

FILOZOFIA I NAUKA  
Studia filozoficzne i interdyscyplinarne  
Tom 8, cz. 1, 2020

Jacek Koronacki

## SZTUCZNA INTELIGENCJA W ODCZAROWANYM ŚWIECIE

<https://doi.org/10.37240/FiN.2020.8.1.1>

### STRESZCZENIE

Niniejsze rozważanie jest pisane przez inżyniera. W pierwszych dwóch punktach artykułu znajdujemy narysowany kilkoma kreskami szkic metodologicznych podstaw sztucznej inteligencji (SI) i czym dziś SI jest. W dalszych punktach zasygnalizujemy kształt najbliższej przyszłości SI, umieścimy SI w kontekście kultury, odnotujemy fenomen tzw. silnej sztucznej inteligencji i zakończymy całość paroma uwagami.

**Słowa kluczowe:** sztuczna inteligencja, silna sztuczna inteligencja, uczenie maszynowe, świat odczarowany.

### NOWY WSPANIAŁY ŚWIAT MAXA WEBERA?

Słynny monachijski wykład Maxa Webera z roku 1917, opublikowany dwa lata później, zawierał takie stwierdzenia:

„Wzrastająca intelektualizacja i racjonalizacja nie oznacza [...] wzrostu powszechnej wiedzy o warunkach życiowych, którym podlegamy. Oznacza ona coś innego: wiedzę o tym, albo wiarę w to, że gdyby tylko człowiek tego *chciał, to mógłby* w każdej chwili przekonać się, że nie ma żadnych tajemniczych, nieobliczalnych mocy, które by w naszym życiu *odgrywały jakąś rolę*, ale wszystkie rzeczy można – w zasadzie – *opanować przez kalkulację*. Oznacza to odczarowanie świata. [podkreślenia autora]

[...] Wszystkie nauki przyrodnicze dają nam odpowiedź na pytanie, co powinniśmy zrobić, aby opanować życie w sposób techniczny. Nie dają one jednak odpowiedzi na pytanie [...]: czy naprawdę chcemy i czy powinniśmy opanowywać życie w ten sposób oraz czy ma to w ogóle jakiś sens?”<sup>1</sup>

Opisany przez Webera proces opanowywania świata przez racjonalizację instrumentalną, zwaną też celowościową (*Zweckrationalität*) lub kalkula-

<sup>1</sup> *Wissenschaft als Beruf* (tłum. polskie *Nauka jako zawód i powołanie*), w: idem, *Polityka jako zawód i powołanie*, przeł. A. Kopacki, P. Dybel, Wyd. Znak 1998.

cyjną ma być kontynuowany w dużej mierze dzięki rozwojowi sztucznej inteligencji (dalej skrótem dla terminu „sztuczna inteligencja” będzie SI albo z angielskiego AI). Opracowana przez międzyresortowy zespół analityczny w roku 2019 i przedstawiona do konsultacji społecznych „Polityka rozwoju sztucznej inteligencji w Polsce (na lata 2020–2027)” zawiera w swojej części wstępnej taką diagnozę:

*„Sztuczna inteligencja jest kolejną – po wynalezieniu druku, broni palnej, maszyny parowej, silnika spalinowym, elektryczności, radiu czy komputerach – technologią przełomową o potencjale zaburzającym (disruptive), która może gwałtownie zmienić świat, w którym żyjemy. Technologie przełomowe są podstawą doniosłych innowacji, które mają naturę kreatywnej destrukcji. Innymi słowy, ich zastosowanie w dłuższym okresie przynosi duży wzrost gospodarczy i dobrobytu ludzi, czy potęgi państw lub bloków międzynarodowych, ale jednocześnie wiąże się z masowym zanikiem dotychczas znanych aktywności gospodarczych lub istniejących przewag politycznych [podkreślenia autorów].”<sup>2</sup>*

Z kolei pierwsze zdania opracowania Fundacji digitalpoland pt. *Strategia rozwoju sztucznej inteligencji w Polsce z roku 2018* brzmią:

*„Stoimy u progu rewolucji, która zmieni nasz świat nie do poznania. Trzy wcześniejsze rewolucje przemysłowe wyrzuciły produkcję do góry nogami, ale ta czwarta, dotycząca sztucznej inteligencji przewartościuje życie każdego z nas. Wywoła wstrząs, który będzie większy niż ten spowodowany rewolucją przemysłową czy upowszechnieniem elektryczności. Ekspert nie mają wątpliwości, że AI zmieni nasz świat nie do poznania.”<sup>3</sup>*

Nieco dalej możemy tamże przeczytać:

*„Państwo Środka do 2030 roku chce być technologicznym liderem w obszarze AI, a chiński rynek sztucznej inteligencji ma być wart blisko 150 miliardów dolarów. A może bliżej nam jednak do strategii Francji? Rząd Emmanuela Macrona, do 2022 roku planuje przeznaczyć 1,5 miliarda euro na rozwój tej technologii i wsparcie nauki. [...] Z kolei rząd USA na nieujawnione B+R w obszarze AI wydał 1,2 miliarda dolarów w 2016 roku, 30 miliardów dolarów sięgnęły inwestycje sektora prywatnego. [...] wydatki federalne wyniosły 2 miliardy dolarów.”*

W stosunku do swojego potencjału ekonomicznego ogromne pieniądze na rozwój SI chcą wydawać wszystkie kraje mające ambicję unowocześniania

<sup>2</sup> Materiał opracowany przez międzyresortowy zespół analityczno-redakcyjny, ustanowiony na podstawie memorandum zawartego w dniu 26 lutego 2019 r. przez Ministra Cyfryzacji, Ministra Przedsiębiorczości i Technologii, Ministra Nauki i Szkolnictwa Wyższego oraz Ministra Inwestycji i Rozwoju.

<sup>3</sup> Zob. <https://www.digitalpoland.org/assets/publications/przeegl%C4%85d-strategii-rozwoju-sztucznej-inteligencji-na-swiecie/przeglad-strategii-rozwoju-ai-digitalpoland-report.pdf>

swoich gospodarek. Jeśli więc SI ma odgrywać tak istotną rolę w opisanym przez Webera procesie, to przynajmniej pobieżnie należy przyjrzeć się, co ów rozwój ma przynieść i zapytać *czy ma to w ogóle jakiś sens*.

Autorem niniejszych rozważań jest inżynier, a ściślej przedstawiciel nauk ścisłych i technicznych, którego zainteresowanie układami automatycznego sterowania, modelowaniem matematycznym i statystyką matematyczną zaprowadziło wiele już lat temu do badań nad metodami uczenia maszynowego, o których dalej będzie jeszcze mowa. W kolejnych dwóch punktach artykułu znajdziemy narysowany ledwie kilkoma kreskami szkic źródeł SI i czym dziś SI jest. W dalszych punktach zasygnalizujemy kształt najbliższej przyszłości SI, umieścimy SI w kontekście kultury, odnotujemy fenomen tzw. silnej sztucznej inteligencji i zakończymy całość paroma uwagami.

## SZTUCZNA INTELIGENCJA – SKĄD JEST I W CZYM JEJ SIŁA

Pojęcie sztucznej inteligencji tak się zadomowiło w świadomości społecznej, że właściwie nie mu już potrzeby jej ścisłego definiowania, tym bardziej, że nie ma w tym przypadku definicji jednej i powszechnie przyjętej. Przypomnijmy jednak jej operacyjne określenie zaproponowane przez jednego z pionierów tej dyscypliny, Nilsa J. Nilssona: „Sztuczna inteligencja jest efektem ludzkiego działania, zmierzającego do uczynienia maszyn inteligentnymi; zaś inteligencja jest tą jakością, która pozwala danemu bytowi działać właściwie i przewidująco w jego otoczeniu.”<sup>4</sup>

To operacyjne i ogólne opisanie samej inteligencji pasuje tak samo dobrze do inteligentnej maszyny jak i do człowieka. Tyle że to człowiek stwarza maszynę. Chcąc ją uczynić inteligentną, człowiek wyposaża maszynę w pamięć, w której zapisana zostaje informacja z przeszłości, np. wszystkie dotąd spisane dzieła oraz tysiące rozegranych w przeszłości gier w szachy czy pochodzącą z Chin grę go. Także w czasie bieżącym zapisywane są kolejne informacje docierające z otoczenia do maszyny. Maszyna musi być także zdolna uczyć się – musi być wyposażona w algorytmy pozwalające tak przetwarzać posiadaną przez nią informację, by osiąść zdolność wykonywania pewnych zadań, np. grupowania podobnych do siebie obiektów czy odróżniania obiektów jednej klasy od innych, czy też grania w szachy albo pisania powieści podobnych do wcześniej załadowanych do jej pamięci. Zdolność maszynowego uczenia się musi zawierać w sobie umiejętność doskonalenia własnego działania dzięki powiększaniu się zasobów posiadanej informacji, w tym na podstawie zdobywanego własnego doświadczenia, np. na podstawie rozgrywanych przez uczącą się maszynę gier w szachy, raz wygranych,

<sup>4</sup> N. J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements*, Cambridge University Press 2010.

innym razem przegranych. Dzisiejsze maszyny są np. zdolne do pisania nowych wierszy, zwłaszcza naśladowujących dzieła konkretnego poety. Zaczynają być zdolne generować własne, zgoła sensowne teksty. Grają dziś w szachy lepiej od najlepszego szachisty na świecie. I zaczynają być również zdolne do modyfikowania wbudowanych w nie algorytmów uczenia się.

