

WACŁAW ISZKOWSKI\*, RYSZARD TADEUSIEWICZ\*\*

## Na marginesie dyskusji o sztucznej inteligencji

### Wprowadzenie

O sztucznej inteligencji (SI) mówi się i pisze bardzo dużo [1], szczególnie wśród osób i instytucji traktujących Ją (pisząc ten zaimek wielką literą, chcemy podkreślić jej ważność) już jako byt istniejący, dobrze zdefiniowany i przyjmowany jako niepodlegający dyskusji. Z tego powodu publikowane pytania i stwierdzenia dotyczą wyobrażeń, jaka jest czy też ma być owa sztuczna inteligencja z punktu widzenia różnych kryteriów oraz jak może być wykorzystana dla dobra ludzkości i czy jej nie będzie w jakiś sposób zagrażać. Można z tym poglądem polemizować, ale takie są fakty.

Przypomnijmy, że najwcześniej o sztucznej inteligencji wypowiadali się informatycy, którzy przez ponad 50 lat tworzyli i doskonalili algorytmy programów komputerowych przetwarzania informacji, uzyskując wyniki coraz bardziej zbliżone do wyników działań odpowiednio kwalifikowanych ludzi. Obecnie mówią o niej wszyscy – nie zawsze do końca wiedząc, o czym mówią...

Ta obecna, rozwijająca się w różnych kierunkach dyskusja skłania nas, informatyków, do nowego spojrzenia na problem SI, nieco odmiennego od poglądów przedstawionych we wstępnym rozdziale w obszernym amerykańskim podręczniku z 2011 r. [2]. Opisywana tam sztuczna inteligencja była dobrze zdefiniowana jako podzielona na podejście oparte na naśladowaniu myślenia (obliczenia symboliczne) i wykorzystujące zbiory rozmyte i logikę rozmytą (czyli zbliżające sposób działania komputera do ludzkiego myślenia z różnymi niuansami, zamiast bardzo kategorycznych podziałów na prawdę i fałsz, właściwych dla działań komputerów) oraz podejście oparte na modelowaniu struktur wyposażonych w inteligencję (sieci neuronowe, algorytmy mrówkowe, systemy ewolucyjne). Burzliwy rozwój sztucznej inteligencji dosłownie w ostatnim roku [3] spowodował jednak konieczność wzbogacenia tego schematu.

### Spór wokół nazwy

Termin „sztuczna inteligencja” nie wszystkim się podoba. Autorzy tego artykułu także nie stronią od uwag krytycznych, chociaż przyjmują stan istniejący. Kontestując tę

---

\* Dr inż. Wacław Iszkowski, (waclaw@iszkowski.eu), Polskie Towarzystwo Informatyczne, Polska Izba Informatyki i Telekomunikacji; \*\* Prof. dr hab. Ryszard Tadeusiewicz, (rtad@agh.edu.pl), członek rzeczywisty PAN, Akademia Górniczo-Hutnicza, Kraków

nazwę, warto przypomnieć, że w okresie ujawniania społeczeństwu istnienia pierwszych elektronicznych maszyn obliczeniowych posługiwano się terminem *electronic brain*, z czego szybko (na szczęście) zrezygnowano, wyjaśniając, że maszyny te zbudowane z lamp i przełączników niewiele mają wspólnego z naszymi mózgami. Dlatego też termin *artificial intelligence (AI)*, przez lata używany wewnątrz tylko wśród zespołów zajmujących tą tematyką, stał się obecnie przez wielu opacznie rozumiany, przez nadawanie mu większego znaczenia pojęciowego, niż na to zasługuje. O ile byłoby prościej, gdyby wtedy przyjęto na przykład termin *electronic intelligence* (inteligencja elektroniczna)!

W Polsce początkowo budowano elektroniczne maszyny cyfrowe lub matematyczne. W literaturze popularnonaukowej, w mediach posługiwano się też terminem „mózg elektronowy”, co miało świadczyć o jakichś szczególnych mocach tych maszyn. Z czasem na szczęście z tego terminu zrezygnowano, chociaż i dzisiaj możemy kupić grę edukacyjną dla dzieci *Mózg Elektronowy*, mieszając maluchom w głowach. Początkowo termin „komputer” zwalczała cenzura, bo pochodził od imperialistów i dopiero Adam Empacher w 1967 roku wybronił użycie tego terminu, a językoznawcy po „odkryciu”, że w XVII i XVIII wieku mieliśmy wojska komputowe – nie zaprotestowali.

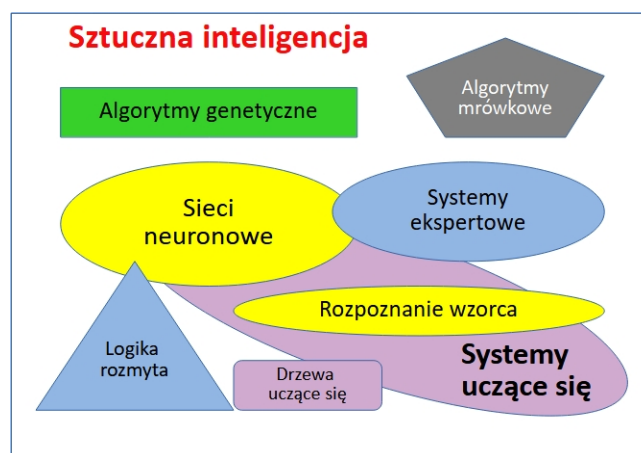
Kłopot polega w dużej mierze na tym, że termin „sztuczna inteligencja” ze skrótem SI (bo AI nie wszyscy rozumieją) jest zbyt szeroko interpretowany, gdy przecież każdy wie, że sztuczne serce nigdy w pełni nie zastąpi serca ludzkiego, choćby nawet przeszczepionego.

### **Parę słów o tym, czym jest sztuczna inteligencja**

Dla konkretyzacji dalszych rozważań korzystne będzie przytoczenie swoistej „mapy” sztucznej inteligencji przedstawionej na rysunku 1 [4]. Tworząc rysunek przyjęto pewne zasady, które warto uwzględnić podczas jego oglądania. Po pierwsze figury geometryczne symbolizujące poszczególne rozważane obszary sztucznej inteligencji mają różny rozmiar, co sygnalizuje, że jedne z tych obszarów są intensywniej rozwijane, a inne mniej. Po drugie, figury te **częściowo** pokrywają się, co oznacza, że pewne obszary mają istotne części wspólne. Na przykład obszar symbolizujący sieci neuronowe nałożony jest częściowo na obszar systemów uczących, co oznacza, że wiele systemów uczących się jest realizowanych jako sieci neuronowe, ale są takie systemy uczące się, które nie są sieciami neuronowymi, oraz są sieci, które nie są systemami uczącymi się. Figury mają różne kształty, co sygnalizuje ich różną naturę. Dominuje na nim obszar opisany jako „systemy uczące się”. Rzeczywiście – największe osiągnięcia w tej dziedzinie uzyskuje się, korzystając z tego, że odpowiednie programy są zdolne do pobierania potrzebnej wiedzy z dużych zasobów danych (których obecnie w systemach informatycznych jest mnóstwo) i do dostosowywania swojego działania do owej zdobytej wiedzy.

