

Mirostław Sopek

MakoLab SA

ORCID: 0000-0003-0378-5125

sopek@makolab.com

METAFORY SZTUCZNEJ INTELIGENCJI

DOI: 10.52097/acapress.9788362475810.73-98

Słowa kluczowe: Streszczenie

szuczna inteligencja, metafony, technologia, historia, teoria, praktyka

Metafony odgrywają w sztucznej inteligencji znaczącą rolę, zarówno w teorii, jak i w praktyce tej dyscypliny. Co więcej, to, w jaki sposób rozumiemy sztuczną inteligencję, jest wynikiem procesu myślowego o głęboko metaforycznym charakterze. Obecność metafor nie jest jednak zauważana ani doceniana przez jej obecnych adeptów, pomimo iż w tradycyjnej, symbolicznej sztucznej inteligencji (zwanej często GOFAI, Good Old Fashioned AI, czyli „staromodną sztuczną inteligencją”) powstawały rozwiązania zdolne do przetwarzania metafor – ich wykrywania i ujawniania ich dosłownych znaczeń. Pierwszym celem niniejszej publikacji jest przede wszystkim wskazanie i zdefiniowanie metaforycznego charakteru rozumienia SI, gdyż niedostrzeżenie jego obecności przyczynia się do wzmacniania nieporozumień dotyczących istoty współczesnej sztucznej inteligencji. Drugim celem publikacji jest opis sposobów, jakimi tradycyjna (symboliczna) sztuczna inteligencja próbowała zmierzyć się z metaforycznym sposobem opisu rzeczywistości, oraz jak tematykę metafor traktuje współczesna sztuczna inteligencja oparta na metodach sztucznych sieci neuronowych i uczenia maszynowego. W podsumowaniu autor postuluje włączenie zagadnień dotyczących metafor zarówno do dydaktyki podstaw teoretycznych (*theoria*), jak i zastosowań sztucznej inteligencji (*praxis*).

Keywords: Abstract

artificial intelligence, metaphors, technology, history, theory, practice

Metaphors play a significant role in artificial intelligence, both in theory and in practice of the discipline. Moreover, how we understand AI is the product of a deeply metaphorical thought process. The presence of metaphors, however, is not noticed or appreciated by its current adepts, despite the fact that in the traditional, symbolic artificial intelligence (often called GOFAI – Good Old Fashioned AI) various solutions were created that were capable of processing metaphors – not only their detection but also the disclosure of their literal meanings. The first goal of this article is, above all, to uncover the metaphorical nature of our thinking about AI. The failure to perceive the presence of metaphors in it, contributes to the amplification of misunderstandings about the very essence of modern artificial intelligence. The second goal of the article is to describe the ways in which traditional (symbolic) artificial intelligence has tried to deal with the metaphorical description of reality, and then, how the subject of metaphors is treated by contemporary artificial intelligence based on the methods of artificial neural networks and machine learning. In conclusion, the author proposes to include the issues of metaphors both in the education of theoretical foundations (theoria) and the applications of artificial intelligence (praxis).

Wstęp – metafory w nauce i technologii

Metafory są wzajemnym, choć zwykle asymetrycznym odzwierciedleniem różnych obszarów konceptualnych ludzkiej wiedzy i jako takie zajmują centralne miejsce w nauce. Są one częścią składową głównego mechanizmu poznawczego, poprzez który „dokonujemy zrozumienia abstrakcyjnych pojęć oraz za pomocą którego przeprowadzamy abstrakcyjne rozumowanie” (Lakoff, 1993, s. 202). Rozumienie ich roli w nauce zawdzięczamy autorowi tych słów, George’owi Lakoffowi, który wydobył na światło dzienne ich zasadnicze miejsce w nauce, ukazując, iż zdecydowana

większość zarówno popularnych zagadnień „technologicznych”, jak i najbardziej zawiłych teorii naukowych, jest rozumiana niemal wyłącznie poprzez konceptualne metafory.

W związku z fundamentalną rolą metafor w nauce, jak też ich powszechnością w języku potocznym, przytaczanie konkretnych przykładów zawsze będzie obarczone pewną tendencyjnością wynikającą z „macierzystej” dyscypliny naukowej przytaczającego przykłady. Tak też zapewne jest w wypadku autora tej publikacji, który jest fizykiem z wykształcenia. Świadom tej tendencyjności, autor zwraca uwagę, iż język, jakim fizycy opisują atomy, korzystający z analogii do Układu Słonecznego, jest silnie metaforyczny. Mówimy o *orbitalach* atomowych, używając wprost metafory orbit planetarnych. W większości prób wyjaśnienia, czym jest spin, stosuje się metaforę *małego* magnesu albo rotującego obiektu. Thomas Kuhn uważał, że metafora kulek ping-pongowych lub bilardowych ma zasadnicze znaczenie dla rozumienia teorii mikroświata. W pracy *Metaphor in Science* Kuhn pisze: „Bez pomocy (metafor kul) nie można nawet dzisiaj zapisać równania Schrödingera dla złożonego atomu lub cząsteczki, ponieważ różne człony tego równania odnoszą się do modelu, a nie bezpośrednio do natury” (Kuhn, 1993, s. 538). Kiedy dziś objaśniamy podstawową jednostkę komputerów kwantowych – qubit – najczęściej stosujemy metaforę wektora zamkniętego w sferze (korzystając tu z pojęcia sfery Blocha). Wiąże się z tym oczywiście również sporo problemów i nieporozumień, gdy zapominamy, że mamy do czynienia z metaforami – co tak często zachodzi w nauczaniu fizyki.

Z bardzo interesującym przypadkiem użycia metafor mamy ostatnio do czynienia na pograniczu psychologii ewolucyjnej oraz epistemologii, gdzie powstała teoria percepcji znana pod nazwą teorii percepcji jako interfejsu (Interface Theory of Perception, ITP) (Hoffman, Singh, Prakash, 2015, s. 1480–1506). Autorzy tej koncepcji zakładają wprost, iż strategię percepcyjną preferowaną przez naturalną selekcję najlepiej postrzegać nie jako *okno* na prawdę, ale jako okno podobne do obiektu znanego z interfejsu systemu Windows na komputerach PC. W tym wypadku siła tej metafory wynika z powszechności rozumienia, czym są okna w systemach komputerowych.

Podobnie rzecz ma się z metaforą sprawności (fitness) w niektórych teoriach ewolucji percepcji. Mowa jest tam o tym, że „sprawność pokonuje prawdę” („fitness beats truth”; Prakash, Stephens, Hoffman, Singh, Fields, 2020). Jest to język jawnie metaforyczny, zrozumiały dzięki popularności pojęcia „fitness” we współczesnej cywilizacji.