Wszystko to jest zgoła normalnym i niemającym w sobie nic tajemniczego efektem naukowego postępu, choć w istocie zadziwiające jest jego tempo. Jest to wszystko owocem pracy naukowców i inżynierów zajmujących się od dekad dyscypliną o nazwie „maszynowe uczenie się”; dziś częściej mówi się o „uczeniu maszynowym”. Zbudowane przez nich maszyny wykonują takie same zadania jak człowiek, przy tym coraz częściej lepiej od człowieka. Podobnie rzecz się już ma lub wkrótce będzie miała z humanoidalnymi robotami.

Rzeczone uczenie maszynowe (w skrócie, z angielskiego ML) jest dziś technicznym filarem sztucznej inteligencji. W bliższej albo dalszej przyszłości kolejnymi filarami SI mają się stać genetyka, biologia molekularna i molekularna nanotechnologia. Algorytmy ML są dziś wbudowane w najróżniejsze urządzenia i stają się wszechobecne w naszym otoczeniu.

Marzenia o stworzeniu humanoidalnych robotów sięgają starożytnej Grecji. Oświecenie przyniosło wiarę w nieograniczony postęp naukowy i przynajmniej od początku doby nowoczesności także technologiczny. W swoim dziele z 1793 roku, *Enquiry Concerning Political Justice, and It's Influence on General Virtue and Happiness*, William Godwin przewidywał „całkowite wyeliminowanie ułomności naszej natury” oraz przedłużenie ludzkiego życia „poza wszelkie granice, jakie jesteśmy w stanie wyznaczyć”. Postęp miał uwolnić człowieka nie tylko od cierpienia i choroby, ale także gnuśności, melancholii, agresji i nienawiści.

Siedemdziesiąt lat później, w artykule *Darwin among the Machines* Samuel Butler pisał już o maszynach wyposażonych w „mechaniczne życie”, wierzył w ich ewolucję i osiągnięcie przez nie dominującej pozycji na Ziemi. Tak jak w onym czasie człowiek posługiwał się zwierzętami hodowlanymi, tak w przyszłości – przewidywał Butler – maszyny przejmą rolę człowieka, któremu pozostanie status podobny do statusu owych zwierząt. Zarazem człowiek miał pozostać spokojny o swój los: maszyny będą dlań dobre, ponieważ będą potrzebować jego usług. Butler wydawał się łączyć w jedno wiarę w ewolucję, która doprowadzi do zastąpienia człowieka przez inne stworzenia, jego zdaniem mechanicznej natury, oraz – zdaniem piszącego te słowa – prorocze spojrzenie na człowieka oddającego się we władanie maszyn.<sup>5</sup>

Jest faktem szeroko znanym, że narodziny SI jako dziedziny wiedzy sięgają roku 1930, gdy myśl o niej niejako musiała się nasunąć Alanowi Turin-

<sup>5</sup> S. Butler, *Darwin among the Machines*, Christchurch, New Zealand, 13 June, 1863; <http://www.gutenberg.org/ebooks/1906>

gowi, kiedy tworzył teorię obliczeń, i ku której prowadziły Johna von Neumanna jego badania. Obydwaj spotkali się w owym czasie w Princeton, by dyskutować nad podwalinami SI.<sup>6</sup>

W latach 40-tych dwudziestego wieku von Neumann dyskutował z Norbertem Wienerem o modelu neuronu. Ten ostatni stworzył wówczas na MIT zespół badań kognitywistycznych, między innymi z Warrenem McCullochem i Walterem Pittsem. Kolejne dwie dekady przyniosły bardziej złożone konstrukcje sztucznych sieci neuronowych (w skrócie ANN). Ówczesne sieci nie przyniosły jednak sukcesów, jakich się po nich spodziewano. Skądinąd wybitni wizjonerzy chcieli osiągnąć zbyt wiele zbyt małymi środkami – potrzeba było czasu na pogłębienie koncepcji oraz na pojawienie się komputerów pozwalających te koncepcje zaimplementować.

Pamiętamy, że termin SI wprowadził w roku 1955 wybitny informatyk John McCarthy, współpracujący m.in. z Allenem Newellem i Herbertem Simonem, którzy w tym samym czasie, z pomocą Johna C. Shawa, opracowali pierwszy komputerowy język programowania do dowodzenia twierdzeń, *Logic Theory Machine* (znany też pod nazwą *Logic Theorist*).<sup>7</sup> Program był zdolny dowodzić prawdziwości elementarnych twierdzeń rachunku zdań. Tak otworzona została metodologia programowania w logice, czyli pisania programów nie w postaci sekwencji kroków prowadzących do rozwiązania, lecz narzucania warunków logicznych, jakie końcowe rozwiązanie musi spełniać. Od tamtego czasu programowanie w logice jest słusznie uznawane za część SI.

Dalsze losy badań SI związały się z konstrukcją nowych paradygmatów uczenia się maszyn oraz komputerowej implementacji tych paradygmatów. Warto odnotować pionierski udział w tym okresie rozwoju SI osiadłych w USA Polaków, przede wszystkim Ryszarda Spencera Michalskiego (1937–2007, w USA od 1971<sup>8</sup>). Ich podejście wyrastało z logiki matematycznej i, jak to ujmował Michalski, za swoją oś miało *naturalną indukcję*.<sup>9</sup> Indukcja, albo lepiej uczenie (rozumowanie) indukcyjne mówi o procesie algorytmicznego generowania wiedzy z dostarczonych maszynie przykładów i zarazem na

<sup>6</sup> Tu należy przypomnieć zupełnie niezwykłą jak na ów czas, niedokończoną i wydaną rok po śmierci von Neumanna książeczkę *The Computer and the Brain*, New Haven 1958; na polski przełożył ją Klemens Szaniawski; została opublikowana pod tytułem *Maszyna i mózg ludzki*, PWN 1963.

<sup>7</sup> Newell był informatykiem i psychologiem kognitywistą, Simon – ekonomistą zainteresowanym teorią i praktyką zarządzania oraz kognitywistką, Shaw był informatykiem, znakomitym programistą systemowym.

<sup>8</sup> Ryszard Michalski, przez dekady profesor George Mason University, przybrał drugie imię po przybyciu do USA, korzystając z możliwości zmiany nazwiska przy okazji uzyskiwania amerykańskiego obywatelstwa i dając w ten sposób wyraz swojej ówczesnej fascynacji myślą Herberta Spencera. Stosunkowo wczesnym, ale już monograficznym ujęciem jego i jego współpracowników podejścia jest książka: R. S. Michalski, J. G. Carbonell, T. M. Mitchell, *Machine Learning: An Artificial Intelligence Approach*, Morgan Kaufmann 1983.

<sup>9</sup> Por. J. Wojtusiak, K. A. Kaufmann, *Ryszard S. Michalski: The Vision and Evolution of Machine Learning*, oraz K. J. Cios, Ł. Kurgan, *Machine Learning Algorithms Inspired by the Work of Ryszard Spencer Michalski*, w: *Advances in Machine Learning*, t. 1, J. Koronacki et al. (red.), Springer 2009.

podstawie posiadanej – również załadowanej do komputera – wiedzy dziedzinowej. Naturalny charakter owej indukcji miał być równoznaczny z przydaniem jej wynikom formy przyjaznej człowiekowi, łatwej do przyswojenia, np. mającej postać czytelnych i zrozumiałych reguł. Mówiąc bardzo ogólnie i jednocześnie trochę upraszczając, obszar ML miał obejmować nie tylko budowę i analizę działania oraz doskonalenie komputerowych systemów ML, które uczą się rozwiązywać konkretne zadania, ale także badanie i komputerową symulację ludzkich procesów uczenia się.

Nieprosta i może mniej znana jest historia związków ML ze statystyką matematyczną i statystyczną analizą wielowymiarową. Przedmiotem statystycznej analizy są zaobserwowane zbiory danych (przykładów), np. zbiór pacjentów, z których każdy opisany jest w jakiś ustrukturalizowany sposób, wspólny dla wszystkich. Początkowo przyjmowano, że oryginalne dane – bez ich jakiegoś przetworzenia – są opisywane przez wektory w pewnej przestrzeni pomiarowej, np. każdy pacjent jest opisany przez swoją płęć, wiek, ciśnienie tętnicze skurczowe, ciśnienie rozkurczowe i wybrane inne „parametry” jego organizmu. Dużo później dopuszczono inne, bardziej złożone struktury danych, które dopiero po odpowiednim przetworzeniu można było poddać działaniu algorytmów ML. Od początku, podstawowym paradygmatem statystyki było i pozostało założenie losowego charakteru danych – ich rozkład w przestrzeni, w której „żyją”, jest jakimś rozkładem prawdopodobieństwa. Wszak praktycznie wszystkie zjawiska i obiekty, które badamy, charakteryzują się jakimś stopniem niezdeterminowania (np. stężenie trójglicerydów w surowicy zdrowej dorosłej osoby nie musi wynosić i prawie nigdy nie wynosi dokładnie 150 mg/dl, mimo że taką wartość uznaje się za normę dla zdrowej osoby).