Na tej zasadzie zbudowany jest między innymi Chat GPT, który narobił ogromnego zamieszania w społecznej percepcji sztucznej inteligencji. W piśmiennictwie poświęconym sztucznej inteligencji popularny jest termin „uczenie maszynowe” (w skrócie ML) który dotyczy oczywiście czynności, jaka jest realizowana w systemach uczących się. Na „mapie” przedstawionej na rycinie 1 zaznaczone są różne rodzaje systemów sztucznej inteligencji, a nie czynności przez nie wykonywane. Dlatego na rycinie przyjęto taką nazwę.

Wśród systemów uczących się najbardziej istotną rolę odgrywają sieci neuronowe, będące mocno uproszczonymi modelami mózgu [5]. Obecnie sieci neuronowe są często wiązane z metodami głębokiego uczenia i termin *deep learning* jest często w użyciu. Nie zmienia to jednak faktu, że chodzi stale o sieci neuronowe, tylko liczba tak zwanych warstw ukrytych i ich rola odróżnia sieci deep learning od najpopularniejszych do niedawna sieci typu *multilayer perceptron*. Drugim ważnym obszarem systemów uczących się są metody rozpoznania wzorca, w polskiej literaturze często określane jako metody rozpoznawania obrazów. Dziedzina ta rozwinęła się jako jeden z najwcześniejszych działów sztucznej inteligencji [6]. Wynikało to z faktu, że rozpoznanie obrazów (głównie fotografii satelitarnych) było silnie wspierane finansami wojskowymi ze względu na przydatność w szpiegostwie. Potem jednak stwierdzono, że te same metody automatycznie wyuczonej klasyfikacji mogą służyć do rozpoznawania różnych wzorców (na przykład w ekonomii [7]).



Ryc. 1. Mapa sztucznej inteligencji [4]

Z systemami uczącymi się korespondują uczące się drzewa (służące do wspomagania procesów podejmowania decyzji), systemy ekspertowe (narzędzia mogące doradzać użytkownikom podobnie, jakby to mógł czynić człowiek będący ekspertem w jakiejś dziedzinie), zaś dla zbliżenia sposobu przetwarzania informacji w programie aspirują-

cym do miana inteligentnego wprowadzono logikę rozmytą. Kształt fiury symbolizującej logikę rozmytą zasadniczo różni się od figur syblizujących metody sztucznej inteligencji wcześniej omawiane, bo sposób naukowego wprowadzenia i praktycznego wykorzystywania logiki rozmytej jest zasadniczo odmienny od tych wszystkich wcześniej omówionych metod. To samo dotyczy kształtów pozostałych figur użytych na rysunku. Dodatkowo można zobaczyć algorytmy genetyczne, pozwalające uzyskać optymalne rozwiązanie problemu decyzyjnego na zasadzie symulowanej w komputerze ewolucji całej populacji różnych metod rozwiązywania, a także algorytmy mrówkowe, które imitują „inteligencję roju” – zdolność do uzyskiwania inteligentnego działania zbiorowości osobników, z których żaden z osobna inteligencji nie wykazuje.

### Jak to się zaczęło?

Cofnijmy się do XVII wieku, kiedy w 1613 roku pojawił się termin *computer* oznaczający osobę, która liczy. Komputery ludzkie były wykorzystywane do obliczeń matematycznych i sporządzania tabel matematycznych, pogodowych, artyleryjskich itp. Coraz częściej były zatrudniane komputery żeńskie, szczególnie podczas II wojny światowej. Wtedy też rozpoczął się rozwój, technicznie nazywanych, *Electronic Calculating Machines*. Komputerki (kobiety) zostały zwolnione do domu, ale część z nich stała się programistkami ręcznie wtykającymi kabelki połączeń w maszynie ENIAC (tak na początku programowano!).

Na temat maszyny myślącej, od początku istnienia tej idei, istniały różnice zdań. Pierwszym naukowcem, który sformułował koncepcję sztucznej inteligencji, był Alan Turing, który w październiku 1949 roku napisał pionierską pracę *Computing Machinery and Intelligence*. Praca ta, opublikowana w 1950 roku, w czasopiśmie „Mind” wydawanym przez Oxford University Press) jako pierwsza wskazywała, że komputery (których było wtedy zaledwie kilka na świecie) obok obliczeń numerycznych mogą także przetwarzać informacje podobnie jak ludzie. W tym historycznym artykule zdefiniowany został między innymi tak zwany (potem) „test Turinga”, czyli empiryczna metoda orzekania o tym, czy to, co robi maszyna, jest odpowiednikiem ludzkiego myślenia [8], czy też jest to wyłącznie sztuczne żonglowanie informacjami i pojęciami.

W niemal każdym opracowaniu dotyczącym sztucznej inteligencji można dziś przeczytać, że nazwa i wstępne sformułowanie zakresu tej dziedziny nastąpiło w 1956 roku podczas konferencji w Dartmouth (USA), co należy przyjmować z pewnym dystansem wobec wcześniejszej koncepcji Turinga. Konferencja zwołana z inicjatywy Johna McCarthy’ego miała jedno zadanie: zdefiniowanie celów i granic obszaru programów zdolnych do inteligentnego działania i ustalenie dla nich jednej wspólnej nazwy. Na temat autorstwa tej nazwy trwał potem przez wiele lat spór. Przypuszczalnie termin *Artificial Intelligence* zdefiniowany jako *obliczeniowa część zdolności do osiągnięcia celów na świe-*

cie zaproponował zaraz na początku John McCarthy. Ale też bardzo aktywny uczestnik konferencji, późniejszy pionier metod uczenia maszynowego, Marvin Minsky, uporczywie twierdził, że to był jego pomysł, definiując *sztuczną inteligencję jako zdolność do rozwiązywania trudnych problemów*. Trzeba przyznać, że w Dartmouth 20 uczonych w osiem tygodni naprawdę zrewolucjonizowało rozumienie relacji między ludźmi a maszynami [9]. Uczestnicy konferencji byli też przekonani, że maszyny myślące zostaną zbudowane w przeciągu najbliższych lat. W istocie nie zostały zbudowane do dziś.

### **Początki budowania zrębów sztucznej inteligencji w Polsce**

Sięgnijmy do wyników badań prowadzonych jeszcze w latach 70. XX wieku w Polsce. Oczywiście będzie to ogląd bardzo subiektywny, zaś odmiennie owe początki opisane są w artykule [10]. W tym artykule warto także wspomnieć, że w pracy [11] opisano jeden z najwcześniejszych w Polsce systemów automatycznego rozpoznawania mowy. W pracy [12] pokazano, że istnieje możliwość skorzystania z komputera przy znajdowaniu optymalnej terapii medycznej, w pracy [13] nakreślono koncepcję inteligentnego systemu doradczego dla sędziów, a w pracach [14] i [15] opisano pierwszy polski działający system prowadzący tekstową konwersację z człowiekiem w języku polskim, czyli był to taki pra-prototyp systemu podobnego do Chat GPT.

Do prac tych kreatywnych informatyków dołączali użytkownicy pierwszych, powiedzmy sztucznie inteligentnych systemów komputerowych wspomagających ich swoją rzeczywistą inteligencją w różnych dziedzinach wiedzy. Można zaryzykować twierdzenie, że przez blisko pół wieku można było rozwijać prace badawcze pod nazwą *sztuczna inteligencja*, bo wykonywali je i wykorzystywali ludzie, którzy naprawdę wiedzieli, co się za tym kryje.