Spośród wielu prób podkreślenia roli metafor w nauce warto zwrócić uwagę na opinie badaczy takich jak Richard Boyd, według którego „metafory mają czasami zasadnicze znaczenie dla tworzenia nowych teorii naukowych”. Takie metafory określa mianem „metafor konstytutywnych dla teorii” (Boyd, 1993, s. 485). Z kolei Andrew Ortony, który na zadane przez siebie pytanie „Do czego służą metafory?” odpowiada: „Pozwalają one na artykulację nowych idei, których teorie naukowe są tylko szczególnym przypadkiem. Jest to funkcja, której nie sposób zawsze zrealizować, używając wyłącznie języka dosłownego” (Ortony, 1993, s. 114). Wreszcie Hugh G. Petrie i Rebecca S. Oshlag argumentują: „Główną funkcją metafory jest umożliwienie zrozumienia nowych pojęć poprzez iteracyjny proces kolejnych przybliżeń, mający swój początek w lepiej znanym metaforycznym wehikule pojęciowym” (Petrie, Oshlag, 1993, s. 15)

Metaforyczne rozumienie sztucznej inteligencji jako funkcji stanu rozwoju technologii

Niezwykle interesujące spojrzenie na rolę metafor w rozwoju sztucznej inteligencji przedstawia George Zarkadakis w pracy *Na nasz obraz i podobieństwo. Czy sztuczna inteligencja nas zbawi czy zniszczy?* (Zarkadakis, 2015). Zarkadakis zauważa, że dominująca w danym okresie historycznym technologia narzuca specyficzne metafory myślenia o sztucznej inteligencji. Oczywiście we wczesnych fazach rozwoju cywilizacji nie odnajdziemy myślenia o sztucznej inteligencji rozumianej w sposób zbliżony do naszego współczesnego rozumienia. Odnajdujemy natomiast silnie metaforyczne myślenie o istocie umysłu i inteligencji jako takiej oraz ich związku z materią lub ogólnie rozumianym „ciałem” i jego

życiem. Metafory te ewoluują skokowo, dzięki zjawiskom znanym w historii nauki jako zmiany paradygmatu (*paradigm shift* – termin ten zawdzięczamy Thomasowi Kuhnowi). Zarkadakis wyodrębnia sześć głównych zmian wpływających na myślenie o istocie umysłu. Zmiany te polegały na cyklicznym zastępowaniu metafor, za pomocą których usiłowano wyjaśnić naturę umysłu nowymi metaforami rodzącymi się wraz ze zmianami paradygmatów technologicznych.

Pierwszą znaną metaforą (występującą zarówno w greckich, jak i żydowskich mitach) jest metafora prochu ziemi lub błota (*dust, mud*). Według Biblii człowiek jest stworzony z „prochu ziemi”, zaś w mitologii greckiej Prometeusz tworzy człowieka z błota (gliny), a Atena swoim tchnieniem go ożywia. Co więcej, mimo wielu oczywistych różnic pomiędzy judaizmem a religią grecką, podobieństwa między tymi mitami stworzenia są uderzające, a mają one swoje źródło w mitach poprzednich cywilizacji z rejonu Mezopotamii. Obserwacja Zarkadakisa wydaje się wiarygodna: w tym czasie dominującą „technologią” było rolnictwo, a w powszechnym rozumieniu, wynikającym z obserwacji natury, życie wyrasta z ziemi albo bezpośrednio (rośliny), albo pośrednio (zwierzęta). Znajdujemy ślady tej metafory również w języku: łacińskie słowo *hominum* jest powiązane ze słowem *humus* (ziemia) – ten związek występuje także w języku angielskim (*human – humus; the organic component of soil* – organiczny składnik gleby).

Drugą metaforą, która ujawniła się kilka wieków później, po raz pierwszy w czasach Hipokratesa, a później Archimedesesa, była metafora hydrauliczno-pneumatyczna. W metaforze tej tym, co ożywia obiekty nieożywione, jest woda i para. Widać to wyraźnie u Hipokratesa i jego teorii czterech soków (humorów): krwi, żółci, śluzu i czarnej żółci. Teoria ta dzięki Galenowi (Claudius Galenus, 130–200 n.e.) wpływała na rozwój medycyny i nauk pokrewnych niemal aż do okresu oświecenia. Pierwszą „sztuczną”

¹ „I tak utworzył Wiekuisty Bóg człowieka z prochu ziemi, i tchnął w nozdrza jego dech żywota, i stał się człowiek istotą żyjącą” (*Biblia hebrajska*, Rdz, 2:7 tłum. I. Cyłkow).

istotą, jaką spotykamy w tym czasie, jest Talos – jeden z bohaterów epickiego poematu *Argonautica* Apolloniusza Rodyjskiego. Talos był jednym z pierwszych mitologicznych automatów – protoplastów późniejszych robotów. Była to postać z brązu, ożywiona dzięki żył, w której płynął „ichor” – eteryczny płyn, jaki odnajdujemy również, jako „krew bogów”, w innych mitach z tego okresu. W czasach cesarstwa rzymskiego i Bizancjum powstało wiele realnych automatów.

W kulturze muzułmańskiej tej epoki sławę zdobywa Ismail Al-Jazari, który w dziele *Encyklopedia wiedzy o genialnych urządzeniach mechanicznych* zasłynął opisem stu mechaniczno-hydraulicznych urządzeń. Jego najsłynniejszym wynalazkiem był zegar wodny w kształcie słonia. Zmiana metafory na hydrauliczno-pneumatyczną, z jaką mamy tu do czynienia, wynika bezpośrednio z rozwoju technologii, w której dominująca rola technologii czysto rolniczych zastąpiona została technologiami korzystającymi ze zjawisk hydraulicznych i pneumatycznych.

Kolejną, trzecią metaforą, jaką zaczynamy obserwować w czasach renesansu, jest metafora maszyny. W tym okresie (XV–XVI wiek) następuje znaczny rozwój technologii czysto mechanicznych, zastępujących rozwiązania hydrauliczne. Paradygmat uległ wyraźnej zmianie. Zmianę tę zapowiadały niezwykle interesujące prace z zakresu logiki i filozofii prowadzone w wiekach średnich. Szczególne znaczenie mają tu prace Ramona Llulla, będące jedną z pierwszych prób mechaniczno-symbolicznego wnioskowania za pomocą kombinacji różnych układów symboli. Jego *Ars Magna* (zwana też *Ars Generalis Ulitima*), czyli *Wielka, ogólna i ostateczna sztuka*, opublikowana w 1305 roku, była inspiracją dla myślicieli takich jak: Juan L. Vives, Blaise Pascal czy też Gottfried W. Leibniz, a współcześnie jest uznawana przez autorytety nowoczesnej informatyki (np. Knuth, 2004) i sztucznej inteligencji (Nilsson, 2009, s. 20) za pionierskie dzieło z zakresu teorii obliczeń. Dokonania te były tak istotne, iż można by się pokusić o klasyfikację ich jako metafor symboliczno-mechanicznych.

Epoka maszyn rozpoczyna się ok. XV wieku, gdy powstają pierwsze mechaniczne zegary sprężynowe. Powoli maszyny zaczy-

nają wkraczać w różne dziedziny życia, rodzi się przemysł oparty na nowych możliwościach, jakie otwierają się dzięki mechanizacji produkcji. Zmiana paradygmatu jest tak znacząca, że zaczyna wpływać na inne obszary życia, w tym na filozofię. Kartezjusz był pierwszym myślicielem, którego dzieło ucieleśnia nową metaforę dla życia – metaforę maszyny. W jego *Rozprawie o metodzie* znajdujemy słynny wywód:

Nie wyda się to zgoła dziwne tym, którzy wiedzą, jak wiele różnych automatów, czyli poruszających się maszyn, ludzka przemyślność zdolna jest wytworzyć, posiłkując się bardzo niewielką liczbą części w porównaniu do wielkiej mnogości kości, mięśni, nerwów, tętnic, żył i wszystkich innych składników znajdujących się w ciele każdego zwierzęcia; będą oni uważali ciało za maszynę, która jako wykonana ręką Boga jest nieporównywalnie lepiej urządzona i ma w sobie ruchy bardziej zadziwiające aniżeli jakakolwiek z maszyn wymyślonych przez człowieka (Descartes, 1981, s. 29).

Najbardziej znaczącymi przykładami ukazującymi istotę metafory mechanicznej są prace Leonarda da Vinci (projekt humanoidalnego robota powstał w 1495 roku), rozważania Thomasa Hobbesa o możliwości zbudowania sztucznego zwierzęcia (1651), czy też słynna mechaniczna kaczka Jacques'a de Vaucanson (1738). O ile na mechaniczną kaczkę można patrzeć nieco z przymrużeniem oka, to według George'a Dysona (2012) Hobbes zasługuje na tytuł „patriarchy sztucznej inteligencji”.