Bujny rozkwit nowoczesnej statystyki nastąpił z początkiem XX wieku. W 1901 roku Karl Pearson, ojciec równie co on wybitnego Egona, zapoczątkował tzw. analizę składowych głównych, wyszukującą w przestrzeni danych te „kierunki”, wzdłuż których skupiają się zaobserwowane przykłady. W 1909 roku antropolog Jan Czekanowski podał oryginalną metodę taksonomiczną – w przypadku jego badań odniesioną do antroposystematyki, ale w istocie stosowaną do dowolnych danych – która była pierwowzorem tych późniejszych metod numerycznej taksonomii, które obecnie noszą nazwę metod analizy skupień. Celem tych metod jest podzielenie zbioru przykładów na podzbiory (skupienia) przykładów do siebie podobnych i zarazem wyraźnie odmiennych od przykładów z innych skupień.<sup>10</sup> Wspomniane w tym miejscu dwie przykładowe metody należą do szerokiej palety metod składa-

<sup>10</sup> W roku 1963 Robert R. Sokal i Peter H. Sneath opisali w monografii *Principles of Numerical Taxonomy* metodę bardzo podobną do metody Czekanowskiego bez odnotowania tego podobieństwa, jakkolwiek w innym miejscu monografii wspomnieli, że „współczynnik Czekanowskiego” został użyty przez matematyków wrocławskich do skonstruowania w 1951 r. tzw. taksonomii wrocławskiej. Widać Sokal i Sneath sięgnęli do prac Wrocławian, ale nie zainteresowali się bliżej pionierską pracą Czekanowskiego.

jących się na dział statystyki zwany statystyczną analizą wielowymiarową. Już te przykłady pokazują, że ów dział ma za swój filar analizy geometryczne; poszukuje się odpowiedzi na takie czy inne pytania o sposób ułożenia się danych w przestrzeni, w której „żyją”. Czynnione założenia probabilistyczne są minimalne. Nie jest ważne, jaki dokładnie rozkład prawdopodobieństwa rządzi ułożeniem się danych w przestrzeni, a tylko, jakie są pewne cechy tego rozkładu, np. w jakich obszarach dane się skupiają, a w jakich ich (prawie) nie ma. Jest jasne, że taka analiza jest naturalnym sposobem maszynowego uczenia się np. tego, co charakteryzuje chorobę A, co chorobę B, i co je od siebie odróżnia. Jest też jasne, że takie „geometryczne” wyjaśnienie samo w sobie nie ma nic wspólnego z modelowaniem przyczynowo-skutkowym, np. nie mówi, dlaczego chorobie A odpowiada ta właśnie wykryta charakterystyka pacjenta. Taka jest cena i zaleta analizy eksploracyjnej, która nie przyjmuje z góry żadnego modelu przynajmniej z grubsza opisującego badane zjawisko. Inną zaletą zaś jest to, że analizę można przeprowadzić nawet wtedy, gdy taki model jest zupełnie nieznan. I dopiero ekspert dziedzinowy, zobaczywszy wyniki analizy eksploracyjnej, może pokusić się o zrozumienie źródeł pochodzenia tych wyników (np. źródeł otrzymanej charakterystyki choroby A).<sup>11</sup>

Co ciekawe, statystycy bardzo długo nie dostrzegali potencjału, jaki dla rozwoju ML ma ich analiza wielowymiarowa. Może dlatego, że jednocześnie – i zresztą słusznie – pracowali także nad rozwojem tego działu statystyki, który dziś nazywamy statystyką matematyczną.<sup>12</sup> Ale ta ostatnia, traktowana jako rdzeń badań statystycznych, a nie jeden z metodologicznych fundamentów tych badań, musiała przegrać z realiami. Zaś te realia mówiły o rosnących w niebywałym tempie zbiorach danych wymagających analizy oraz o tym, że dane te są generowane przez zjawiska coraz bardziej złożone i coraz mniej znane. Kto potrafi podać model charakteryzujący członków grupy terrorystycznej i przewidzieć zamiar dokonania terrorystycznego zamachu? W sieci – nie dla wszystkich i z pewnymi ograniczeniami – widoczne są przepływy pieniędzy między osobami, rozmowy telefoniczne, zakupy biletów lotniczych itd., itp., ale zachowania typowego dla terrorystów i nietypowego dla innych osób może się nauczyć jedynie inteligentny program ML. Podob-

<sup>11</sup> Rzecz jasna, jeśli to tylko możliwe, z modelowania przyczynowo-skutkowego nie powinno się rezygnować. *Ad spór między pragmatystami-nominalistami (rezygnacja z modelowania) a arystotelesowskimi realistami (modelować, jeśli się potrafi)* zob. J. Koronacki, M. Draminski, *Empirical Model Building Revisited*, w: *Models and Reality: Festschrift for James Robert Thompson*, J. Dobleman (red.), Chicago, IL: T&NO Company 2017.

<sup>12</sup> Bardzo dynamiczny rozwój statystyki matematycznej rozpoczął się w latach 20-tych ubiegłego wieku. Jej punktem wyjścia jest przyjęcie z góry ogólnego modelu probabilistycznego – jakiejś klasy rozkładów prawdopodobieństwa – rządzącego badanym zjawiskiem i, na podstawie zebranych (zaobserwowanych) danych zweryfikowanie adekwatności przyjętego modelu oraz oszacowanie wartości tych parametrów modelu, które zostały uznane za *a priori* nieznanne. Taki paradygmat statystyki matematycznej przenosił zadanie uczenia się na człowieka – dziedzinowego eksperta. To on stawiał hipotezę o modelu, zaś statystyk tylko tę hipotezę weryfikował. I jeśli ją statystyk sfalsyfikował, to eksperta zadaniem stawało się zaproponowanie innego modelu.

nie rzecz się ma z wykrywaniem kradzieży kart kredytowych. Przykłady tego typu, przykłady z przemysłu i biznesu można by mnożyć. Na polu nauki np. dane genetyczne i epigenetyczne pacjentów są tak dziś bogate, że bez pomocy programów ML nie sposób im się dokładniej przyjrzeć, by lepiej poznać mechanizm danej choroby. Jednym słowem, klasyczne podejście matematyków statystyków do problemów współczesnego świata i współczesnej nauki było anachronizmem.<sup>13</sup> Co innego podejście wyrastające ze statystycznej analizy danych, które dziś leży u podstaw obecnych sukcesów ML w SI i poza nią.

Kilku wybitnych badaczy statystyków i probablistów dostrzegło ten fakt w latach 80–tych dwudziestego wieku<sup>14</sup>, w 90–tych rozpoczął się imponujący rozwój statystycznych metod ML, zresztą realizowany wspólnie przez nich i informatyków jak też tych, którzy już wówczas zajmowali się ML od dekad. Na przełomie lat 80–tych i 90–tych na tyle już była rozwinięta kognitywistyka, w tym, co najważniejsze, neurokognitywistyka, że odżyła myśl o wykorzystaniu wyników tych badań w obszarze ML. Maszynowe uczenie się jako filar SI jest obszarem wedle starych podziałów interdyscyplinarnym.

Zaiste, obecny sukces tego, co wyrasta z myśli o ANN oraz statystycznych metodach maszynowego uczenia się, nie byłby możliwy bez rozwoju mocy obliczeniowych, technologii budowy i działania współczesnych komputerów (zaczynając od wielordzeniowych i wielowątkowych procesorów oraz organizacji i pojemności pamięci wewnętrznej i zewnętrznej). Bez tego rozwoju nie byłoby dziś mowy o paradygmatach tzw. głębokiego uczenia się ANN i np. uczenia z przenoszeniem (ang. *transfer learning*; umożliwia ono komputerowi przeniesienie wiedzy zdobytej przy rozwiązywaniu jednego zadania na uczenie się, jak rozwiązać podobne, ale inne zadanie). Także uczenie ze wzmocnieniem (czyli „nagradzaniem” uczącej się maszyny za jej sukcesy) nie mogłoby przebiegać tak efektywnie, jak przebiega. Te nowe paradygmaty pozwoliły np. na zbudowanie nowej reprezentacji wiedzy o języku naturalnym, nieznaney jeszcze kilkanaście lat temu. I to ta reprezentacja, oparta na statystycznym modelu produkcji językowych i zaimplementowana w głębokich ANN<sup>15</sup>, daje już dziś niezwykle obiecujące wyniki m.in. na polu prze-

<sup>13</sup> Por. L. Breiman, *Statistical Modeling: The Two Cultures*, *Statistical Science*, 16 (2001), s. 199–231.

<sup>14</sup> Niech tu za przykład posłuży książka Leo Breimana, Jerome’a H. Friedmana, Richarda A. Olshena i Charlesa J. Stone’a, *Classification and Regression Trees*, Wadsworth, Belmont, CA 1984, o drzewach decyzyjnych, czyli znakomicie przyjaznych użytkownikowi regułowych systemach decyzyjnych. W Polsce pierwsza konieczność zmiany podejścia do statystyki dostrzegła Elżbieta Peczeyńska, prof. em. Instytutu Podstaw Informatyki PAN.