Również w świecie naukowym i z czasem w innowacyjnych firmach przez te lata pracowano nad rozwojem różnych metod uzyskiwania coraz wyższego potencjału wiedzy pozyskiwanej metodą uczenia z coraz większych zbiorów danych. Główną przeszkodą w szybkim rozwoju tych metod były zbyt małe moce przetwarzania dostępnych komputerów oraz wysokie koszty pamięci komputerowej, coraz bardziej potrzebnej o większych pojemnościach. Pewnym przełomem w uzyskaniu pożądanej mocy obliczeniowej było zastosowanie z początkiem lat 20. procesorów graficznych. Z czasem były też dostępne taniejące moduły pamięci o coraz większych pojemnościach, liczonych już w petabajtach (1PB jest równy 1 mln GB lub 1 Mld MB). Przed sztuczną inteligencją szeroko otworzyły się drzwi nowoczesnej technologii informatycznej.

### **Nowa sytuacja**

Niespodziewanie zaistniała nowa sytuacja, gdy w listopadzie 2022 roku do powszechnego użytku został udostępniony za darmo Chat GPT-3.5 (czat wstępnie prze-

szkolonego transformera generatywnego), opracowany przez firmę OpenAI [16]. Jego fenomen opiera się na tym, że użytkownik, mając tylko podstawowe umiejętności zadawania pytań w jego własnym języku, może uzyskać na nie odpowiedzi oparte na prawie pełnej aktualnie dostępnej wiedzy na ten temat. To spowodowało, że tym narzędziem, udostępnionym ludziom całego świata pozwolono się pobawić, porozmawiać z czatem, zarejestrować przykładowe dialogi i potem pokazywać to naokoło. W dwa miesiące narzędziem tym zainteresowało się ponad 100 mln osób na całym świecie. Początkowy entuzjazm „badaczy” wyposażonych w podpowiedzi sztucznej inteligencji, którzy niewiele wiedząc, mogli, zadając czatowi GPT pytania i zadania, pochwalić się „własną inteligencją”, dość szybko przeminął. Za to firma OpenAI skrzętnie zbierała wszelkie uwagi, opinie i komentarze, doskonaląc swój produkt [17].

Sztuczna inteligencja dzięki systemowi GPT-3.5 stała się bezspornie popularna. Ale to nie jest jej prawdziwa istota. To jest wersja dla ludzi ciekawych nowości, marketingowa. Konieczne jest poszerzenie dyskusji w kierunku jej praktycznych rozwiązań i zastosowań profesjonalnych, bo te dopiero będą miały istotne znaczenie gospodarcze i społeczne. Natychmiast pada też pytanie o możliwe zagrożenia wynikające z jej coraz szerszego wykorzystywania. Systemy sztucznej inteligencji są bowiem w istocie papugą stochastyczną, bo bezmyślnie powtarzają fragmenty tekstów wydobytych z analizowanych zasobów danych, kierując się głównie wskaźnikami statystycznymi i generując odpowiedź na podstawie poszukiwania kolejnego elementu pasującego do poprzednich z największym prawdopodobieństwem.

W marcu 2023 roku firma OpenAI udostępniła kolejną wersję Chat GPT-4, wyposażonego w znacznie bogatszą wiedzę. Tym razem wersja ta jest już dostępna tylko w subskrypcji za 20\$/mc. Również inne firmy uruchomiły swoje podobne systemy SI: Bard z Google’a, Bing z Microsoft, Codewhisper z Amazon, Llama z Meta (FB) oraz wiele innych, w tym specjalistycznych w różnych dziedzinach [16]. Mamy też już polskie rozwiązania w opracowaniu: GPT z Sentione oraz Trurl z VoiceAI. Część z nich jest bezpłatna, ale też mają różne mniejsze zasoby wiedzy.

### **Zajrzyjmy za kurtynę – czyli jak to działa?**

Dokładne wyjaśnienie zasad działania modelu językowego GPT wymagałoby sporo miejsca, a i tak nie wszystko byłoby możliwe do pokazania, gdyż wiele szczegółów jest tajemnicą firmy. W skrócie jest to duży model językowy, który jest „uczony” danymi z wielu źródeł, przeważnie dostępnych w sieci przez Internet. Głównym źródłem są artykuły z Wikipedii, książki z biblioteki Google Books, strony internetowe z Google Search, posty na Facebooku oraz świergot z Twittera (obecnie X). Tak, to **nasza wiedza** jest przejmowana przez taki model.

Pierwszy, wydany w 2018 roku, wstępnie przeszkolony transformer GPT-1 przetwarzał język naturalny na podstawie 117 mln parametrów pozyskanych z danych pobranych ze stron internetowych oraz zbioru książek.

Drugi GPT-2 z 2019 roku zawierał 1,5 mld parametrów i już pozwalał na proste spójne konwersacje, ale miał kłopoty z dłuższymi fragmentami.

GPT-3, wydany w 2020 roku, miał już 175 mld parametrów i skorzystał do swej nauki z zasobów Wikipedii, ale dość łatwo prezentował nieprawdziwe lub stronnicze informacje, zwane halucynacjami, przez brak możliwości ich weryfikacji. Na rynku udostępniono nieco poprawioną wersję GPT-3.5

I najnowszy GPT-4, korzystający (podobno) z 1,7 biliona parametrów, uzyskanych z 45 terabajtów danych źródłowych, na których był trenowany, weryfikowany oraz testowany do marca 2023 roku. Podobno szkolenie GPT-4 kosztowało 100 mln dolarów, głównie kosztów pracy komputerów.

Niestety nie udało się w pełni zweryfikować wszystkich zależności w modelu językowym i GPT często wykazuje halucynacje, czyli może prezentować wyniki niemające potwierdzenia w danych źródłowych. Obecnie jest największym modelem językowym, ale też cały czas się doucza, czerpiąc wiedzę z aktualności w Internecie oraz z poprawek dostarczanych przez jego użytkowników.

Podstawową zasadą działania GPT jest formułowanie odpowiedzi poprzez przyrostowe uzupełnianie pierwotnego tekstu pochodzącego z zapytania. Kolejne elementy są wybierane na podstawie wag parametrów, przy czym ich wartości i ich dostrajanie jest pilnie strzeżoną tajemnicą producenta.

Jest jednak problem moralny, który będzie wymagał interwencji prawnej. Systemy GPT korzystają do woli z dorobku społeczności, która przez lata gromadziła swoją wiedzę, artykuły, posty, dokumenty, opisy, zdjęcia w sieci internetowej. I teraz za darmo dorobek ten zostaje przejęty dla tworzenia wielkich baz danych modeli językowych. Ciekawe, ile wart jest ten zasób?

Kluczowym źródłem są zasoby Wikipedii. Obecnie Wikipedia zawiera ponad 60 mln artykułów w 335 językach (w tym dialektach), najwięcej 6,7 mln w języku angielskim, a 1,5 mln w języku polskim. Zasadą Wikipedii jest publikowanie treści w formie encyklopedycznej, opracowywanych wyłącznie na podstawie zweryfikowanych źródeł, przeważnie odnajdywanych w Internecie, ale również w postaci papierowej. Każdy z ponad 100 mln zarejestrowanych użytkowników może opublikować artykuł lub poprawić fragmenty już opublikowanego artykułu. Działania te muszą być zatwierdzone przez jednego z 200 tys. redaktorów. A jeszcze istnieją administratorzy do rozstrzygania spraw spornych. Przyjmuje się, że większość treści jest prawdziwa, chociaż istnieją też treści niezweryfikowane lub umyślnie zniekształcone (tzw. wandalizmy). Nikt z autorów, redaktorów, administratorów nie jest wynagradzany – dostęp do Wikipedii jest też darmowy.