Czwartą metaforą związaną z kolejną zasadniczą zmianą paradygmatu w nauce i technologii była metafora elektryczno-telegraficzna, będąca skutkiem XVIII- i XIX-wiecznych odkryć związanych z elektrycznością i jej industrialnymi zastosowaniami. Jednocześnie z wytwarzaniem i używaniem energii elektrycznej fundamentalne dla powstania tej nowej metafory były odkrycia bioelektryczności. Słynne są eksperymenty Luigiego Galvaniego z roku 1791 roku ze „zwierzęcą elektrycznością”. W ślady Galvaniego poszedł Alessandro Volta, a około pół wieku później Emil du Bois-Reymond, prowadzący w systematyczne badania nad zjawiskami elektrycznymi u zwierząt i ludzi.

Odkrycia te doprowadziły do powstania koncepcji witalizmu w rozumieniu życia i inteligencji ludzkiej. Witalizm postulował istnienie tajemniczej siły życiowej (*élan vital*) kojarzonej właśnie z nieuchwytnymi i wtedy jeszcze tajemniczymi zjawiskami elektrycznymi. W 1851 roku du Bois-Reymond wygłasza wykład, w którym stwierdza:

Cud naszych czasów, telegrafia elektryczna, został już dawno temu wymodelowany w zwierzęcej maszynie. Ale podobieństwo między tymi dwoma systemami, nerwowym i telegraficznym, ma znacznie głębsze podstawy. To coś więcej niż zwykłe podobieństwo; jest to pokrewieństwo nie tylko skutków, ale być może także przyczyn (Otis, 2002, tłum. M.S.).

Uczonym, któremu ostatecznie przypisuje się sformułowanie metafory elektryczno-telegraficznej, był słynny fizyk Hermann von Helmholtz. Twórca prawa zachowania energii miał wręcz niezwykle osiągnięcia w fizjologii i medycynie. Już w 1850 roku dokonał pomiaru prędkości impulsów nerwowych, wynalazł oftalmoskop – przyrząd do badania dna oka – i sformułował teorię widzenia kolorów, jaką posługujemy się do dziś. Helmholtz, pomimo świadomości (wynikającej z jego starannych badań) podstawowych różnic pomiędzy systemem telegraficznym a układem nerwowym (np. w szybkości przekazywania sygnałów), napisał w 1868 roku:

Włókna nerwowe są często porównywane z przewodami telegraficznymi pokrywającymi obszar kraju, a porównanie to dobrze ilustruje uderzającą i ważną osobliwość w ich sposobach działania. W sieci telegraficznej znajdziemy wszędzie te same miedziane lub stalowe kable niosące ten sam rodzaj ruchu, a mianowicie strumień energii elektrycznej, ale wytwarzający bardzo różne wyniki w różnych stacjach, w zależności od urządzeń, do jakich są podłączone (...) Zarówno włókna nerwowe, jak i przewody telegraficzne są uderzającymi przykładami ilustrującymi doktrynę, że te same przyczyny mogą, w zależności od warunków, wywoływać różne skutki (von Helmholtz, 1995, s. 150).

Dalszy rozwój nauki i technologii w pierwszej połowie XX wieku przyczynił się do kolejnej zmiany paradygmatu. We wczesnych latach czterdziestych technologie „elektryczne” stały się już na tyle dojrzałe, że doprowadziły do powstania prekursorów komputerów. Niewątpliwie pierwszym urządzeniem zasługującym na miano „komputera” była maszyna ABC, czyli Atanasoff-Berry Computer. Co prawda, nie ma do dziś pełnej zgody, czy ABC był rzeczywiście uniwersalnym komputerem, ale nie ma też wątpliwości, iż jego konstrukcja wyznaczyła kierunek rozwoju wczesnych komputerów. Wczesnymi następcami, posiadającymi już cechy uniwersalności, były wynalazki: Z₃ Konrada Zusego – zaprojektowany w 1935 roku i zbudowany w 1941 roku – oraz Colossus – brytyjski komputer zaprojektowany w celach kryptoanalizy szyfrów używanych przez Niemców. Twórca teoretycznych podstaw technologii komputerowej, Alan Turing, choć miał istotny wpływ na projekt Colossusa, jego osiągnięcia były większe niż jego udział w konstrukcji tego komputera. Po pierwsze, był niewątpliwym autorem projektu systemu o nazwie Bombe (generalizacja projektu Bomba polskiego kryptologa Mariana Rejewskiego), z pomocą którego Turing w 1940 roku dokonał ostatecznego złamania szyfrów niemieckiej Enigmy. Po drugie, jego prace teoretyczne de facto stworzyły teoretyczne podstawy nowego paradygmatu, szczególnie prace *On Computable Numbers, with an Application to the Entscheidungsproblem*, opublikowana w 1936 roku, oraz *Computing Machinery and Intelligence* z roku 1950.

Pierwszym uniwersalnym komputerem był niewątpliwie ENIAC (Electronic Numerical Integrator and Computer) skonstruowany w 1945 przez Johna P. Eckerta oraz Johna W. Mauchly’ego z University of Pennsylvania z pomocą słynnego Johna von Neumanna, autora obowiązującego do dziś wzorca architektury komputera (architektura von Neumanna). Historia komputerów od tego właśnie momentu gwałtownie przyspiesza. Jest ona oczywiście niezmiernie interesująca, ale z naszego punktu widzenia najważniejsze jest powstanie zupełnie nowego paradygmatu technologicznego z centralnym miejscem maszyny liczącej o architekturze von Neumanna, zdolnej do uniwersalności przewidzianej przez Turinga (maszyna Turinga).

Zmiana ta wywołała powstanie kolejnej, piątej metafory, w której z jednej strony komputery nazywane były „mózgami elektronowymi”, a z drugiej strony w opisie mózgu odwoływano się do architektury komputerów. Wraz z rozwojem technologii powstały maszyny, które były zdolne do systematycznego i szybkiego odtwarzania najbardziej nawet złożonych procesów logicznych. W połączeniu z przetwarzaniem danych maszynowe procesy logiczne stały się zdolne do tworzenia nowej wiedzy. Nie sposób się dziwić, że dla bardzo wielu ludzi były zdolne do tego, co nazywamy myśleniem.

Wreszcie w drugiej połowie XX wieku, gdy technologia komputerowa jest już w pełni rozwinięta, powstają sieci komputerowe i słynny slogan „Dopiero sieć to komputer” („The Network is the Computer”) stworzony przez Johna Gage’a, pracownika firmy Sun Microsystems. Pierwsze sieci komputerowe łączyły komputery pracujące w obrębie pewnych lokalizacji. Zwiększało to ich moc przetwarzania, ale dopiero gdy w latach siedemdziesiątych i osiemdziesiątych XX wieku powstawał Internet, a szczególnie od powstania sieci Web (1989), zaczynamy obserwować zjawisko, jakie możemy nazwać „szóstą metaforą”. Zarkadakis wyraża to pięknie, pisząc: „Internet jest «jak» mózg, a mózg jest «jak» Internet” („The Internet is «like» a brain and the brain is «like» the Internet”) (Zarkadakis, s. 71). Kiedy na początku XXI wieku powstała sieć o nazwie Semantic Web, połączone w sieć komputery były zdolne do wykrywania i kodowania znaczeń. Autor niniejszego artykułu tłumaczył nawet nazwę tego zjawiska jako „sieć pełną znaczeń” (Sopek, 31.01.2020; Zykubek, 24.11.2016), mając oczywiście pełną świadomość tego, iż mówimy tutaj wyłącznie za pomocą pewnej metafory (jedynie człowiek jest w stanie pojmować i nadawać w sensie dosłownym znaczenia).