<sup>15</sup> Punktem wyjścia dla takiej reprezentacji jest zauważenie, że jakkolwiek wypowiedane i zapisywane zdania nie mają same w sobie nic wspólnego z jakimiś obliczeniami, to przecież już sama „gola” analiza milionów dokumentów tekstowych (lepiej miliardów, a te są w zasobach takich potęg jak Google) pozwala – via modele probabilistyczne – przenieść słowa ze wszystkich dokumentów do wielowymiarowej przestrzeni, w której słowa są tej przestrzeni punktami, tym sobie bliższymi, w im węższych kontekstach wspólnie występują. I to jest tylko początek budowy reprezentacji języka umożliwianej przez przejście od takiego „płytkiego” paradygmatu opartego na „odległościach” między słowami do paradygmatu „głębokiego”, który np. pozwala na znalezienie przez maszynę kolejnego słowa w dotąd wygenerowanym przez nią tekście.



tworzenia języka naturalnego – maszynowego tłumaczenia tekstów, odpowiadania na pytania, streszczania dokumentów, czy generacji tekstów w języku naturalnym.<sup>16</sup> Wykorzystywanie stale rosnącej wiedzy neurokognitywistycznej – modeli pamięci krótkotrwałej, przenoszenia informacji z pamięci długotrwałej do krótkotrwałej – sprawia, że można być pewnym dalszego, najpewniej dramatycznego postępu na polu maszynowego „rozumienia” i przetwarzania języka naturalnego.<sup>17</sup>

Ostatnie zdanie jest równie słuszne w odniesieniu do wszystkich innych pól zastosowań ML w SI. Jeśliby chcieć opisać paradygmat głębokiego uczenia się w możliwie najprostszy sposób, to wystarczyłoby powiedzieć, że reprezentacja wiedzy zawartej w danych powstaje wielowarstwowo. Jest to, powtarzając za Ianem Goodfellowem, Joshuą Bengio i Aaronem Courville’em „zagnieżdżona hierarchia pojęć, gdzie każde pojęcie jest zdefiniowane w relacji z prostszymi pojęciami, i reprezentacje bardziej abstrakcyjne są obliczane w terminach reprezentacji mniej abstrakcyjnych”<sup>18</sup> (taka zagnieżdżona hierarchia nie wyklucza wywoływania przez wyuczone, bardziej ogólne pojęcia modyfikacji pojęć mniej ogólnych). Autorzy ci zwracają uwagę, że tym sposobem pojęcie głębokiego uczenia się wykracza poza perspektywę znaną neuronaukom i buduje ramy ML, które nie muszą być inspirowane przez te nauki.

Na koniec tego punktu taka jeszcze uwaga. Rewolucja informacyjna ... dopiero nabiera rozpędu. A to za sprawą niesłychanej rozbudowy farm serwerów i centrów danych największych korporacji (Google, Microsoft, Amazon i innych) oraz niezwykłych prędkości przesyłania danych w sieci coraz dokładniej oplatającej świat.

## SI DZISIAJ...

Jest rok 1997. Specjalnie zaprogramowany komputer Deep Blue koncernu IBM wygrywa sześciomecz w szachy z arcymistrzem Garrim Kasparovem 3,5:2,5. Zafascynowany tym wynikiem arcymistrz otwiera typ gier, w których człowiek może posługiwać się pomocą komputera w grze z innym człowiekiem. Z inicjatywy arcymistrza zaczynają być organizowane gry szachowe „w stylu wolnym”, gdzie przeciwko sobie występują „centaury”, inaczej zwane „cyborgami” – przeciwko sobie stają drużyny szachistów wspomaganych w czasie gry komputerowymi programami. Nie ma dziś liczącego się na

<sup>16</sup> Chodzi tu o system GPT-2, zob. A. Radford, J. Wu et al., *Language Models Are Unsupervised Multitask Learners*, Technical Report, OpenAI 2019;

[https://cdn.openai.com/better-language-models/language\\_models\\_are\\_unsupervised\\_multitask\\_learners.pdf](https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf)

<sup>17</sup> Por. J. L. McClelland et al., *Extending Machine Language Models toward Human-Level Language Understanding*, Technical Report, 2019; <https://arxiv.org/pdf/1912.05877.pdf>

<sup>18</sup> I. Goodfellow, J. Bengio, A. Courville, *Deep Learning*, MIT Press 2016; przekład polski: *Głębokie uczenie. Systemy uczące się*, PWN 2017.

świecie szachisty, który nie rozgrywałby treningowych partii z komputerem, albo lepiej powiedzieć komputerami, ponieważ każdy program uczy się grać na podstawie innych gier z przeszłości i przeto każdy program rozgrywa partie szachowe w (nieco) innym stylu. Dotyczy to także aktualnie najlepszego szachisty na świecie, Magnusa Carlsena. *Nota bene*, niejako w opozycji do Garri Kasparowa, stary arcymistrz Ronald Fisher zaproponował wprowadzenie do szachów elementu losowego, by „zwrócić” tę grę człowiekowi.

Czternaście lat później, w roku 2011 komputer IBM Watson pokonuje dwóch amerykańskich mistrzów gry w *Jeopardy*, quizie wymagającym rzeczywiście szerokiej wiedzy ogólnej, w której trójka zawodników słyszy odpowiedzi, zaś ich rola polega na zadaniu właściwego pytania, przy czym głos udzielony zostaje zawodnikowi, który pierwszy poprosił o prawo sformułowania pytania. Poziom quizu, szybkość reakcji zawodników, zwłaszcza Watsona i to zwłaszcza w rundzie trudniejszej, zapierały dech w piersiach. Zwycięstwa mistrzów w łatwiejszej rundzie stąd się głównie brały, co przyznał jeden z nich, że zgłaszali chęć odpowiedzi natychmiast, nim zdążyli pomyśleć nad właściwym pytaniem. Twórcy Watsona przeprogramowali go, w ostatnich latach załadowali doń wiedzę onkologiczną i dziś Watson dobrze sobie radzi wspomagając onkologów w diagnozowaniu różnych przypadków chorobowych.

Trzeba tu podkreślić, że – jak to w systemach ML czy szerzej SI – Watsona baza danych jest ciągle wzbogacana, Watson się doskonali (uczy) i tym sposobem jego wiedza stale rośnie. Uznany obserwator technologicznego postępu Kevin Kelly ma na pewno rację stwierdzając, że „[j]eśli SI potrafi pomóc ludziom w staniu się lepszymi graczami w szachy, to ma rację bytu myśl, iż może także pomóc nam zostać lepszymi pilotami, lekarzami, sędziami, czy nauczycielami (tłum własne)”<sup>19</sup>

W roku 2015 program AlphaGo firmy DeepMind (przejętej w 2014 przez Google) rozgromił zawodowego gracza w chińska grę go, nieporównanie trudniejszą niż szachy. Dwa lata później badacze firmy Facebook udostępnili program zdolny do prowadzenia dialogu i negocjacji w języku naturalnym.<sup>20</sup> Podajemy ten przykład m.in. dlatego, że algorytm uczący się negocjacji okazał się zdolny dokonywać pewnego oszustwa, przy czym przeciwnikiem algorytmu byli ludzie nie wiedzący, że negocjują z komputerem.

Zadanie polegało na wynegocjowaniu przez dwie strony akceptowalnego dla nich podziału zadanego zbioru towarów, przy czym wartość każdego z tych towarów może być inna dla każdej ze stron i dana strona nie wie, jaką wartość ma dany towar dla strony drugiej. Każda ze stron maksymalizuje

<sup>19</sup> K. Kelly, *The Inevitable: Understanding the 12 Technological Forces That Will Shape Our Future*, Viking Press 2016; wydanie polskie: *Nieuniknione. Jak inteligentne technologie zmienią naszą przyszłość*, Poltext 2017.

<sup>20</sup> Patrz M. Lewis, D. Yarats et al., *Deal or No Deal? End-to-End Learning for Negotiation Dialogues*, 2017; <https://arxiv.org/abs/1706.05125>; oraz D. Yarats, M. Lewis, *Hierarchical Text Generation and Planning for Strategic Dialogue*, 2017, <https://arxiv.org/abs/1712.05846>

sumaryczną wartość wynegocjowanych przez nią towarów pod warunkiem dojścia przez obydwie strony do porozumienia i zgodnego zakończenia negocjacji. Ku zaskoczeniu jego autorów, algorytm (wykorzystujący ANN i oparty na paradygmacie uczenia się ze wzmocnieniem) nauczył się sugerować w pewnych sytuacjach drugiej stronie (negocjującemu człowiekowi), wbrew prawdzie, że dany towar jest dla niego bezwartościowy, by potem łatwiej ów towar od drugiej strony odzyskać i zakończyć za obopólną zgodą negocjacje.

Drugą przyczyną, która sprawiła, że opisaliśmy powyższe zadanie, jest odnotowanie, iż był to pierwszy poważny krok w kierunku otrzymania chatbotów (programów komputerowych zdolnych do prowadzenia dialogu z ludźmi), które mają być w stanie prowadzić w zastępstwie człowieka działania bardziej złożone niż np. udzielanie odpowiedzi zadawanych telefonowi komórkowemu czy prowadzenie wstępnych rozmów z kandydatami do pracy (jak to się dziś w Polsce mówi, rozmowami prowadzonymi przez dział HR).