Każdy ma prawo do pobierania treści z Wikipedii, ale z podaniem źródła. Wikipedia nie publikuje żadnych reklam, a utrzymuje się z funduszy Wikimedia Foundation, a ta z donacji użytkowników, grantów oraz firm. Fundacja prowadzi jeszcze kilkanaście innych specjalistycznych projektów gromadzących informacje.

Przejmując zasoby Wikipedii i innych źródeł, firma OpenAI (twórca kolejnych wersji GPT) wywołuje pytania o prawa majątkowe i autorskie, gdyż w modelu językowym „ginie” identyfikacja źródeł. Firmy SI nie zamierzają też nikomu płacić za pobrane treści, podczas gdy parę lat temu Google był zmuszany do wnoszenia opłat za linkowanie do treści witryn w Internecie. Pojawiają się też istotne zastrzeżenia, że systemy SI będą „zabijać” Wikipedię i inne informacyjne witryny, gdyż użytkownicy zawarte tam informacje otrzymają bezpośrednio z systemu SI. Witryny te tracą odwiedzających, a tym samym utracą dochody z reklam.

Drugim zagrożeniem dla Wikipedii i innych witryn jest obecnie jeszcze często mała wiarygodność publikacji pochodzących z systemów SI, gdyż są one składane z wymieszanych w modelu językowym danych. Te publikacje z możliwymi halucynacjami oraz szybkość ich powstawania mogą prowadzić do „zatrucia” prawdziwości informacji publikowanych w Internecie, a to może wpływać na jakość informacyjną artykułów w Wikipedii i w innych portalach, w rezultacie również w systemach SI. Warto dodać, że terminu „zatrucie” częściej używa się do wskazywania informacji celowo fałszywych, produkowanych na skutek uczenia systemów SI na bazie specjalnie preparowanych baz danych, ale to jest temat na osobne opowiadanie.

Znaczącym zastrzeżeniem są obawy twórców: dziennikarzy, autorów książek, scenarzystów, copywriterów itp., że systemy SI będą ich zastępować, najpierw gromadząc ich dorobek w modelu językowym, a potem tworząc na życzenie „własne” przetworzone utwory. Z setek scenariuszy filmów można wygenerować co najmniej kilka „nowych”.

Wydaje się, że droga prawna do dyskusji z firmami SI jest otwarta. Dokonano już pierwszego ustalenia, że systemy SI nie są autorami, a więc nie przysługują im prawa majątkowe i autorskie do generowanych przez nich utworów: tekstów, obrazów, obiektów itp. Ale jeżeli publikacja wygenerowanej przez system SI treści nie narusza czyjegoś prawa, to czy przez to nie będą wyciekać treści, do których ktoś ma te prawa?

I drugie ciekawe pytanie, czy treści wygenerowane przez system SI mogą być potraktowane jako plagiat innych treści? Zostawmy to prawnikom.

### **Chwasty do wyplewienia**

Termin „sztuczna inteligencja” stał się tak nośny, że pokusie formułowania różnych sensacyjnie wyglądających teorii ulegli badacze mający autorytet w różnych dziedzinach, którzy jednak nigdy nie zajmowali się tworzeniem ani merytorycznym stosowaniem sztucznej inteligencji. Otóż badacze ci z wielką pewnością siebie lansowali różne, często



dość słabo uzasadnione poglądy. Na temat rzekomych zagrożeń, jakie dla całej ludzkości ma nieść rozwój sztucznej inteligencji już w 2015 roku wypowiadał się bardzo kategorycznie znany astrofizyk Stephen Hawking. Był on jednym z sygnatariuszy listu otwartego *Future of Life Institute*, apelującego o zaniechanie prac nad sztuczną inteligencją. Warto dodać, że w niektórych debatach na temat sztucznej inteligencji wprowadzany jest podział na tak zwaną „silną sztuczną inteligencję” (całkowicie równoważną inteligencji ludzkiej) oraz na tak zwaną „słabą sztuczną inteligencję” (zdolność do sprawnego rozwiązywania praktycznych problemów). Dla inżynierów budujących systemy sztucznej inteligencji – a takimi są autorzy tego artykułu – ważna jest tylko ta słaba sztuczna inteligencja.

Jeszcze bardziej rewolucyjne poglądy prezentował Roger Penrose, laureat Nagrody Nobla za teoretyczny opis czarnych dziur w kontekście teorii względności i kosmologii. Nie ulega wątpliwości, że w fizyce jest on wielkim odkrywcą – i tam należy uznać jego kompetencje. Ale gdy Penrose zaczął publikować także na temat filozofii umysłu, dość bezceremonialnie wypowiadając się na temat sztucznej inteligencji, to można było mieć wątpliwości. W 1989 roku opublikował on książkę *Nowy umysł cesarza*, gdzie starał się dowieść, że stworzenie sztucznej inteligencji jest wykluczone. Książka ta była mocno krytykowana, więc Penrose wydał w 1994 roku jeszcze bardziej radykalne *Cienie świadomości*, a potem pracę zbiorową *Makroświat, mikroświat i ludzki umysł*.

Badacze zajmujący się mózgiem też czasem wypowiadają dość kontrowersyjne tezy na temat sztucznej inteligencji. Giulio Tononi z Uniwersytetu Wisconsin jest neurobiologiem i psychiatrą, więc z inteligencją w jej naturalnej (ludzkiej) postaci jest bardzo dobrze zaznajomiony. Jednak bardziej niż w medycynie znany jest ze swojej *Zintegrowanej Teorii Informacji* (IIT) oraz matematycznej teorii świadomości, której symbolem jest Phi. Na ten temat, razem z Geraldem Edelmanem, wydali w 2000 roku książkę zatytułowaną *Świadomość: Jak materia staje się wyobraźnią*. Gerald Edelman, jest również laureatem Nagrody Nobla, którą otrzymał w 1972 roku za odkrycie struktury cząsteczek chemicznych białek przeciwciał odgrywających kluczową rolę w mechanizmach odpornościowych. Z całym szacunkiem trzeba stwierdzić, że do wypowiedzianych opinii na temat świadomości, a w dalszej kolejności także sztucznej inteligencji to nie upoważnia.

Warto dodać, że w następstwie euforii wokół Chat GPT-3.5 pojawił się w marcu 2023 list otwarty *Pause Giant AI Experiments* [18], podpisany między innymi przez niebyle kogo: Elona Muska, Steve’a Wozniaka, Yuvala Noaha Harariego oraz przez 33 711 osób [dane z 22.09.2023]. Abstrahując od faktycznych intencji autorów listu, zawarte w nim treści są naiwne, a jednocześnie świadczą o braku rzetelnej wiedzy, czym tak naprawdę są techniki sztucznej inteligencji oraz czym są generatory tekstów, które to mają tak bardzo zagrażać ludzkości, że trzeba aż wstrzymać prace nad ich rozwojem na

6 miesięcy! [19]. Obecne możliwe efekty działania systemów SI nie zagrażają społeczeństwom, bo oprócz „zalewania” ekranów swoją twórczością, nie mają bezpośredniego wpływu na jakiegokolwiek decyzje czy działania, które mogą dotyczyć pojedynczych osób lub większych grup społecznych. Takie stwierdzenie ma sens tak długo, póki jacyś mniej rozumni użytkownicy tych systemów, mający zdolności decyzyjne w jakimś zakresie, nie zaczną bezkrytycznie korzystać z uzyskiwanych wyników tych generatorów do podejmowania i wdrażania w życie podpowiedzianych, ale swoich – podkreślamy: swoich – decyzji. Można tu wspomnieć o fakcie, że dla większości użytkowników systemy sztucznej inteligencji dostępne są na zasadzie „czarnej skrzynki”. Widoczne są zadawane pytania oraz dostępne są gotowe odpowiedzi – natomiast użytkownik nie może prześledzić, co się tak dzieje „w środku”. Dlatego pojawiają się skrajne opinie, że sztuczna inteligencja po prostu nie istnieje!