Metafora inteligentnej sieci jest szczególnie widoczna w neuro-nauce. To na jej gruncie dokonuje się dziś konstrukcji tzw. konektomów (*connectomes*) jako kompletnej mapy połączeń neuronowych. O ile samo badanie połączeń neuronowych jest dobrze zdefiniowaną procedurą naukową, o tyle już ich interpretacje pokazują, jak silna jest ta szósta metafora. Sebastian Seung (July 2010), znany przedstawiciel neuronauki obliczeniowej, używa powiedzenia

„Jestem moim konektorem” („I Am My Connectome”), a jego słynna książka *Connectome: How the Brain's Wiring Makes Us Who We Are* [Konektom: jak połączenia mózgowe czynią nas tymi, którymi jesteśmy], wyraża tę metaforę bardzo dosadnie.

Nie ma wątpliwości, że żyjemy dziś pod urokiem tej metafory, choć wydaje się, że częściej działa ona niejako w drugą stronę. Nie-samowite sukcesy, jakie odnosi współczesna sztuczna inteligencja, zdolna do rozpoznawania obrazów i dźwięków, analizy tekstów, sterowania autonomicznymi samochodami, przewidywania pogody i struktury białek (a listę tę można by bardzo wydłużyć), są wynikiem zastosowania metafory struktur mózgu w informatyce. Pojęcia takie jak sieci neuronowe, maszynowe uczenie, głębokie uczenie czy wreszcie inżynieria neuromorficzna są składowymi języka ściśle metaforycznego! Każde z tych pojęć de facto opisuje pewne specyficzne algorytmy i rozwiązania komputerowe, często bardzo złożone, ale jednak pomimo to „zwykłe” maszynowe działania, używając języka neuronauki!

Co przyniesie przyszłość? Jaka będzie siódma metafora? Według autora tej pracy będzie nią „mózg jako komputer kwantowy”, a stworzy ją nowa dziedzina informatyki, mianowicie obliczenia kwantowe (*quantum computing*). Po pierwsze, już dziś bardzo wyraźnie widać, że dziedzina ta przyniesie kompletną zmianę paradygmatu technologicznego. Każdy, kto zetknął się choćby krótko z tą dziedziną, wie, że nie polega ona na prostym przyspieszeniu zwykłych, klasycznych obliczeń. Programowanie kwantowe stanowi zupełnie nową jakość i nie jest sprowadzalne do znanych nam obecnie sposobów tworzenia algorytmów.

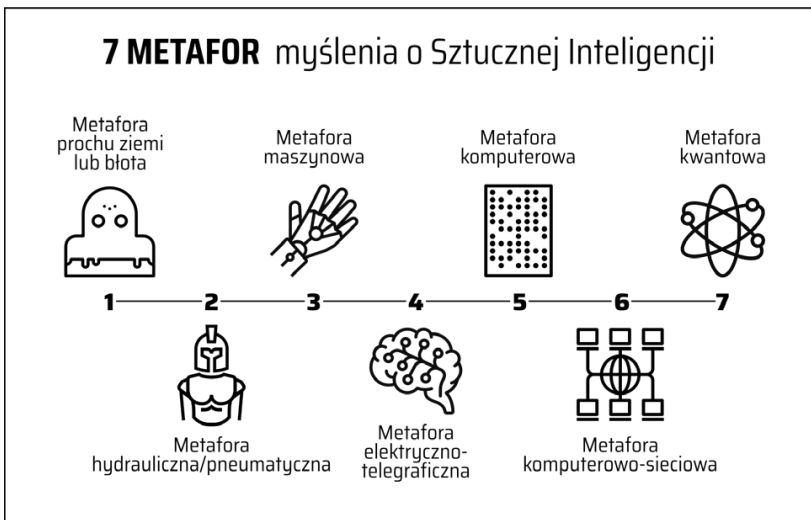
Są też sygnały, że rozumienie zjawisk mózgowych, a przede wszystkim świadomości, może zacząć korzystać z tej przyszłej metafory – „Siódmej Pieczęci”² technologii. Słynny fizyk, Roger Penrose, wraz z psychologiem i anestezjologiem Stuartem Hame-

² Metafora nawiązująca zarówno do znanego filmu Ingmara Bergmana, jak i do Księgi Apokalipsy jest zamierzona. Siódma Pieczęć reprezentuje w nich „milczenie niebios” – sytuację ostatecznej niewiedzy. Współcześnie nie mamy żadnej uznanej teorii tłumaczącej mechanikę kwantową – odwołanie się zatem do niej w wyjaśnianiu ludzkiej świadomości jest naznaczone ostateczną niepewnością i niewiedzą.

roffem, stworzyli teorię zorkiestrowanej obiektywnej redukcji. Jest to biologiczna teoria umysłu, która zakłada, że świadomość powstaje na poziomie zjawisk kwantowych wewnątrz neuronów, a nie w wyniku połączeń między neuronami – czyli nie jako konketom. Analiza tej teorii i implikacji z niej płynących wykracza poza zakres niniejszego artykułu – przywołano ją tutaj, aby pokazać, jak silny jest i będzie w przyszłości *metaforyczny* związek między rozumieniem mózgu, inteligencji i w końcu – świadomości z aktualną technologią.

Rysunek 1 przedstawia siedem opisanych metafor w historii sztucznej inteligencji.

Rysunek 1. Siedem metafor w rozwoju AI



Źródło: Zbigniew Zieliński (MakoLab SA).

Czego uczy historyczna analiza metafor sztucznej inteligencji?

Przede wszystkim... pokory. W każdym z okresów rozwoju ludzkości naukowcom i filozofom spekulującym na temat natury inteligencji wydawało się niewątpliwym to, iż jest ona realizo-

wana środkami, jakie były im znane. Wszystko jedno, czy tymi środkami były zjawiska hydrauliczne, mechaniczne, elektryczne czy cyfrowe – w każdej z tych epok było pozornie oczywiste, że inteligencja powstaje z tych właśnie środków. Ale wraz z rozwojem nauki i zmianą paradygmatów technologicznych okazywało się, że środki poprzedniej epoki są przestarzałe i po prostu nieskuteczne.

Tak jest i dziś – wyjaśnianie działania mózgu poprzez konektomy, czy też sieci neuronowe, rozumiane tak, jak dziś w metodach maszynowego uczenia, zapewne okaże się prymitywne, gdy będziemy już po zmianie paradygmatu. I o ile dziś, w roku 2021, wiemy już, że ta zmiana przyjdzie ze strony komputerów kwantowych i zapewne bardzo pomoże wyjaśnić naturę inteligencji – możemy być też pewni, że o ile ludzkość nie dokona samozagłady, ta zmiana nie będzie ostatnią.

W tym kontekście należy wskazać, paradoksalnie jako optymistyczne, zjawisko, jakie można w ostatnich latach zaobserwować, polegające na rosnącej świadomości fundamentalnych ograniczeń sztucznej inteligencji rozwijanej w ramach obecnego (szóstego) paradygmatu. Spośród wielu przykładów tego zjawiska warto wskazać znakomitą pracę *Rebooting AI – Building Artificial Intelligence We Can Trust* Gary’ego Marcusa i Ernesta Davaisa (2019). Autorzy, bardzo znane osoby w świecie AI, przyznają:

Pomimo że danych jest coraz więcej, a klastry komputerów stają się coraz szybsze, a inwestycje coraz większe, ważne jest, aby zdać sobie sprawę, że wciąż brakuje czegoś fundamentalnego. Mimo całego tego postępu, pod wieloma względami maszyny nadal nie mogą się równać z ludźmi (s. 10).