W poprzednim punkcie wspomnieliśmy już o wielkim postępie na polu zastosowania SI do przetwarzania i generacji tekstów w języku naturalnym. Dodajmy, że przywołany tam system GPT-2 ma wbudowanych w siebie około półtora miliarda parametrów dostrajanych przez algorytm w trakcie jego uczenia się. Jakość generowanych na zadany temat tekstów była tak wysoka, że twórcy udostępnili ogółowi użytkowników sieci jedynie uproszczoną wersję systemu o złożoności około 150 milionów parametrów. System GPT-2 mógłby być np. użyty do produkowania fałszywych wiadomości (*fake news*). Nie trzeba dodawać, że taki system jest produktem znacznie bardziej wyrafinowanym niż program piszący wiersze podobne do wierszy wybranego poety, czy muzykę podobną do utworów wybranego kompozytora.

Dodajmy też, że *New Scientist* doniósł w październiku 2019 o odczytaniu przez program firmy DeepMind zniszczonych starożytnych tablic greckich, z którym to zadaniem gorzej sobie radzili badacze.<sup>21</sup> W tym przypadku trzeba było sięgnąć po głębokie ANN nie tylko uczące się analizować tekst, ale także analizować obrazy. Warto jeszcze nadmienić, że korporacja Google dysponuje dobrze się sprawdzającym programem, który automatycznie generuje podpisy pod obrazami (zdjęciami). No i jeśli o analizie obrazów mowa, to dodajmy na koniec, że rozpoczęty został nowy rozdział maszynowego uczenia się – mianowicie uczenia się kształtów – przedmiotów, liter – na podstawie nie bardzo wielu, a jednego, czy kilku przykładów. W tym przypadku znowu istotną rolę odgrywa głębokie uczenie się ANN i/albo modelowanie probabilistyczne.<sup>22</sup>

<sup>21</sup> <https://www.newscientist.com/article/2220438-deepmind-ai-beats-humans-at-deciphering-damaged-ancient-greek-tablets/>

<sup>22</sup> Zob. G. Koch, R. Zemel, R. Salakhutdinov, *Siamese Neural Networks for One-shot Image Recognition*, Proc. 32nd International Conference on Machine Learning, Lille, France, 2015, t. 3, JMLR: W&CP; oraz B. M. Lake, R. Salakhutdinov, J. B. Tenenbaum, *Human-level Concept Learning through Probabilistic Program Induction*, Science, 11, t. 350, zeszyt 6266, 2015.

Cele, którym służą systemy SI/ML, nie zawsze są bezsprzecznie światlane. Wspomniani właśnie użytkownicy sieci to miliardy osób. Dzisiejsze serwery pozwalają na gromadzenie wprowadzanych do nich przez te osoby najróżniejszych danych. Tym razem chodzi o ogół, a nie np. o pacjentów onkologicznych, o których informacje zbiera IBM Watson. I oto najpotężniejsze koncerty usług internetowych wiedzą dziś o nas prawie wszystko za sprawą algorytmów ML. Nie znają naszych myśli, ale wiedzą, co lubimy, a czego nie, jakie są nasze zainteresowania, potrafią stworzyć mniej lub bardziej trafny profil psychologiczny każdego z nas. Wszystko to na podstawie gromadzonych w pamięciach komputerów naszych listów elektronicznych, pisanych przez nas (jeśli je piszemy) tekstów na portalach Facebooka, Twittera, zdjęć ładowanych do Instagramu, czy sposobów wykorzystywania przez nas internetowych wyszukiwarek oraz platform do elektronicznego handlu. Ludzie uzależniają się od SI, ale też najczęściej czynią to z ochotą, radzi przyjąć sugestie systemu rekomendacyjnego polecającego film, czy systemu podsuwającego reklamę butów podobnych do oglądanych wcześniej w internetowym sklepie.

Lepszą ilustracją możliwości uzależnienia człowieka od SI jest opracowana przez Nicka Fostera (badacza z Google X) w 2016 roku koncepcja systemu informatycznego o nazwie *Selfish Ledger*. Koncern zarzeka się, że nie ma zamiaru projektu realizować, ale pomysł odnotować warto. Zapamiętane przez sieć informacje o każdym z nas – naszych preferencjach, sposobach reagowania na zdarzenia zewnętrzne, sposobach radzenia sobie z nadwagą albo nadmierną chudością, leczeniem depresji, itd., itp. – przestają być prywatną własnością. Odpowiednio przetworzone przez komputerowy program mogą stać się nieomal niezastąpionym doradcą użytkownika sieci, tego z nadwagą, albo nie wiedzącego w co zainwestować oszczędności, albo szukającego dowolnej innej, w tym psychologicznej porady. I takim doradcą, stale się doskonalącym na podstawie ocen jego dotychczasowego działania oraz powiększania się jego zasobów wiedzy, mógłby stać się system *Selfish Ledger*.

Nie trzeba wielkiej wyobraźni, by dostrzec, jak potężny potencjał władzy znalazł się w rękach panujących nad Internetem. Potencjał dziś jawnie wykorzystywany przez rząd Chin, niejawnie przez inne rządy albo osoby ubiegające się o władzę. Nie przypadkiem wśród najbliższych doradców Donalda Trumpa, wówczas kandydata na urząd Prezydenta, znalazły się – o co nie mam zresztą do niego pretensji – najważniejsze osobistości z Doliny Krzemowej i rynków finansowych, i nie przypadkiem znakomita firma software'owa jednej z nich profilowała obywateli USA z myślą o wspomżeniu kampanii przyszłego prezydenta. Rząd Chin zainicjował w roku 2014 powstanie komputerowego, inteligentnego systemu oceny obywateli Państwa Środka o nazwie *China Social Credit*, z myślą o uzależnieniu np. wydania

pozwolenia na budowę domu czy licencji na prowadzenie rikszy od oceny wygenerowanej przez system.<sup>23</sup>

### ...I JUTRO WSZĘDZIE

Ostatnie dekady przynoszą stale postępującą wirtualizację rzeczywistości, w której żyje człowiek. Jasne jest, że SI ma w tym ciągle rozwijanym procesie coraz bardziej istotny udział, co już zostało zasygnalizowane, ale też wspomnieć jeszcze należy np. o tworzeniu wirtualnych światów (albo życia w świecie wirtualnych gier). Nie zatrzymując się dłużej nad analizą owego procesu, tu przypomnijmy takie jego podsumowanie dane u schyłku ubiegłego wieku przez Manuela Castellsa:

„W ostatnim ćwierćwieczu XX wieku, skoncentrowana wokół informacji rewolucja technologiczna zmieniła sposób, w jaki myślimy, produkujemy, konsumujemy, handlujemy, zarządzamy, komunikujemy się, żyjemy, umieramy, kochamy się. [...] Kultura rzeczywistej wirtualności, skonstruowana na bazie coraz bardziej interaktywnego świata audiowizualnego, przeniknęła wszędzie mentalne zdolności reprezentacji i komunikacji, scalając różnorodne kultury w jednym elektronicznym hipertekście. Przestrzeń i czas, będące materialnymi podstawami ludzkiego doświadczenia, także uległy przemianie, ponieważ przestrzeń przepływów dominuje dziś nad przestrzenią miejsc, zaś bezczasowy czas zastępuje zegarowy czas epoki przemysłowej.”<sup>24</sup>

Życie w środowisku informacji przesyłanych przez komputerową sieć zrywa z kulturą dotąd znaną, zawsze mocno osadzoną w danym miejscu i czasie. Przez całą historię wartości i cele społeczeństwa rodziły się na styku wiary i objawienia oraz codziennego doświadczenia, zmagania się człowieka ze światem, związków między ludźmi, którzy żyli w określonym miejscu i czasie. Ale paradygmaty wieku informacji są zupełnie inne. Jak pisał Castells, zresztą z pewną tego aprobatą, zaczął się marsz ku rzeczywistej wirtualności. A rzeczywista wirtualność jest częścią nowej kultury niesionej przez SI.

Na naszych oczach nabiera rozmachu proces o równie fundamentalnym znaczeniu. To proces inteligentnej robotyzacji, czyli otaczania człowieka coraz większą liczbą robotów, botów (programów zdolnych wykonywać pewne czynności zamiast człowieka) i innych urządzeń o pewnych zdolnościach poznawczych. Wspomniany już Kelly nazywa ten proces kognifikacją (*cognifying*). Owe zdolności poznawcze będą za każdym razem bardzo ogra-

<sup>23</sup> Ad inne przykłady „inteligentnego” działania państw autorytarnych: A. Kendall-Taylor, E. Frantz, J. Wright, *The Digital Dictators: How Technology Strengthens Autocracy*, Foreign Affairs, March–April 2020.