W tym artykule nie uda się definitywnie przesądzić, czy sztuczna inteligencja rzeczywiście istnieje, z całą pewnością można jednak wykazać, że istnieje pragnienie sztucznej inteligencji.

### **Pradawne pragnienie sztucznej inteligencji**

Ludzie już w starożytności marzyli o robotach, które będą mogły zastępować ich w pracy. Biorąc pod uwagę fakt, że roboty (te antropomorficzne, przeznaczone do bezpośredniej współpracy z ludźmi, a nie te czysto mechaniczne, wykorzystywane w przemyśle, zwłaszcza motoryzacyjnym) zwykle wyobrażamy sobie jako wyposażone w sztuczną inteligencję, można zaryzykować stwierdzenie, że początki SI miały miejsce już w starożytności [20].

W wielu mitach greckich wspomniane są maszyny, które z dzisiejszego punktu widzenia nazwalibyśmy robotami. Homer w *Iliadzie* (VIII w p.n.e.) opisuje dwie złote dziewice – maszyny zbudowane przez Hefajstosa, które potrafiły chodzić, obsługiwać uczującym oraz śpiewać. Nawet dziś zbudowanie takich robotów byłoby trudne, ale Hefajstos był bogiem, więc sobie z tym poradził 8 wieków przed Chrystusem.

Robotem ze sztuczną inteligencją była też ulepiona z gliny i wody Pandora (oczywiście także twór Hefajstosa), która jednak tej inteligencji miała za mało i otworzyła puszkę, a z niej wydostały się wszystkie nieszczęścia, które nękają ludzi do dziś. Można powiedzieć, że to był pierwszy przypadek zaszkodzenia ludziom przez sztuczną inteligencję.

Przykłady starożytnych robotów (oczywiście wymyślonych), wyposażonych w sztuczną inteligencję, można mnożyć. Wybrzeży starożytnej Krety strzegł Talos, olbrzym z brązu (też dzieło Hefajstosa), który obiegał wyspę dookoła i ciskał kamieniami w zbliżające się statki. Rozpoznawał statki przyjazne i wrogie i rzucał bardzo celnie ogromne głazy. Niepokojąco podobny do niektórych współczesnych robotów bojowych!

Unieszkodliwiła go Medea, która chciała ułatwić lądowanie na Krecie Jazona i Argonautów, więc mamy tu do czynienia z przykładem, że korzyści ze sztucznej inteligencji mogą być unicestwione przez ludzi realizujących swoje partykularne interesy.

Te przykłady sięgające starożytności odzwierciedlają ludzkie marzenia o zbudowaniu maszyny myślącej i działającej za nas.

### Co potrafi inteligentny awatar?

W sytuacji gdy wiele kontaktów między ludźmi oraz instytucjami odbywa się za pomocą komunikacji elektronicznej, stworzenie informatycznego alter ego jest bardzo atrakcyjne. W zakresie komunikacji alfanumerycznej (mejl, sms, komunikacja z serwisami firm, aktywność w mediach społecznościowych) różnych narzędzi automatycznie generujących treści jest już co niemiara. W imieniu człowieka, naśladowując jego zachowanie, działają boty rozsiewające informacje lub generujące pozorowane duże poparcie dla jakichś oficjalnie publikowanych treści oraz wiadomości fałszywe i zwykle szkodliwe (*fake-news*). Są też boty pozytywne, powielające i rozsyłające ostrzeżenia meteorologiczne, komunikujące o utrudnieniach w transporcie, albo też powiadamiające o różnych wydarzeniach – na przykład konferencjach naukowych.

Znacznie bogatsze możliwości oferują tworzone przez sztuczną inteligencję awatary. Obecnie możliwe jest stworzenie wizualizacji, na której wygenerowana sylwetka imitowanego określonego człowieka będzie poruszała się zgodnie z poleceniami pochodzącymi od sterującego nią programu. Obejmują między innymi sztucznie imitowaną mimikę twarzy i ruchy artykulacyjne związane ze sztucznie wytwarzaną mową i to nawet charakterystyczną dla danego języka. Naśladownictwo jest tak doskonałe, że rozmówca potraktuje kontakt z awatarem jako rozmowę z rzeczywistym rozmówcą. W efekcie możliwe jest przekazanie treści, z pozoru prawdziwej, żeby odbiorcę zmanipulować w określonym celu. Wykrycie takiego fałszerstwa jest bardzo trudne, bo nieufny odbiorca nie może równocześnie sprawdzić w innych źródłach wiarygodności słownego przekazu.

Użycie awatarów budzi niepokój. Znany jest eksperyment, jaki przeprowadziła Joanna Stern, dziennikarka z *Wall Street Journal* [21], która stworzyła za pomocą programu *Synthesia* swojego cyfrowego sobowtóra. Ten awatar zastępował ją podczas prowadzonych w trybie zdalnym posiedzeń redakcji, a także logował się w jej imieniu do banku i wydawał dyspozycje. O ile koledzy w redakcji byli trochę zdziwieni niektórymi zachowaniami i wypowiedziami awatara i byli bliscy odkrycia fałszerstwa, o tyle owe dyspozycje system bankowy przyjmował i realizował.

### Prawne ograniczenia eksploatacji systemów SI

W poprzednich punktach pokazaliśmy różne aspekty systemów SI – bardziej te, które niepokoją społeczeństwa, a tym samym polityków. Niewiele wspominaliśmy o możli-

wych pozytywnych zastosowaniach tych systemów obecnie i w przyszłości, gdyż takich możliwości jest i będzie znacząco dużo. Jednakże będzie to możliwe tylko wtedy, gdy zaniepokojone społeczeństwa nie zażądadają nałożenia zbyt restrykcyjnych prawnych ograniczeń na eksploatację i rozwój systemów SI.

W opracowaniu i zatwierdzaniu są takie regulacje mające ograniczać prawdopodobne negatywne skutki działania systemów SI. Takie działania są prowadzone równoległe w Kongresie i Białym Domu US, w gremiach Komisji i Parlamentu Europejskiego, w rządzie Chin i w innych krajach. W każdym przypadku przyjmowane są nieco inne założenia wyjściowe – w USA celem jest zachowanie wiodącej pozycji amerykańskich korporacji, w UE ochrona danych osobowych obywateli przed ich wyprowadzeniem z Europy, a w Chinach uzyskanie nowych możliwości inwigilacji obywateli. Otwarte pozostaje pytanie, co z innymi regionami świata – z Indiami, Izraelem, krajami arabskimi czy afrykańskimi? Przecież tam również istnieje dostęp do Internetu, a wiele z systemów SI potrafi komunikować się w wielu językach.