Oczywiście bardzo istotne jest pytanie, czym jest to coś, co jest fundamentalne?

Czy sztuczna inteligencja może „rozumieć” metafory?

Niezwykle interesujący i stosunkowo mało znany jest fakt, iż świadomość znaczenia metafor towarzyszyła rozwojowi systemów sztucznej inteligencji od początków symbolicznej sztucznej inteligencji. Kiedy w 2001 roku znane wśród uczonych zajmujących się metaforami czasopismo *Metaphor and Symbol* poświęciło zagadnieniu metafor w sztucznej inteligencji całe specjalne wydanie, stało się jasne, iż mają one duże znaczenie. I dokładnie taki właśnie był też tytuł pracy wybitnego specjalisty z tej dziedziny, Johna Barndena (*Metaphor and Artificial Intelligence: Why They Matter to Each Other*) (2008) opublikowanej w monografii *The Cambridge Handbook of Metaphor and Thought* (Gibbs, 2008).

Historia systemów symbolicznej sztucznej inteligencji zdolnych do interpretacji i generowania metafor sięga wczesnych lat dziewięćdziesiątych XX wieku. W University of California w Berkeley powstał w 1990 roku system MIDAS (Metaphor Interpretation, Denotation, and Acquisition System), stworzony przez Jamesa Martina – był częścią systemu Unix Consultant System. System ten wyposażony był w zbiór wbudowanych metafor, które John Barnden opisał jako „z grubsza Lakoffskie metafory pojęciowe” (Barnden, 2008, s. 9). I tak np. MIDAS rozumiał wyrażenia takie jak „Jestem w programie Emacs” albo pytania, jak np. „Jak mam wejść do programu VI?”, które są oczywiście całkowicie metaforyczne. Było to możliwe dzięki wyposażeniu MIDAS w podsystem MES (Metaphor Extension System), który rozumiał „mapowanie” o postaci np.: „być-w-miejscu \leftrightarrow używać-procesu-systemu”. Martin jest też autorem pracy *A Computational Model of Metaphor Interpretation* (1990), w której zawarł idee rozwinięte później w jego przetwarzaniu języka naturalnego *Speech and Language Processing* (Jurafsky, Martin, 2008).

W instytucie SRI (Stanford Research Institute) przy University of Stanford, w latach 1991–1993 Jerry Hobbs wdrożył wyniki swoich badań nad metaforami i sztuczną inteligencją w systemie TACITUS (Hobbs, 1989). System oparty był na trzech węzłowych

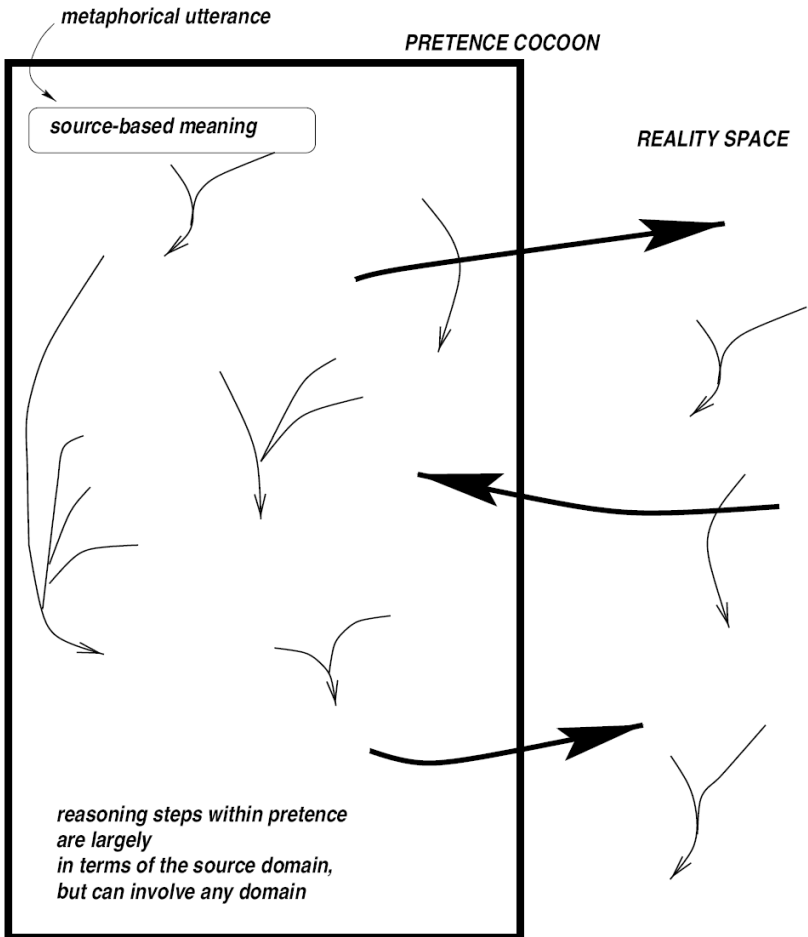
metodach wynikających z podejścia do języka naturalnego, które uwzględniało metafory, a nawet pokrewne im metonimie. Metody te: transfer niemodyfikowanej własności, transfer za pomocą odwzorowania z wnioskowaniem oraz odkrywanie odwzorowań za pomocą analogii uczyniły system TACITUS jednym z pierwszych systemów stosujących rozumowanie abdukcyjne w sztucznej inteligencji.

W latach dziewięćdziesiątych powstaje w Simon Fraser University (British Columbia, Canada) bardzo zaawansowane oprogramowanie Meta5 (Fass, 1997). Osiągnięciem jego autora, Dana Fassa, było automatyczne wykrywanie analogii pomiędzy dwoma domenami konceptualnymi, powiązanymi wyrażeniami metaforycznymi. Przykładem takiego działania jest zdolność do odkrycia metaforycznego sensu zdania: „My car drinks gasoline”. System stosował w tym celu wbudowaną w niego wiedzę mówiącą, że z jednej strony zwierzęta piją płyny nadające się do picia, a z drugiej strony, że samochody konsumują („piją”) benzynę. Meta5 potrafił w tym wypadku „odkryć”, że czasownik *to drink* może mieć poprawne znaczenie tożsame z czasownikiem *to use*.

Bardzo interesującą cechą Meta5 była zdolność do wykrywania nie tylko metafor, ale również metonimii. Idee zawarte w Meta5 pomogły stworzyć Iversonowi i Helmreichowi (1992) jego udoskonaloną wersję Metallel, która obsługiwała zarówno metonimie, jak i metafory.

Prawdopodobnie jednym z najdoskonalszych programów interpretujących metafory był system ATT-Meta, stworzony przez Johna Barndena z University of Birmingham (Lee, Barnden, 2001). ATT-Meta był de facto prekursorem semantycznych reasonerów (algorytmów zdolnych do maszynowego rozumowania). Nie dokonywał on analizy języka naturalnego, ale opierał się na logicznych zdaniach, w jakich zakodowane były znaczenia wyrażen języka naturalnego. Procesowanie metafor w ATT-Meta odbywało się dzięki zaimplementowaniu interesującego rozwiązania, zwanego kokonem pozorów (*pretence cocoon*) (Barnden, 2001). Rolę kokonu pozorów w programie ATT-Meta ilustruje rys. 2.

Rysunek 2. Ogólne podejście do wnioskowania w systemie ATT-Meta



Źródło: Barnden (2001, s. 17).

