1. Edited by Lesław H. Haber. Kraków: Akademia Górniczo-Hutnicza, 2004.

ABSTRAKT / ABSTRACT

Anna MAJ – Ewolucja inteligencji i jej badanie. Oswajanie sztucznej inteligencji w perspektywie przemian kultury i komunikacji

DOI 10.12887/36-2023-3-143-11

Celem artykułu jest przedstawienie ewolucji inteligencji (od biologicznej, przez sztuczną, po rozszerzoną) i dotychczasowych interdyscyplinarnych badań nad sztuczną inteligencją, między innymi rozwoju sposobów rozumienia terminu „inteligencja” w interesującym obszarze. W tym celu wybrano kilka podstawowych koncepcji. Ukazano także różnorodne konteksty osvajania sztucznej inteligencji w perspektywie przemian kultury i komunikacji. Wskazano na przemiany komunikacji w obrębie relacji międzyludzkich, przestrzeni domowej i publicznej, a także popkulturowych i artystycznych wizji ciała z obszaru sztuki mediów, na które znaczny wpływ mają koncepcje transhumanizmu i cyborgizacji. Wykorzystano analizę dyskursu oraz wyniki wieloletniej obserwacji uczestniczącej prowadzonej w obszarze sztuki mediów i technologii informatycznych. Ewolucji sposobów rozumienia pojęcia „inteligencja” towarzyszy zmiana w obszarze projektowania sztucznej inteligencji i samej wyobraźni twórców. Protechnologiczny dyskurs zachęca użytkowników technologii do interakcji z agentami konwersacyjnymi AI i robotami, prowadząc do humanizacji technologii, a jednocześnie do pozytywnego wartościowania maszyny oraz postrzegania cyborgizacji jako naturalnego i dobroczynnego procesu. Badania

te mogą być rozwijane w kierunku dalszej analizy przemian dyskursu dotyczącego sztucznej inteligencji, jej rozwoju i wzrastającej komplikacji relacji człowiek–technologia, na co wskazano, przedstawiając koncepcję samoudomowienia i interpretację memetyczną ewolucji.

Słowa kluczowe: sztuczna inteligencja, cyborgizacja, robotyzacja, interfejs konwersacyjny, zmiany zachowań komunikacyjnych

Kontakt: Instytut Nauk o Kulturze, Wydział Humanistyczny, Uniwersytet Śląski, ul. Uniwersytecka 4, 40-007 Katowice

E-mail: anna.maj@us.edu.pl

Tel. 32 2009323

<https://us.edu.pl/institut/inok/osoby/anna-maj/>

ORCID: 0000-0003-3958-267X

Anna MAJ, *The Evolution of Intelligence and Its Research: Taming Artificial Intelligence, as seen from the Perspective of Changes in Culture and Communication*

DOI 10.12887/36-2023-3-143-11

The article aims to present the evolution of intelligence (from biological, through artificial, to augmented) and the existing interdisciplinary research on artificial intelligence, including developing ways of understanding the term “intelligence” in this area. For this purpose, several basic concepts have been selected. The paper also shows various contexts of taming artificial intelligence from the perspective of changes in culture and communication. The author discusses changes in communication within interpersonal relationships, home, and public space, as well as pop culture and artistic visions of the body in the area of media art, which are significantly influenced by the concepts of transhumanism and cyborgisation. The discussion is based on discourse analysis and results of many years of participant observation conducted in media art and information technology. The evolution of the ways of understanding the term “intelligence” is accompanied by a change in artificial intelligence design and in the very imagination of its creators. The pro-technology discourse encourages technology users to interact with conversational AI agents and robots, leading to the humanisation of technology and simultaneously to the positive evaluation of the machine and the perception of cyborgisation as a natural and beneficial process. Further observation and analysis of changes in the discourse on artificial intelligence, its development, and the increasing complication of the human–technology relationship will help understand the evolution of the phenomenon in question. The presented theories, the concept of self-domestication, and the memetic interpretation of evolution provide promising theoretical tools in this field.

Keywords: artificial intelligence, cyborgisation, robotisation, conversational interface, changes in communication behavior

Contact: Institute of Cultural Studies, Faculty of Humanities, University of Silesia, ul. Uniwersytecka 4, 40-007 Katowice, Poland

E-mail: anna.maj@us.edu.pl

Phone: +48 32 2009323

<https://us.edu.pl/institut/inok/osoby/anna-maj/>

ORCID: 0000-0003-3958-267X