W tych pracach prym wiodą politycy, prawnicy, socjologowie, ekonomiści oraz lobbyści korporacji zajmujących się SI – informatyków raczej nikt tam o zdanie nie pyta, a to jednak sami informatycy powinni się czuć co najmniej szczególnie zaniepokojeni. Bo też w tych regulacjach to od nich oczekuje się takiego opracowania systemów, aby można było kontrolować, według jakich zasad działają, generując kolejne wyniki, a to jest zbyt daleko idące oczekiwanie. Natomiast możliwa jest tak zwana wyjaśnialna sztuczna inteligencja (*explainable AI*), w której użytkownik po otrzymaniu potrzebnej odpowiedzi może zadać pytanie: „A dlaczego tak?” – i otrzyma w miarę wyczerpującą odpowiedź.

W USA, oprócz kilku aktów prawnych mających regulować korzystanie z systemów teleinformatycznych przez agencje federalne, administracja podejmuje też działania w zarządzaniu ryzykiem stwarzanym przez systemy w ochronie praw i bezpieczeństwa Amerykanów. Przytoczymy oficjalną definicję terminu „sztuczna inteligencja” przyjętą w „National Artificial Intelligence Initiative Act of 2020” [22]:

Sztuczna inteligencja oznacza system maszynowy, który w przypadku danego zestawu celów określonych przez człowieka może prognozować, rekomendować lub podejmować decyzje mające wpływ na środowiska rzeczywiste lub wirtualne. Systemy sztucznej inteligencji korzystają z danych wejściowych pochodzących od maszyn i ludzi, aby: (A) postrzegać środowiska rzeczywiste i wirtualne; (B) streszczać takie spostrzeżenia w modelach poprzez analizę w sposób zautomatyzowany; oraz (C) używać wnioskowania modelowego do formułowania opcji informacji lub działań.

W lipcu 2023 roku odbyło się w Białym Domu [23] spotkanie z wiodącymi w rozwoju SI firmami (przedsiębiorstwami): Amazon, Anthropic, Google, Inflection, Meta, Micro-

soft i OpenAI, aby ogłosić sygnowanie dobrowolnych zobowiązań tych firm do przestrzegania bezpiecznego i przejrzystego rozwoju technik sztucznej inteligencji. I tak:

Firmy zobowiązują się do wewnętrznych i zewnętrznych testów bezpieczeństwa ich systemów SI przed ich udostępnieniem. Nie ma wielu szczegółów co do wymaganych testów. Testowanie będzie częściowo przeprowadzone przez niezależnych ekspertów, którzy skupią się na zagrożeniach bezpieczeństwa biologicznego, cyberbezpieczeństwa oraz szerszych skutków społecznych. Oczekiwane jest, żeby to rząd federalny sfinansował tego rodzaju testy, które mogą być kosztowne i będą wymagać zatrudnienia inżynierów o znacznej wiedzy technicznej. Obecnie wiele testów bezpieczeństwa jest finansowanych i nadzorowanych przez firmy, co rodzi oczywiste pytania o konflikt interesów.

Firmy zobowiązują się do wymiany informacji w całej branży oraz z rządami, społeczeństwem obywatelskim i środowiskiem akademickim na temat zarządzania sztuczną inteligencją. Celem Białego Domu jest zachęcanie firm do wymiany informacji między sobą na temat ryzyka, jakie stwarzają (lub nie) ich modele.

Firmy zobowiązują się do inwestowania w zabezpieczenia cybernetyczne i przed zagrożeniami wewnętrznymi w celu ochrony zastrzeżonych i niepublikowanych parametrów modelu systemu SI, przed szpiegostwem przemysłowym.

Firmy zobowiązują się do ułatwiania stronom trzecim wykrywania i zgłaszania luk w zabezpieczeniach ich sztucznej inteligencji. Nie jest jasne, jak Biały Dom widzi procedury skutecznego mechanizmu zgłaszania tych luk.

Firmy zobowiązują się do opracowania solidnych mechanizmów technicznych (np. „systemu znaków wodnych”), aby użytkownicy oraz inni odbiorcy wiedzieli, kiedy dana treść jest efektem działania systemu SI.

Firmy zobowiązują się do publicznego zgłaszania możliwości, ograniczenia oraz obszarów właściwego i niewłaściwego użytkowania swoich systemów SI. Nie jest określone, jak często i w jaki sposób miałyby to czynić.

Firmy zobowiązują się do nadania priorytetu badaniom dotyczącym zagrożeń społecznych, które mogą stwarzać systemy SI, w tym unikanie szkodliwych uprzedzeń i dyskryminacji oraz ochrona prywatności. A równocześnie zobowiązują się do opracowania i wdrożenia zaawansowanych systemów SI, które pomogą sprostać największym wyzwaniom społeczeństwa. Tutaj Biały Dom wymienia zwalczanie nowotworów oraz łagodzenie zmian klimatycznych – obszary, w których chciałby widzieć największy postęp.

Ogólnie rzecz biorąc, umowa Białego Domu z siedmioma firmami SI wydaje się bardziej symboliczna niż merytoryczna. Nie ma mechanizmu egzekwowania, który zapewniłby, że firmy będą przestrzegać tych zobowiązań. Same muszą zadbać, aby się do nich stosować, pilnując się wzajemnie. A co z innymi firmami (nie tylko z USA), które rozpoczęły prace nad aplikacjami systemów SI? Wyraźnie też widać, że Biały Dom chce wspomagać rozwój amerykańskich firm nie tylko w USA, ale również na świecie.

W Komisji Europejskiej w 2021 roku powstał projekt Rozporządzenia Parlamentu Europejskiego i Rady ustanawiający zharmonizowane przepisy dotyczące sztucznej inteligencji (Akt w sprawie sztucznej inteligencji – ang. AI Act) [24]. Celem tego projektu jest *wypracowanie odpowiednio wyważonego podejścia (...) aby Europejczycy mogli korzystać z nowych technologii opracowanych i funkcjonujących zgodnie z unijnymi wartościami, prawami podstawowymi i zasadami*. Celami szczegółowymi jest zapewnienie wprowadzania bezpiecznych i zgodnych z unijnymi wartościami systemów SI do obrotu w Unii oraz ułatwienie inwestycji i innowacji w tej dziedzinie. Celem nadrzędnym jest też zapewnienie bezpieczeństwa obywateli przy równoprawnych ułatwieniach rozwoju jednolitego rynku.

W podstawowym paragrafie projektu zamieszczono wykaz zakazanych praktyk w działaniu systemów SI ze względu na fakt, że naruszają one prawa podstawowe. Zakazane są praktyki manipulowania ludźmi z wykorzystaniem technik podprogowych działających na ich podświadomość lub wykorzystujących ich słabości mogące spowodować szkody psychiczne i fizyczne. Równie zakazane są praktyki naruszające przepisy ochrony danych osobowych, konsumentów i usług cyfrowych, z obowiązkiem odpowiedniego informowania osób, z zapewnieniem im swobody decydowania o swoim zachowaniu. W kolejnym paragrafie zdefiniowano systemy SI wysokiego ryzyka mogące zagrażać zdrowiu lub bezpieczeństwu użytkowników. Takie systemy mogą być wprowadzone do obrotu tylko po weryfikacji zgodności z ustanowionymi zasadami przez podmioty zewnętrzne dla firmy je wprowadzającej. Dalsze przepisy precyzują zasady wdrażania i kontrolowania tych systemów, ustanawiając wysokie kary za ich nieprzestrzeganie.