Interesującą cechą ATT-Meta było uwzględnienie przy projektowaniu, a następnie przy testowaniu programu, metafor pochodzących z popularnej pracy Andrew Goatly'ego *The Language of Metaphors* (1997), będącej jedną z najważniejszych współczesnych prac dotyczących metafor. Dzięki takiemu podejściu ATT-Meta posiadał interesującą cechę możliwości wkraczania poza proste odwzorowania domeny źródłowej i docelowej. Autor ATT-Meta twierdził, że pozwalało to użytkownikom na rozszerzenie rozu-

mienia metafory jako instrumentu przenoszącego od domeny pozorów do rzeczywistości (Barnden, 2001, s. 9).

Na zakończenie przeglądu dokonań symbolicznej sztucznej inteligencji w wykrywaniu i generowaniu metafor warto jeszcze wspomnieć dwa znaczące projekty.

Pierwszy z nich to KARMA (Knowledge-based Action Representations for Metaphor and Aspect), którego autorem jest Srini Narayanan z University of California w Berkeley. System KARMA był stosowany do procesowania tekstów o tematyce ekonomicznej pochodzących z artykułów prasowych i posiadał duże możliwości wykrywania subtelnych, metaforycznych aspektów tych tekstów (Narayanan, 1997).

Drugi to program SAPPER, stworzony w 1992 roku przez Tony'ego Veale'a (1998) z Dublin City University. SAPPER implementował hybrydowy model (łączył metody symboliczne i metody koneksjonizmu) do znajdowania analogii strukturalnych w analizowanych tekstach. Implementacja programu opierała się na analizie semantycznej struktury sieciowej, w której węzły reprezentują pojęcia.

Opisane tutaj dokonania warto podsumować, przywołując niezwykle interesującą analizę Johna Barndena, ukazującą znaczący wkład opisanych powyżej badań w domenę sztucznej inteligencji na czysto teoretyczne badania metafor. Barnden wylicza kilkanaście elementów tego wkładu, z których warto przytoczyć analizy pospolitości (*mundaneness*), analizy metafor nieoznajmających, np. metafor zawartych w pytaniach (*non-assertional metaphor*), pogłębioną analizę mapowania pomiędzy domenami (*details of mapping*) czy wreszcie szczegóły związku pomiędzy metaforami a metonimiami.

Przedstawiona tu analiza ukazuje, iż w obrębie symbolicznej sztucznej inteligencji znajdujemy jawne i bardzo świadome użycie metafor, jak również tworzenie metod ich wykrywania i generowania. Według opinii niektórych badaczy prace w tym zakresie przyczyniły się także do pogłębienia ogólnego rozumienia metafor.

Współczesna sztuczna inteligencja i jej metafory

Dla autora niniejszej pracy dużym zaskoczeniem było odkrycie, iż tematyka metafor jest bardzo słabo reprezentowana we współczesnej (koneksjonistycznej) sztucznej inteligencji, w której powszechne jest stosowanie metod sieci neuronowych i opartych na nich metod maszynowego uczenia, w tym metod uczenia głębokiego (*deep learning*). Co więcej, temat metafor jest niemal nieobecny w słynnej Biblii AI, pracy *Artificial Intelligence. A Modern Approach* (Russell, Norvig, *Artificial Intelligence*). Pierwsze wystąpienie terminu „metafora” znajdujemy dopiero na stronie 906 (z 1150 stron) i w dodatku występuje błąd definiujący metaforę jako rodzaj metonimii. W tej ważnej pracy nie znajdziemy informacji o systemach takich jak opisane wcześniej ATT-Meta, Meta5, TACITUS, SAPPER, MIDAS, KARMA. Temat metafor ignoruje też słynna pozycja Nilsa J. Nilssona *The Quest for Artificial Intelligence* (2009). Co więcej, zagadnienie metafor nie pojawia się nawet w pracy *The Deep Learning Revolution – Artificial Intelligence Meets Human Intelligence* autorstwa Terrence’a J. Sejnowskiego (poza metaforą *pandemonium* niemającą jednak w niej istotnego znaczenia poza odwołaniem historycznym).

Dopiero w ostatnich latach, głównie przy wsparciu organizacji RaAM (International Association for Researching and Applying Metaphor), organizowane są warsztaty „On Metaphor in NLP” (ich obecna nazwa to „Workshops on Figurative Language Processing”). W 2018 roku podczas warsztatów w New Orleans wypracowano „wspólne zadania dotyczące wykrywania metafor na poziomie tokenów” („shared task on token-level metaphor detection”), które przyczyniły się do większego zainteresowania metaforami w pracach stosujących metody *machine learning* (w tym *deep learning*). Warto w tym kontekście wskazać na pracę K. Swarnkara i A.K. Singha *Di-LSTM Contrast. A Deep Neural Network for Metaphor Detection* (<https://www.aclweb.org/anthology/W18-0914.pdf>) oraz pracę z ubiegłorocznej, 34. konferencji „AAAI Conference on Artificial Intelligence”, pod znamienym tytułem *Multi-Task Learning for Metaphor Detection with Graph Convolutional Neural Networks and Word Sense Disambiguation*

(Le, Thai, Nguyen, 2020). Podstawy zastosowania sieci neuronalnych do wykrywania metafor zostały opracowane m.in. przez zespół naukowców z Allen Institute for Artificial Intelligence oraz University of Washington w pracy *Neural Metaphor Detection in Context* (Gao, Choi, Zettlemoyer, 2018).

Należy mieć nadzieję, że najbliższe lata przyniosą dynamiczny rozwój zastosowań metod uczenia maszynowego do wykrywania i procesowania metafor. Nie ma jednak wątpliwości, że stopień zainteresowania tematem metafor daleki jest jeszcze od tego, jaki obserwowaliśmy w Dobrej Staromodnej Sztucznej Inteligencji.

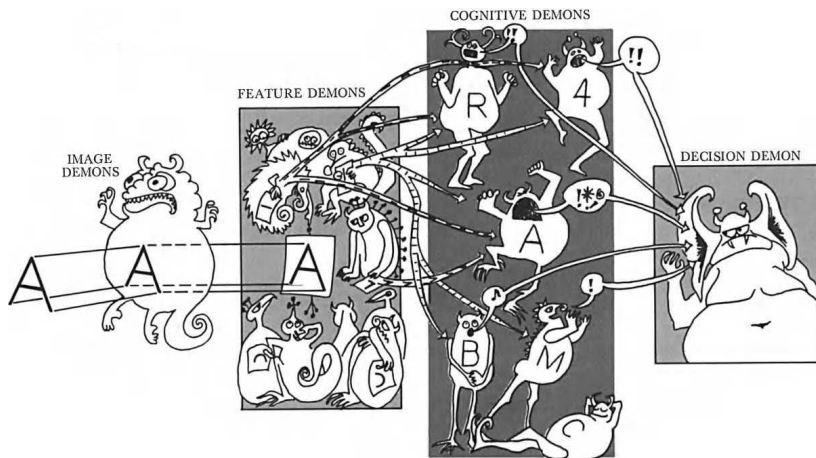
Nieuświadomiona obecność metafor we współczesnej sztucznej inteligencji

Paradoksalnie metaforyczny sposób opisu algorytmów i specyficznych rozwiązań w sztucznej inteligencji jest bardzo powszechny. To, że przetwarzanie metafor nie zadomowiło się w niej jeszcze na dobre, nie oznacza, że nie są one obecne. Niestety bardzo wielu adeptów SI posługuje się nimi nieświadomie, co wywołuje często problemy we właściwym rozumieniu tak możliwości, jak i słabości tej domeny nauki i techniki.