<sup>24</sup> M. Castells, *The Information Age: Economy, Society and Culture*, t. 1–3, Wiley-Blackwell 1996 – 1997. Przekład polski pod kierunkiem Mirosławy Marody, PWN 2007–2009.

niczone, inteligentne urządzenie będzie – i takie już są – zdolne zastępować człowieka w wykonaniu bardzo konkretnej czynności. Ma rację Kelly twierdząc, że w dającej się przewidzieć przyszłości nie będzie mowy o urządzeniach przypominających inteligencję ogólną, a więc zdolnych do działania na różnych polach, wcześniej nie wpisanych w software urządzenia, i w różnych warunkach. Urządzenia takie – jak dzieje się już obecnie – będą „myśleć” po swojemu i „myślenie” to nie będzie miało nic wspólnego z myśleniem człowieka. Najczęściej będą to urządzenia o bardzo ograniczonych możliwościach poznawczych, jeśli w ogóle ich zdolności będzie można nazwać poznawczymi. Na przykład już dziś mamy pacjentów wyposażonych w różne sensory, mierzące różne parametry organizmu, przekazujące ich wartości programowi zdolnemu wychwycić sytuacje zagrażające właściwemu stanowi zdrowotnemu pacjenta. Nieporównanie bardziej złożonymi bytami, ale nadal bardzo ograniczonymi w swoich poznawczych zdolnościach i również już dziś obecnymi, są roboty zdolne wykonywać zadania u boku i wspólnie z człowiekiem, „ostrożne” i dbające o to, by nie potracić swojego ludzkiego współpracownika.

Możemy oczekiwać, że w niedalekiej przyszłości pojawią się na szosach flotyle ciężarówek, z których tylko pierwsza i ostatnia będą kierowane przez ludzi, zaś pozostałe nie będą miały kierowców. Innym zadaniem o podobnym stopniu skomplikowania jest opracowywany przez badaczy z MIT system „przyjaznego” wspomaganie kierowcy w jego jeździe, gdzie ów przyjazny charakter pomocy ma oznaczać „wczuwanie się” w emocjonalny stan kierowcy – dostrzeżenie spadku stopnia koncentracji czy zdenerwowanie kierowcy (a to przez obserwowanie ruchu gałek ocznych, grymasów twarzy itd.).<sup>25</sup> *Nota bene*, pojawienie się na szosach prawdziwie autonomicznych pojazdów nie będzie możliwe jeszcze przez dekady, jeśli w ogóle będzie ziszczalne bez przebudowy szos oraz zreorganizowania zasad poruszania po nich i wokół nich.

Ale społeczne problemy, jakie owa kognifikacja będzie stwarzać, leżą gdzie indziej. Po pierwsze, w tym, że normalny człowiek – właśnie, normalny – nie żył nigdy otoczony robotami, np. przygotowującymi jedzenie w McDonalddie oraz innymi obsługującymi go i sprzątającymi w tymże. Osoba ludzka woli ludzką osobę serwującą jej lekarstwo w szpitalu, a nie robota wykonującego tego typu obowiązki. Nade wszystko woli osobę wyrażającą emocje, a nie robota te emocje (do jakiegoś stopnia) naśladowującego.<sup>26</sup>

Po drugie, parafrazując Kelly’ego, kognifikacja obejmie swoim zasięgiem tyle urządzeń, ile kiedyś objęła elektryfikacja. Będziemy tak samo otoczeni

<sup>25</sup> L. Fridman, *Human-Centered Autonomous Vehicle Systems: Principles of Effective Shared Autonomy*, Technical Report, 2018; [https://www.researchgate.net/publication/328063361\\_Human-Centered\\_Autonomous\\_Vehicle\\_Systems\\_Principles\\_of\\_Effective\\_Shared\\_Autonomy](https://www.researchgate.net/publication/328063361_Human-Centered_Autonomous_Vehicle_Systems_Principles_of_Effective_Shared_Autonomy)

<sup>26</sup> Por. A. Stępnik, *Emocje a maszyny. O możliwościach i konsekwencjach stworzenia inteligentnych emocjonalnie maszyn*, *Studia Metodologiczne*, s. 38, 2017.

takimi urządzeniami, jak dziś żyjemy pośród urządzeń pobierających energię elektryczną. Problem natomiast w tym, że człowiek będzie doświadczać ich wszechobecności w niemałej mierze w zastępstwie kontaktu z innymi ludźmi.

Po trzecie, jest oczywiste, że społeczeństwa rozwiniętych państw czeka rewolucja w strukturze zatrudnienia. Roboty, chatboty, koboty (*collaborative robots*, czyli roboty współpracujące z człowiekiem), w tym urządzenia wyłącznie software'owe, odbiorą pracę znacznej liczbie robotników przemysłowych, pracowników najszerzej pojętych usług (od pomocniczego personelu medycznego i przemysłu farmaceutycznego, poprzez pracowników sektora transportu po cały sektor usługowy *sensu stricto*), czy pracowników dbających o dobro finansów w biznesie (księgowość, prowadzenie transakcji biznesowych, zapobieganie oszustwom na tym polu, i nie tylko). *Novum* w tej materii na tym będzie polegać, że nie jest jasne, gdzie te bezrobotne osoby zostaną skierowane przez przyszły rynek pracy. Prawdą jest, że u początków Stanów Zjednoczonych Ameryki tylko kilka procent obywateli nie pracowało na roli (na marginesie, do USA sprowadzono około 400 tysięcy niewolników z Afryki, nie więcej), że procent osób żyjących z rolnictwa wynosił około 60% w roku 1850, a dziś wynosi około 2%. Ale wówczas farmerzy i ich pomocnicy byli wysysani przez rynek, na których czekała praca gdzie indziej. W Anglii, kolebce rewolucji przemysłowej, ludzie kierowali się ze wsi do miast w poszukiwaniu szans na lepsze życie. Obecna rewolucja ma na razie jedynie odebrać wielu ludziom możliwość pracy.

I po czwarte, co będzie pełnym *novum* – jeśli się spełnią zapowiedzi – również albo szczególnie dojmującym, mają stać się pewne zasady przyszłej współpracy ludzi i robotów. Na przykładzie: Na przełomie wieków amerykańska agencja DARPA (*Defense Advanced Research Projects Agency*) zainicjowała projekt skonstruowania domózgowych implantów (*brain/machine interfaces*), które pozwolą ludzkiemu operatorowi sterować na odległość np. robotem albo dźwigiem. Przypuszcza się, że doprowadzi to do rozmycia granicy oddzielającej człowieka od mechanicznego sprzętu – z czasem zaangażowany w to sterowanie fragment kory mózgowej uzna maszynę za kolejną przez nią poruszaną kończynę.<sup>27</sup>

Krótko zaś podsumowując czym jest i będzie SI, należy zacząć od tego, że inteligencja nie odwołuje się do niczego, co istniałoby jako takie – w sobie, z siebie i dla siebie. Odnosi się ta uwaga do inteligencji naturalnej i, niejako tym bardziej, do inteligencji sztucznej. Inteligencja, o której tu mowa, jest względną miarą efektywności wykonywania zadań. Następnie należy podkreślić, że SI w niczym dziś nie wykracza – i w dającej się przewidzieć przyszłości nie wykroczy – poza to, co daje zwykły, jakkolwiek imponująco szybki, postęp technologiczny, który korzysta z nauk ścisłych, inżynierskich

<sup>27</sup> Ad plany DARPA: C.T. Rubin, *Man or Machine*, The New Atlantis, 4, 2004.

i biologicznych (neurokogniwistyki). Nieprawdą jest, iż urządzenia ML albo szerzej SI, realizując zadania im powierzone, naśladowują sposób rozwiązywania przez człowieka podobnych zadań. Dalece czym innym jest naśladownictwo czegoś oraz wykorzystywanie wiedzy o jakimś aspekcie czegoś. Korzystając z przykładu użytego przez Kelly'ego, samolot nie dlatego lata, że naśladuje ptaka. To fizyczne prawa rządzące zdolnością ptaka do fruwania zostały użyte do zbudowania samolotu, który nie macha skrzydłami, lecz inaczej z owych praw korzysta. *Nota bene* z tego, że samolot wykonuje swoje zadania pod pewnymi względami lepiej od ptaka – np. szybciej, potrafi unieść ciężar ważący tony – nie wynika, że jest od ptaka lepszy. Jest samolot bytem istotowo różnym od ptaka, i tyle.

### W ŚWIECIE BEZ LOGOSU

Jest wiele racji w twierdzeniu, że obecne i nadchodzące dekady są czasem wielkich, fundamentalnych zmian cywilizacyjnych i kulturalnych. Kelly zwraca uwagę, że rozwój SI stawia przed nami pytanie, po co jest człowiek. Píše też, że to SI pomoże nam zdefiniować, czym jest człowieczeństwo. Ostatnie sformułowanie o tyle jest nietrafne, że nie chodzi tu o jakąś redefinicję, zdefiniowanie człowieczeństwa na nowo, ale o przypomnienie sobie, czym człowieczeństwo jest i że potrafiła to przez całe tysiąclecie wyrażać chrześcijańska cywilizacja Zachodu. Jeśli ma w tym pomóc SI, niech tak się stanie. Ale moc jej przekonywania może mieć charakter jedynie negatywny – uświadamiając człowiekowi, że stoi na krawędzi przepaści.<sup>28</sup>

Książka Kelly'ego jest jednym z wielu dowodów na to, że *techne* zastąpiło *logos*. Na początku XX wieku to człowiek kontynuował proces odczarowywania świata, jeszcze ciągle widział się w centrum kosmosu, czyniąc Ziemię sobie poddaną. Ale też nie mylił się Weber bolejąc nad tym, że procesowi „odczarowania świata” będzie towarzyszyć utrata poczucia sensu jednostki i szeroko opisywał nieuniknione konsekwencje tego procesu – postępujące jednostki zniewolenie. Nie wiedział tylko, że obecna, pochrześcijańska cywilizacja Zachodu uzna, iż wszystko co jest technologicznie możliwe – być może poza nowymi środkami prowadzenia wojen – jest dobre. I pozwoli ta cywilizacja prowadzić się technologii, zgodzi się uznać, iż samym rdzeniem cywilizacji Zachodu jest technologiczny rozwój.