Rozporządzenie to jeszcze nie zostało formalnie przyjęte, ale już pewien polski obywatel złożył skargę do Urzędu Ochrony Danych Osobowych (UODO) na działanie GPT-4, który jego zdaniem naruszył prawo unijne, prezentując błędnie jego dane osobowe oraz nie podał skąd je pobrał [25]. UODO zamierza się w tej sprawie zwrócić do OpenAI (chyba w USA, bo ta firma nie ma przedstawicielstwa w Europie) oczekując wyjaśnień. Wydaje się, że nie będzie to prosta sprawa. Nie znamy jej szczegółów, ale możliwe, że bot GPT miał trudności z identyfikacją tego obywatela, którego nazwisko nosi w Polsce ponad 8000 mężczyzn, a razem z imieniem jest ich aż 10 na LinkedIn. W każdym razie mamy przedsmak tego, czego możemy się spodziewać w przyszłości.

Tutaj chwilę jeszcze zatrzymajmy się na definicji „systemu sztucznej inteligencji”, która w projekcie brzmi:

„system sztucznej inteligencji” oznacza oprogramowanie opracowane przy użyciu co najmniej jednej spośród technik i podejść wymienionych w załączniku I, które może – dla danego zestawu celów określonych przez człowieka – generować wyniki, takie jak treści, przewidywania, zalecenia lub decyzje wpływające na środowiska, z którymi wchodzi w interakcję;

Z kolei załącznik I wymienia:

- a) mechanizmy uczenia maszynowego, w tym uczenie nadzorowane, uczenie się maszyn bez nadzoru i uczenie przez wzmacnianie, z wykorzystaniem szerokiej gamy metod, w tym uczenia głębokiego;
- b) metody oparte na logice i wiedzy, w tym reprezentacja wiedzy, indukcyjne programowanie (logiczne), bazy wiedzy, silniki inferencyjne i dedukcyjne, rozumowanie (symboliczne) i systemy ekspertowe;
- c) podejścia statystyczne, estymacja bayesowska, metody wyszukiwania i optymalizacji.

Jest to definicja głęboko techniczna, odwołująca się do klasycznych technik realizacji systemów SI. Dla nie techników była trudna do zrozumienia, co się potem szybko okazało w dyskusjach nad projektem tego rozporządzenia.

W dniu 14 czerwca 2023 roku Parlament [26] przegłosował 771 poprawek do tego projektu rozporządzenia Parlamentu Europejskiego i Rady. Poprawki te były wynikiem dyskusji w wielu gremiach. W tych poprawkach dwie z nich są szczególnie istotne, gdyż znacząco zmieniają postrzeganie systemów sztucznej inteligencji.

Pierwsza z nich wprowadza całkowicie nową definicję:

„system sztucznej inteligencji” (system SI) oznacza system oparty na maszynach, zaprojektowany do działania z różnymi poziomami autonomii i który może, dla celów jawnych lub ukrytych generować wyniki, takie jak przewidywania, zalecenia lub decyzje, które mają wpływ środowiska fizyczne lub wirtualne [26].

Już nie ma stwierdzenia, że jest to oprogramowanie wykorzystujące mechanizmy uczenia maszynowego, reprezentacji wiedzy czy też narzędzi statystycznych [27]. Nie ma nawet stwierdzenia, że jest to system teleinformatyczny czy korzystający z komputera. Jest to system maszynowy, a maszyną może być każdy autonomiczny układ, w jakim zachodzi przekształcanie zasilania lub informacji. A na jego wyjściu mogą być sterowania lub wartości wyników wpływające na środowisko fizyczne lub wirtualne. Może to być nawet regulator temperatury klimatyzacji w budynku. Tym samym pojawia się ogromne pole do interpretacji, gdy możemy stwierdzić istnienie systemu SI w danej maszynie [27].

Tutaj trzeba dodać uwagę, że podobnie nie jest to jeszcze ostateczna wersja tej definicji systemu SI, bo będzie ona teraz dyskutowana i modyfikowana łącznie również z modyfikacją definicji systemu SI przyjmowanej przez OECD (stamtąd zresztą pochodziła poprzednia wersja tej definicji).

Warto ją porównać z przytoczoną wcześniej definicją US National Artificial Intelligence Act of 2020 [22]. Czytając ją, widzimy, że była źródłem tej nowej przyjętej w UE, ale bez istotnej frazy wskazującej na decyzyjność człowieka w definiowaniu celów mających być spełnionymi przez system SI!

Druga z poprawek dokonuje zamiany ponad 100 wystąpień terminu „user” (użytkownik) terminem „deployer” (wdrożeńowiec) z następującym wyjaśnieniem:

„wdrożeńowiec” (poprzednio „użytkownik”) oznacza każdą osobę fizyczną lub prawną, organ publiczny, agencję lub inny podmiot korzystający z systemu AI pod swoim nadzorem, z wyjątkiem sytuacji, gdy system AI jest wykorzystywany w ramach osobistej działalności niezawodowej [26].”

Ta zmiana ma znaczący głębszy sens. W tym przypadku system SI pozostaje pod kontrolą wdrożeńowca, który wtedy odpowiada za prawidłowość i skutki jego działania, zdejmując tę odpowiedzialność z bezpośrednich użytkowników. Ta odpowiedzialność wdrożeńowca przewija się przez cały tekst projektu rozporządzenia, nakładając na niego różne obowiązki oraz ewentualnie kary. Nie wiadomo przy tym kto konkretnie, z osoby czy podmiotu, jest rzeczywiście tym wdrożeńowcem i jakie ma środki na swoje działania. Instytucjonalnie będzie mógł wywierać większy wpływ na dostawcę, aby ten dopilnował tych wszystkich narzuconych zasad. A dostawca będzie starać się je przerzucić na producenta systemu SI, a ten na projektantów, programistów, analityków danych uczących – w skrócie na informatyków.

W rezultacie dążeniem autorów tego rozporządzenia jest sformułowanie i narzucenie twórcom takich systemów SI wielu istotnych obowiązków. Ale tych twórców przy tworzeniu tego rozporządzenia raczej nie było – i w tym jest problem. Teraz otwiera się okres na dyskusję nad zapisami tegoż rozporządzenia w poszczególnych krajach i dobrze by było, aby uczestniczyło w nim środowisko oraz firmy informatyczne tym się zajmujące, starając się wyjaśnić wszystkie wątpliwości i zadbać o swoje interesy. Dobrze by też było, aby przyszli potencjalni użytkownicy tych systemów SI również przyglądali się i ewentualnie uczestniczyli w tych dyskusjach. To będą bowiem prawne podstawy korzystania z systemów SI, ustalone na kilka następnych lat.

### **Miara sztucznej inteligencji**

Poznając kolejne definicje systemu SI, możemy odnieść wrażenie, że jeszcze nie ma spójnego poglądu, czym jest w rzeczywistości taki system. Co więcej, nie bardzo wiadomo, jak go odróżnić od zwykłego systemu teleinformatycznego nieprzejawiającego cech inteligentnych.

Może więc zacznijmy od przypomnienia, czym jest angielski termin *intelligence* oraz polski „inteligencja”. Według słowników mamy następujące definicje:

[Słownik OXFORD]

Intelligence – the ability to learn, understand and think in a logical way about things; the ability to do this well (Inteligencja: zdolność uczenia się, rozumienia i logicznego myślenia o rzeczach; umiejętność zrobienia tego dobrze)



[Słownik PWN]

Inteligencja – zdolność rozumienia, uczenia się oraz wykorzystywania posiadanej wiedzy i umiejętności w sytuacjach nowych.