Pojęcia, z jakimi spotykamy się często, takie jak: *pandemonium*, *perceptron*, *brain-style computing*, *neocognitron*, *computer vision*, *moral machine* – są oczywiście silnie metaforyczne, a ich źródłowymi domenami są najczęściej neuronauka, psychologia, a nawet etyka. Przyjrzyjmy się zatem nieco bliżej tym metaforom.

Pandemonium – opisując jeszcze w 1959 roku (a więc w czasach GOFAI!), jak mogłoby w przyszłości działać rozpoznawanie obrazów, Oliver Selfridge stworzył koncepcję *pandemonium*, metaforyczne urządzenie do rozpoznawania kształtów, w którym „demony” wykrywające pewne cechy obrazu rywalizowały ze sobą o prawo do reprezentacji obiektów w obrazach, co ilustruje rys. 3.

Rysunek 3. Metaforyczne pandemonium Selfridge'a – metafora przyszłych metod deep learning



Źródło: Lindsay Norman (1977). Pobrane z: Wikipedia Commons, <https://commons.wikimedia.org/wiki/File:Pande.jpg> (dostęp: 5.07.2021).

Metaforyczne demony Selfridge'a to demony obrazu, demony cech, demony kognitywne i demony decyzyjne. Jest oczywiste, że w roku 1959 ta wizualizacja mogła być wyłącznie metaforyczna, tym niemniej zadziwiające jest, jak bardzo dobrze ilustrowała różne funkcje warstw, jakie znajdujemy dziś w metodach *deep learning*.

Perceptron – w metodach *machine learning* perceptron jest stosunkowo prostym algorytmem do nadzorowanego uczenia tzw. klasyfikatorów binarnych. Ale w opinii autorów fundamentalnej pracy *Perceptrons*, Marvin'a Minsky'ego i Seymour'a Paperta (1969), perceptron to uproszczony model biologicznego neuronu. Ponieważ złożoność biologicznych neuronów jest ogromna, a jej rozumienie jest warunkiem koniecznym badania zachowania układu nerwowego, mówienie o perceptronie jako o „modelu biologicznego neuronu” ma wysoce metaforyczny charakter.

Brain-style computing to bardzo silna metafora dla pewnej klasy algorytmów z dziedziny neuronauki obliczeniowej (*computational neuroscience*). Oczywiście, pomimo istnienia pewnych analogii pomiędzy strukturami układu nerwowego a złożonymi algorytmami obliczeniowymi, mówienie tutaj o „stylu mózgowym” ma silnie metaforyczny charakter. Jak przyznaje słynny Terry Sejnowski:

„Co kilka lat dostaję telefon od Geoffreya (Hinton), który zaczyna się od «Już wiem, jak działa mózg». Za każdym razem opowiada mi o jakimś sprytnym nowym schemacie ulepszania modeli sieci neuronowych” (Sejnowski, 2018, s. 49). Oczywiście w tym wypadku metafora „Już wiem, jak działa mózg” rzeczywiście przyczynia się do postępu w konstrukcji algorytmów uczenia maszynowego. Ale to są nadal algorytmy komputerowe. Przypisywanie im stylu mózgowego ma wyłącznie metaforyczny charakter!

Neocognitron to wielowarstwowy model sieci neuronowej wynaleziony w 1979 roku przez Kunihiko Fukushima z Osaka University. Model ten oparty był na pewnym wyobrażeniu o architekturze systemu wizualnego, wykorzystywał filtry konwolucyjne i proste modele plastyczności sieci. Był niewątpliwie bezpośrednim prekursorem metod *convolutional neural network*. Jest jednak jasne, że tak jak i inne algorytmy w tej dziedzinie, model ten jest daleki od reprezentacji poznania (*cognition*), będącego udziałem człowieka, a zatem należy uznać, że sama nazwa modelu, jak i wiele odniesień do sieci neuronów mają w nim metaforyczny charakter.

Computer vision jest dobrze zdefiniowaną interdyscyplinarną dziedziną naukową zajmującą się zagadnieniem wysokopoziomowego rozumienia (rozpoznawania) obrazów i filmów przez algorytmy. Język, jakim o tej dziedzinie mówią jej rzecznicy, jest jednak wysoce metaforyczny i pełen odniesień do ludzkiego widzenia. Raz jeszcze oddajmy głos Terry’emu Sejnowskiemu: „W latach sześćdziesiątych mało kto podejrzewał, że potrzeba pięćdziesięciu lat i milionkrotnego wzrostu mocy komputerów, zanim «Computer Vision» osiągnie poziom funkcjonowania właściwy ludziom”. Oczywiście, nie zaprzeczając wybitnym osiągnięciom tej dziedziny, mówienie o „poziomie funkcjonowania właściwym ludziom” („human levels of performance”) ma szczególnie metaforyczny charakter – przede wszystkim dlatego, że przestrzeń, jaka dzieli nawet najdoskonalsze algorytmy od funkcjonowania narządów wzroku oraz części układu nerwowego odpowiedzialne za widzenie, jest ogromna i nadal nie mamy pojęcia, jak te dwie rzeczywistości przybliżyć.

Moral machine to inicjatywa znanego MIT Media Lab (<https://www.media.mit.edu/>), będąca platformą dyskusji dotyczącej ludz-

kiego spojrzenia na decyzje moralne podejmowane przez *maszynową inteligencję*, taką w jaką np. wyposażone będą pojazdy autonomiczne. Badania raportowane przez tę platformę są niezwykle interesujące i ważne. Tym niemniej należy zauważyć, że język, jakim posługują się autorzy raportowanych tam prac, jest wysoce metaforyczny (np. mówi się o „decyzjach etycznych”, „eksperymentach moralnych”, „psychologicznych aspektach algorytmów” itp.). Według autora niniejszej pracy używanie tych metafor samo w sobie nie jest ryzykowne, o ile istnieje świadomość metaforyczności tych pojęć w odniesieniu do maszynowych algorytmów oraz świadomość, że tak naprawdę z bardzo fundamentalnych powodów nie istnieje coś takiego jak moralna maszyna!

Podsumowanie

Metafory i metaforyczne myślenie są głęboko zakorzenione w teorii i praktyce sztucznej inteligencji. O ile język *praxis* sztucznej inteligencji (algorytmy i metody postępowania) jest w przeważającej większości literalny, o tyle jej *theoria* jest bardzo metaforyczna – ale często ta metaforyczność nie jest uświadomiona przez jej adeptów.

Możemy tu wskazać na następujące węzłowe problemy.

Po pierwsze, samo myślenie o naturze inteligencji, a więc i jej sztucznej odmianie, zależy od metafor wynikających z aktualnych paradygmatów w nauce i technologii. Analiza historyczna, jaką zaprezentowaliśmy w pierwszej części pracy, ukazuje z całą mocą, iż jest poznawczo niebezpieczne absolutyzowanie aktualnego paradygmatu i próba wyjaśniania w jego obrębie natury inteligencji, albowiem każdy kolejny paradygmat pokazywał poważne braki, a często nawet zupełną naiwność wyjaśnień dokonywanych w ramach poprzednich perspektyw.

Po drugie, w dydaktyce, zarówno uniwersyteckiej, jak i uprawianej współcześnie na platformach MOOC (Massive Open Online Courses), tematyka metafor w sztucznej inteligencji jest niemal nieobecna. Jak to wykazaliśmy, w obrębie symbolicznej (tradycyjnej) sztucznej inteligencji mieliśmy do czynienia z wielkimi

osiągnięciami tak w rozpoznawaniu, jak i generowaniu metafor. Pomimo że współczesna SI jest w tym obszarze zapewne dopiero w początkowej fazie – znaczenie prac opisujących metody przetwarzania metafor jest dla niej wręcz fundamentalne.