W tym samym kontekście Yuval Levin, znakomity obserwator społecznej współczesności, zwłaszcza amerykańskiej, pisze o odrzuceniu antropologii ciągłości pokoleń na rzecz antropologii innowacji. Ta pierwsza implikuje kultywowanie cnót i życie we wspólnocie wiążącej pokolenia, ta druga czyni

<sup>28</sup> Tak stan Zachodu przedstawiali Jürgen Habermas i Joseph Ratzinger podczas debaty w Katolickiej Akademii Bawarii w styczniu 2004 roku. Patrz też J. Ratzinger, *Europa Benedykta w kryzysie kultur*, Libreria Editrice Vaticana oraz Edycja Świętego Pawła 2005.



fetyszem innowacje z przyszłością niezakotwiczoną w przeszłości i terażniejszości?<sup>29</sup>

Rzecznik antropologii innowacji jest dzieckiem nowożytnego gnostycyzmu, ale żyjącym już po końcu historii, ku któremu kiedyś prowadziła kontemplacja Hegla przemieniona w działanie przez Marksa, Engelsa i innych.<sup>30</sup> Nowość tej gnozy, zwłaszcza w wydaniu transhumanizmu na tym polega, że w dzisiejszym świecie z Bogiem uśmierconym i bez *logosu* prowadzi ku obaleniu człowieka i nicości.

Szerzej tę ostatnią kwestię poruszam gdzie indziej<sup>31</sup> i w tym miejscu potwórę tylko tyle:

„Filozofia, zrodzona z umiłowania bytu, została zastąpiona przez antyfilozofię, która ogłosiła się posiadaczką całej wiedzy o świecie i w nim człowieku. W doskonale zsekularyzowanym świecie ubóstwiła autonomicznego człowieka i ostatecznie [...] musiała zawołać o nadczłowieka.<sup>32</sup> [...] Co jednocześnie wcale nie zaskakujące, niektórzy z »proroków« SI, chwalców społeczeństwa sieci i guru transhumanizmu postulują nowe niby-religie. No cóż, bez mitu – w rozumieniu Leszka Kołakowskiego<sup>33</sup> – ani rusz. Słusznie Alan Jacobs (w *The New Atlantis*, 58, Spring 2019) zwraca uwagę, że oto na naszych oczach technologiczny rdzeń naszej cywilizacji ma się przekształcić w jej rdzeń quasi-mityczny. Kiedyś miał to być kult Rozumu, potem religia Ludzkości, dziś ma to np. być dataizm Yuvala Harariego<sup>34</sup>. Erik Davis cieszy się, że jego książka *Techgnosis: Myth, Magic & Mysticism in the Age of Information* z końca lat 1990–tych miała swoje kolejne wznowienie w roku 2015. Davis widzi w sieci ogarniającej nasz glob i wszystkie społeczeństwa źródło jakiegoś »technomistycyzmu«. W przedmowie do wydania książki z 2015 roku, filozof i medioznawca Eugene Thacker odnajduje w Davisa przedstawieniu technologii – jej historii i dnia dzisiejszego – coś boskiego, »religię wyrażoną innymi środkami«.»<sup>35</sup>

<sup>29</sup> Patrz Y. Levin, *Imagining the Future, The New Atlantis*, 4, 2004. Levin odnosi swoją analizę głównie do postępu w obszarze biotechnologii, ale równie dobrze obejmuje ona problematykę nas tu interesującą.

<sup>30</sup> Ów koniec historii „co do zasady” zdawał się dobrze dostrzegać Alexandre Kojève, por. P. A. Lawler, *Postmodernism Rightly Understood: The Return to Realism in American Thought*, Lanham-Boulder, New York 1999; p. też J. Koronacki w portalu Teologii Politycznej: *O końcu historii i transhumanizmie* oraz *Idee mają konsekwencje*; <https://teologiapolityczna.pl/prof-jacek-koronacki-o-koncu-historii-i-transhumanizmie> oraz <https://teologiapolityczna.pl/prof-jacek-koronacki-idee-maja-konsekwencje>

<sup>31</sup> J. Koronacki, *Posthumanizm a technologia*, w: *W poszukiwaniu nowego człowieka*, M. Szulakiewicz (red.), XXIV Colloquia Torunensia, Toruń 2020 (w druku).

<sup>32</sup> Pisząc w tamtym omówieniu o ambicjach filozofii świadomie pominąłem kwestię filozofii skończoności. Ale też chyba trudno nie zgodzić się, że najważniejsze, co filozofowie skończoności mieli nam do powiedzenia, to zdekonstruowanie naszych teoriopoznawczych ograniczeń (choć wszyscy mówili na różne sposoby prawie to samo) oraz – paradoksalnie – *de facto* przyznanie, że bez spojżenia ku nieskończoności i odwołania do osądu moralnego nic o sobie sensownego powiedzieć nie jesteśmy w stanie (por. np. doskonała synteza Włodzimierza Lorenca, *Filozofie skończoności*, Warszawa 2016).

<sup>33</sup> Zwłaszcza *Obecność mitu*, Paryż 1972.

<sup>34</sup> Por. jego *Homo Deus: A Brief History of Tomorrow*, San Francisco 2017.

<sup>35</sup> Vide też M. Hetmański, *Informatyka jako nauka społecznie użyteczna*, *Studia Metodologiczne*, s. 38, 2017.

Podsumowując i w ramach tego podsumowania przywołując jedynie diagnozę Rémięgo Brague'a, choć można by tu zacytować myślicieli wielu: W ramach dominujących prądów myślowych nie jest możliwe uprawomocnienie istnienia człowieka. Pal licho godność człowieka, albo jej brak, chodzi o coś więcej – nie można odpowiedzieć na pytanie, czy to dobrze, czy źle, że człowiek istnieje. I na pewno jest człowiek z technologicznej perspektywy bytem przestarzałym, jeśli nie zbędnym. Z tym, że ta perspektywa jest chora i – miejmy nadzieję – jest jeszcze czas, by ją zmienić.<sup>36</sup> Nie jest to pewne. Antropologia innowacji działa jak zasłona oddzielająca prawdę o człowieku od rzeczywistości sztucznej, zaprasza człowieka do przejścia na stronę tej drugiej, bez możliwości spojrzenia za siebie.<sup>37</sup>

Rozwój technologiczny jest nieunikniony i jest potrzebny. Jak i w przeszłości, rozwój ten niesie wiele dobra, ale jednocześnie ma też dziś wyjątkowo groźne oblicze. Z tym że to groźne oblicze jest konsekwencją upadku kultury i kryzysu cywilizacyjnego, a nie wynikiem immanentnego zła ukrytego w technologicznym rozwoju.

### CZY CZŁOWIEK JEST MASZYNĄ?

Gdy mówimy o ML i SI, mamy na myśli ten obszar badań i zastosowań, które John Searle nazwał 40 lat temu słabą sztuczną inteligencją. Ta słaba SI, obecna wśród nas dziś i której rozwój potrafimy przewidywać, nie ma nic wspólnego z tym, co Searle nazwał mocną sztuczną inteligencją. Ta pierwsza mówi o tym, o czym tu dotąd była mowa i co doskonale mieści się w palecie naszych inżynierskich dokonań i planów rozwoju, i co w tym artykule zostało zasygnalizowane. Ta druga to tak naprawdę my i ewentualnie, jeśli już *sensu stricto* sztuczna, to superinteligencja wymarzona przez trans- i posthumanistów. Dlaczego my? Ponieważ u podstaw myśli o mocnej sztucznej inteligencji leży założenie, że człowiek jest li tylko maszyną przetwarzającą informację, komputerem, jakkolwiek my sami jeszcze nie potrafimy tego komputera skonstruować. Naturze już to się udało, jakoby *via* ewolucja, my musimy to zrobić na nowo albo jakoś odtworzyć tamtą historię, by – tego chcą trans- i posthumanieści – oddać władzę owej mocnej sztucznej inteligencji, która ewolucję poprowadzi dalej, ku rzeczonyj superinteligencji.

Ale zostawmy na boku marzenia – czy rojenia – trans- i posthumanistów.<sup>38</sup> Spójrzmy na to, co proponują nam filozofie umysłu.<sup>39</sup> Znakiem na-

<sup>36</sup> Por. *Le Propre de l'Homme: Sur une légitime menace*, Paryż 2013 (tłum. ang., z którego korzystałem: *The Legitimacy of the Human*, South Bend, Indiana 2017); zob. też np. zbiór wykładów Brague'a wydany w 2019 pod wspólnym tytułem *Curing Mad Truths: Medieval Wisdom for the Modern Age*, Notre Dame, Indiana, oraz *Modérément moderne*, Paryż 2014; tłum. ang. *Moderately Modern*, South Bend, Indiana 2017.

<sup>37</sup> Tę ostatnią uwagę zawdzięczam prof. Beniaminowi Goldysowi z University of Sydney.