Po trzecie, istnieje w domenie sztucznej inteligencji nie tylko wiele silnie metaforycznych pojęć, ale także w powszechnym dziś popularnonaukowym dyskursie o jej metodach, zastosowaniach oraz wpływie na społeczeństwa i gospodarki świata obecne jest silnie metaforyczne myślenie przy braku wystarczającej świadomości.

Wydaje się zatem, że niezbędne są daleko idące zmiany w dydaktyce SI. Po pierwsze, konieczne jest ukazanie, jak mocna jest zależność rozumienia sztucznej inteligencji od aktualnego paradygmatu w nauce i technologii. Po drugie, niezbędne jest ukazanie adeptom sztucznej inteligencji, jakie miejsce w historii tej dziedziny ludzkiej wiedzy odgrywają metafory – tak aby były one zauważone, w pełni uświadomione i właściwie rozumiane.

Bibliografia

- Barnden, J. (2001). *Application of the ATT-Meta Metaphor-Understanding Approach to Selected Examples from Goatly*. Technical Report CSRP-01-01. Birmingham: University of Birmingham.
- Barnden, J.A. (2008). Metaphor and Artificial Intelligence. Why They Matter to Each Other. W: R.W. Gibbs, Jr. (red.). *The Cambridge Handbook of Metaphor and Thought*, s. 311–338. DOI: 10.1017/CBO9780511816802.020.
- Boyd, R. (1993). Metaphor and Theory Change. What Is „Metaphor” A Metaphor For? W: A. Ortony. *Metaphor and Thought*, ed. 2 (s. 485). Cambridge: Cambridge University Press.
- Descartes, R. (1981). *Rozprawa o metodzie*. Warszawa: PWN, s. 29.
- Dyson, G. (2012). *Turing’s Cathedral. The Origins of the Digital Universe*. New York: Penguin.
- Fass, D. (1997). *Processing Metonymy and Metaphor*. Greenwich: Ablex Publishing.

- Gao, G., Choi, E., Choi, Y., Zettlemoyer, L. (2018). Neural Metaphor Detection in Context, EMNLP. Pobrane z: <https://arxiv.org/abs/1808.09653> (dostęp: 5.07.2021).
- Gibbs Jr., R.W. (2008). *The Cambridge Handbook of Metaphor and Thought (Cambridge Handbooks in Psychology)*. Cambridge: Cambridge University Press.
- Goatly, A. (1997). *The Language of Metaphors*. London–New York: Routledge.
- Hobbs, J., Appelt, D., Bear, J., Stickel, M., Tyson, M. (1989) *TACITUS: A Message Understanding System*. SRI International Menlo Park, California. Pobrane z: <https://www.aclweb.org/anthology/H89-2075.pdf> (dostęp: 5.07.2021).
- Hoffman, D.D, Singh, M., Prakash, Ch. (2015). The Interface Theory of Perception. *Psychon. Bull Rev*, 22, s. 1480–1506.
- Knuth, D.E. (2004). *The Art of Computer Programming. A Draft of section 7.2.1.7. History of Combinatorial Generation*. Boston: Addison-Wesley. Pobrane z: <http://www.antiquark.com/blogging/fasc4b.pdf> (dostęp: 5.07.2021).
- Iverson, E., Helmreich, S. (2007). Metallel. An Integrated Approach to Non-Literal Phrase Interpretation. *Computational Intelligence*, 8(3), s. 477–493. DOI: 10.1111/j.1467-8640.1992.tb00376.x.
- Jurafsky, D., Martin, J. (2008). *Speech and Language Processing*, ed. 2. Hoboken: Prentice Hall.
- Kuhn, T. (1993). Metaphor in Science. W: A. Ortony. *Metaphor and Thought*, ed. 2 (s. 538). Cambridge: Cambridge University Press.
- Lakoff, G. (1993). The Contemporary Theory of Metaphor. W: A. Ortony, *Metaphor and Thought* (s. 202). Cambridge: Cambridge University Press.
- Le, D.M., Thai, M., Nguyen, T.H. (2020). Multi-Task Learning for Metaphor Detection with Graph Convolutional Neural Networks and Word Sense Disambiguation. The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI-20).
- Lee, M.G., Barnden, J.A. (2001). Reasoning About Mixed Metaphors Within an Implemented AI System. *Metaphor and Symbol*, 16(1–2), s. 29–42.
- Lindsay, P.H., Norman, D.A. (1977). *Human Information Processing: An Introduction to Psychology*. New York: Academic Press.

- Marcus, G., Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. New York: Vintage Books.
- Martin, J.H. (1990). *A Computational Model of Metaphor Interpretation (Perspectives in Artificial Intelligence)*, ed. 1. San Diego: Academic Press.
- Minsky, M.L., Papert, S.A. (1969). *Perceptrons*, Cambridge: The MIT Press.
- Narayanan, S. (1997). *Knowledgebased Action Representations for Metaphor and Aspect KARMA*. University of California at Berkeley. Pobrane z: <http://www1.icsi.berkeley.edu/~snarayan/thesis.pdf> (dostęp: 5.07.2021).
- Nilsson, N.J. (2009). *The Quest for Artificial Intelligence*. Cambridge: Cambridge University Press, s. 20.
- Ortony, A. (1993). Metaphor, Language, and Thought. W: *Idem. Metaphor and Thought*, ed. 2 (s. 14). Cambridge: Cambridge University Press.
- Otis, L. (2002). The Metaphoric Circuit: Organic and Technological Communication in the Nineteenth Century. *Journal of the History of Ideas*. DOI: 10.2307/3654260. Tłum. M. Sopek.
- Petrie, H.G., Oshlag, R.S. (1993). Metaphor and Learning. W: A. Ortony, *Metaphor and Thought*, ed. 2 (s. 15). Cambridge: Cambridge University Press.
- Prakash, Ch., Stephens, K.D., Hoffman, D.D., Singh, M., Fields, Ch. (2020). Fitness Beats Truth in the Evolution of Perception. *Acta Biotheor*. DOI: 10.1007/s10441-020-09400-0.
- Russell, S., Norvig, P. *Artificial Intelligence: A Modern Approach*. Pobrane z: <http://aima.cs.berkeley.edu/> (dostęp: 7.07.2021).
- Sejnowski, T.J. (2018). *The Deep Learning Revolution*. Cambridge: The MIT Press, s. 49.
- Seung, S. *I Am Connectome*. TEDGlobal 2010. Pobrane z: https://www.ted.com/talks/sebastian_seung_i_am_my_connectome (dostęp: 5.07.2021).
- Sopek, M. *Naszą prywatność ocali mechanika*, 30.01.2021. Pobrane z: <https://www.sztucznainteligencja.org.pl/tag/siec-pelna-znaczen/> (dostęp: 5.07.2021).

- Veale, T. „Just in Time”. *Analogical Mapping, an Iterative-Deepening Approach to Structure-Mapping*. The Proceedings of ECAI '98, the 13th European Conference on Artificial Intelligence, Brighton, UK. von Helmholtz, H. (1995). *Science and Culture. Popular and Philosophical Essay*. Chicago: University of Chicago Press, s. 150.
- Zarkadakis, G. (2015). *In Our Own Image. Will Artificial Intelligence Save or Destroy Us*. New York: Random House.
- Zykubek A. *Od sieci pełnej znaczeń do sztucznej inteligencji – Mirostaw Sopek*. *Galerie zdjęć i nagranie wideo*, 24.II.2016. Pobrane z: <https://kognitywistyka.kul.pl/sieci-pelnej-znaczen-sztucznej-inteligencji-miroslaw-sopek-galerie-zdjec-nagranie-wideo/> (dostęp: 5.07.2021).