<sup>38</sup> Por. J. Koronacki, *Posthumanizm a technologia*, op. cit.

szego czasu, właściwie począwszy od oświecenia, jest naturalizm, albo raczej scjentyzm. Niczego, co wykraczałoby poza naturę, nie ma, nie ma zewnętrznej racji istnienia i nie ma innego sensu niż obecny w naturze (czyli sens praw natury, jeśli ten określić owym mianem). Wybitni filozofowie umysłu potrafią bronić mechanicyzmu i komputacjonizmu, trafnie widząc w nich dobry fundament, na którym można oprzeć modele funkcjonowania mózgu i pewnych aspektów poznania, ale nie twierdzą zarazem, że tym sposobem docierają do istoty umysłu. Inni, jak John Searle, wolą wyjść ponad teorie obliczeniowe i kierują się ku biologicznemu naturalizmowi oraz emergentyzmowi. Fascynujące analizy tych stanowisk można znaleźć u Edwarda Fesera, czy Jamesa D. Maddena.<sup>40</sup> Wyjątkowej wagi analizę znajdujemy u Thomasa Nagela<sup>41</sup>, który będąc ateistą i naturalistą, jednocześnie w sposób fundamentalny sprzeciwia się obowiązującej tradycji myśli nowożytnej i nowoczesnej. Przekonująco i precyzyjnie uzasadnia, iż to co mentalne i to co fizyczne istnieje obiektywnie, ale jedno jest doskonale nieredukowalne do drugiego – są to rzeczywistości różne, z tym co mentalne nie dającym się opisać przez współczesne nauki szczegółowe. Nagel jest przekonany, że za pokolenie czy dwa będący już zresztą w kryzysie neodarwinizm – co ciekawe, kryzysie wymuszonym przez imponujący rozwój zwłaszcza biologii – zostanie ostatecznie wyśmiany. Liczy na powstanie nowych paradygmatów nauki, które zbliżą nas do wyjaśnienia tajemnicy umysłu – świadomości, poznania, przekonania o obiektywnej prawdziwości czegoś – choć może tej tajemnicy nie pojmujemy nigdy. Można rzec, że intuicja podpowiada Nagelowi to, co doskonale ujął Stanisław Krajewski pisząc o filozoficznych konsekwencjach twierdzenia Gödla.<sup>42</sup>

Kończąc te rozważania warto jeszcze przypomnieć dziełko Rogera Scrutona *On Human Nature* (Princeton–Oxford 2017) i słowa Leszka Kołakowskiego, którymi kończył swój esej *Horror metaphysicus*: „I czyż nie narzuca się podejrzenie, że gdyby »być« nie było »do czegoś«, a świat pozbawiony sensu, to nie tylko nigdy nie zdołalibyśmy wyobrazić sobie, że jest inaczej, lecz nawet pomyśleć nie bylibyśmy w stanie tego właśnie, że być nie jest w istocie »do czegoś«, a świat pozbawiony sensu?”<sup>43</sup>

<sup>39</sup> Zob. np. M. Miłkowski, R. Poczobut (red.), *Przewodnik po filozofii umysłu*, Karków 2012; zob. też np.: M. Miłkowski, *From Computer Metaphor to Computational Modelling: The Evolution of Computationalism*, *Minds and Machines*, 28, 2018, 515–541; idem, *Objections to Computationalism: A Survey*, *Roczniki Filozoficzne*, LXVI, 3, 2018.

<sup>40</sup> Patrz E. Feser, *Philosophy of Mind: A Beginner's Guide*, Oxford 2005 oraz: idem, *Aristotle's Revenge: The Metaphysical Foundations of Physical and Biological Science*, 2019 editiones scholasticae, Neunkirchen-Seelscheid; J. D. Madden, *Mind, Matter & Nature: A Thomistic Proposal for the Philosophy of Mind*, Washington, DC, 2013.

<sup>41</sup> T. Nagel, *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*, Oxford 2012.

<sup>42</sup> S. Krajewski, *Umysł a metalogika*, w: *Przewodnik po filozofii umysłu*, op. cit., oraz *Twierdzenie Gödla i jego interpretacje filozoficzne*, Warszawa 2003.

<sup>43</sup> L. Kołakowski, *Horror metaphysicus*, Res Publica, Warszawa 1990.

## ZAMIAST ZAKOŃCZENIA

Andrew Lytle, ważny reprezentant amerykańskiego Południa, pisał w 1930 roku, iż

„...dążenie do szczęścia zostało zastąpione kręceniem się w kółko, pozbawionym nawet tej logiki, jaką kieruje się pies goniąc własny ogon.

[...]

Jest to sprawa konfliktu między ludźmi a nienaturalnym pokłosiem geniuszu inwencji. Jest to wojna na śmierć i życie między techniką a zwykłym ludzkim funkcjonowaniem. Prawa do takiego funkcjonowania są naturalnymi prawami człowieka, ale dziś, w XX a nie XVIII wieku, są to prawa zagrożone po raz pierwszy w historii. [...] jesteśmy świadkami spektaklu zaiste tragicznego: widzimy ludzi oszalałych pod wpływem własnej inwencji, zastępujących samych siebie nieożywionymi przedmiotami. Ostatecznie jest to moralne i duchowe samobójstwo, po którym nastąpi destrukcja fizyczna.”<sup>44</sup>

Podobnie wówczas pisali angielscy dystrybucjoniści, wielki Gilbert Keith Chesterton i jego przyjaciel Hilaire Belloc, zaś 18 lat później tak pisał Wilhelm Röpke – współtwórca niemieckiego cudu gospodarczego, ekonomiczny liberał, ale też ordoliberal i orędownik ekonomii „na ludzką skalę”:

„Nie ma wątpliwości, że w końcu osiągniemy punkt, w którym niemożliwe już będzie sprawienie, by postęp techniczny nadal służył ludzkemu szczęściu w jego najszerszym i najmądrzejszym znaczeniu. [...] Nadchodzi czas, w którym trzeba będzie przypomnieć sobie, iż maszyna jest tylko instrumentem, i konieczne będzie zadanie pytania czy w interesie ludzkości „postęp” techniczny nie powinien zostać odrzucony, tak jak przedkładamy najbardziej nieracjonalną kuchnię domu rodzinnego nad najbardziej racjonalną stołówkę.”<sup>45</sup>

Ludzkie losy są złożone. Gdyby Amerykanie posłuchali Lytle’a, to w końcu XX wieku ich południowe stany byłyby republikami bananowymi. Nie należy też postępu technicznego odrzucać, ale potrzebna jest siła kultury służącej okiełznaniu tego postępu. Warto na przykład zapytać, czy w ogóle chcemy mieć roboty do złudzenia przypominające piękne dziewczyny, pełniące rolę czułych pielęgniarek w szpitalu. I należy pochylić się nad tym, co w świeżo wydanej książce pisze nie jakiś nostalgiczny marzyciel, lecz wybitny amerykański politolog, George Friedman, analizując źródła oraz fakt pogłębiającego się kryzysu społeczno-ekonomicznego oraz instytucjonalnego:

<sup>44</sup> Andrew Nelson Lytle, *The Hind Tit*, w: *I'll Take My Stand: The South and the Agrarian Tradition*, New York 1930.

<sup>45</sup> Wilhelm Röpke, *Civitas Humana: A Humane Order of Society*, Londyn 1948 (tłum. z niemieckiego); pod zmienionym tytułem *The Moral Foundations of Civil Society*, Londyn i Nowy Jork 1966.

„Adwokaci SI myślą, że byłby to [chodzi o wszechobecność SI w ludzkim świecie – przyp. JK] triumf ludzkości. Ale poza takimi problemami jak utrata stanowisk pracy oraz komputerowe błędy, SI zredukowałaby ludzką potrzebę życia we wspólnocie, jednego człowieka z drugim. Tworzymy wspólnoty z konieczności ekonomicznej i z innych konieczności. Pozostajemy związani ze sobą, ponieważ czerpiemy radość z tego, iż obydwójce jesteśmy ludźmi. Gdyby zadziałała, SI podkopałaby te konieczności, które nas wiążą. W sercu ludzkiej inteligencji leży przypadkowy, nieoczekiwany wgląd i także spotkanie. Sztuczna inteligencja, gdyby mogła być stworzona, zniszczyłaby je swoją nieludzką skutecznością.”<sup>46</sup>

### ***ARTIFICIAL INTELLIGENCE IN THE DISENCHANTED WORLD***

#### ***ABSTRACT***

This is a modest endeavour written from an engineering perspective by a non-philosopher to set things straight if somewhat roughly: What does artificial intelligence boil down to? What are its merits and why some dangers may stem from its development in this time of confusion when, to quote Rémi Brague: “From the point of view of technology, man appears as outdated, or at least superfluous”?

**Keywords:** artificial intelligence, strong artificial intelligence, machine learning, disenchanted world.

O AUTORZE — prof. dr hab. inż., prof. em. Instytutu Podstaw Informatyki PAN, ul. Jana Kazimierza 5, 01-248 Warszawa.

Email: jacek.koronacki@ipipan.waw.pl

---

<sup>46</sup> G. Friedman, *The Storm Before the Calm: American Discord, the Coming Crisis of the 2020's, and the Triumph Beyond*, Doubleday, New York 2020.