Definicje są podobne, ale nie takie same. Inteligencja jest zdolnością, umiejętnością rozwiązywania nowych problemów na podstawie posiadanej wiedzy. Inaczej mówiąc, ucząc się, mamy więcej tej zdolności. Tylko jak możemy zmierzyć ilość posiadanej inteligencji w danym momencie. Istnieje miara, tak zwany iloraz inteligencji IQ wyznaczany na podstawie dość złożonych testów myślenia logicznego oraz obliczeń. Osoby z  $IQ > 148$  są uważane za bardzo inteligentne. Ale też nie zawsze taka inteligencja przekłada się na jakieś specjalne osiągnięcia.

Poważniejszym problemem jest brak miary sztucznej inteligencji, powiedzmy ilorazu sztucznej inteligencji (AIQ/SIQ). Byłoby to szczególnie użyteczne do określania, na ile dany system teleinformatyczny albo maszyna jest systemem SI, bo nie każda musi nim być. Obecnie to firma, wprowadzając swój system na rynek, twierdzi, że jest on systemem AI, traktując to jako efektywny element marketingowy [tzw. AI Inside]. W zasadzie przy każdym odwołaniu się wprowadzanego w życie rozporządzenia [26] prawnicy będą musieli powoływać biegłych, których zadaniem będzie stwierdzenie, czy dana maszyna jest rzeczywiście systemem SI i o jakiej wartości – tylko w jakiej skali?

Dla oceny mocy komputerów mamy zdefiniowane tak zwane benchmarki (testy sprawności obliczeniowej). Dla systemu SI istnieją już testy określające ich sprawność „intelektualną”. Często są to arkusze testowe dla osób ubiegających się o jakieś stanowisko, certyfikat czy uznanie specjalizacji (np. medycznej). Ale na dłuższą metę to zagadnienie także trzeba będzie uporządkować.

### Cyber bezpieczeństwo wobec SI

Bezpieczeństwo systemów teleinformatycznych (przezwanego jako „cyberbezpieczeństwo”) jest obecnie najpoważniejszym problemem eksploatacji tych systemów dla ich administratorów oraz użytkowników. Liczba różnego typu cyberataków jest tak znacząca, że praktycznie każdy z użytkowników powinien czuć się zagrożony [27].

Dla obrony system SI, ze względu na szybkość działania, ma ogromną przewagę nad ludzkimi obserwatorami funkcjonowania sieci. System SI szybciej może wykryć oznaki lub początki takich ataków, co umożliwia zapobieganie zbyt daleko idącemu skutkowi. Systemy SI szybko też mogą się uczyć, jakie modyfikacje w swoich działaniach wprowadzają atakujący.

Niestety również atakujący, korzystając z systemów SI, mogą szybciej analizować struktury systemów teleinformatycznych, wyszukując w nich punkty podatne na nieautoryzowane wniknięcie w ich strukturę, a tym samym dokonanie zamierzonych działań –

pozyskanie zbiorów danych, zaszyfrowanie plików czy zablokowanie działania systemu. Zmierzamy do konstatacji, że już lub niebawem na tym polu będą zmagaly się ze sobą systemy SI, a wygrana będzie zależała od lepszej sprawności działania jednego z nich. Może to też prowadzić do niekończącej się eskalacji działań, które będzie można przerwać tylko przez wyłączenie dopływu prądu.

Zasadnym jest tutaj pytanie, kto fizycznie jest sprawcą inicjacji tych cyberataków. Często początkowo nie jesteśmy w stanie nawet określić miejsca źródła generującego atak, a tym samym sprawcy, przyczyny i rzeczywistego celu ataku. Ogólnie mogą to być cyberprzestępcy ukierunkowani na pozyskanie wartościowych danych do późniejszej sprzedaży, pobranie środków z kont w bankach czy też doprowadzenie do możliwości szantażu – odszyfrowanie danych za okup (*ransomware*). I tutaj mogą pomóc specjalizowane systemy SI, które zbierając drobne z pozoru informacje, mogą określić, kto fizycznie stoi za danym atakiem. Potem tylko pozostaje wysłać oddział policji, jeżeli mamy jurysdykcję nad tym adresem.

Poważniejszym działaniem jest operacja cyberwojskowa jednego państwa przeciwko innemu. W tym przypadku celem jest zniszczenie lub tylko zakłócenie infrastruktury krytycznej atakowanego, lub też pozyskanie informacji wojskowych czy też wpłynięcie na decyzje polityczne lub nastroje społeczne [28]. W tym przypadku systemy SI mogą również pomóc zidentyfikować źródła ataku oraz określić prawdopodobny cel. Sprawa odwetu jest równie trudna, gdyż w zasadzie powinien to być podobny cyberkontratak lub atak siłami konwencjonalnymi. Obecnie każdy kraj szybko rozbudowuje komponenty wojsk obrony cyberprzestrzeni, gdyż cyberwojna już trwa [29].

I wreszcie same systemy SI mogą być poddane cyberatakowi polegającym na zdezaktualizowaniu modelu językowego poprzez usunięcie z niego wielu danych albo zastąpienie ich danymi fałszywymi prowadzącymi do halucynacji w działaniu systemu SI. Skutki takich działań, w porę nie wykryte, mogą być znaczące – np. kto inny może wygrać wybory.

### **Ku czemu to zmierza?**

Ślepy zaułek związany z Chat GPT, któremu poświęcono na początku tego artykułu sporo uwagi, bo jest on bardzo widoczny – nie zmieni faktu, że ta prawdziwa, poważnie rozumiana sztuczna inteligencja jest rozwijana i stosowana. Międzynarodowe firmy konsultingowe starają się ująć ten rozwój liczbowo, analizując działania firm wdrażających lub zamierzających wdrożyć aplikacje systemów SI. Podawane dane wartości inwestycji są znaczące, stawiając sektor SI w czołówce rozwoju gospodarczego, głównie w Chinach, USA oraz w Unii Europejskiej. Tak – większość tych inwestycji będzie w Chinach, które od 2021 roku są światowym liderem w rozwoju AI. Jest to wynikiem systemu zachęt stosowanych przez chiński rząd w stosunku do prywatnych firm pra-

cujących nad rozwojem sztucznej inteligencji. Ale „po piętach deptają im” ośrodki naukowe USA, a w Europie przodują Niemcy i Wielka Brytania. Duże aspiracje mają Indie i Kanada.

Firmy oferujące systemy SI notują silny przyrost przychodów oraz wzrost wartości swoich akcji na giełdach. Nie podajemy tutaj liczb, bo te łatwo wyszukać, choćby pytając GPT. Trzeba też do nich podchodzić z pewnym krytycyzmem, gdyż wobec tej rewolucji SI analizowanie przepływów finansowych – kto zyska, kto straci – pomiędzy firmami oferującymi systemy SI a firmami je użytkującymi jest jeszcze trudnym zadaniem.

Istotnym pytaniem jest, jaka będzie przyszłość rozwoju SI? Konsultanci przewidują silny wzrost wartości rynku SI, prawie 40% rocznie. Ale my jesteśmy sceptyczni, gdyż taka prosta aproksymacja przyszłych dochodów czy inwestycji nie zawsze się spełnia, gdyż zależy od wielu czynników, o których teraz może nawet nie mamy pojęcia. Zbyt wiele bowiem na tej drodze rozwoju jawi się przeszkód – od sytuacji polityczno-gospodarczej po możliwości wytwórcze sektora informatycznego. Wystarczy wspomnieć, że główne wytwórnie czipów NVIDIA – czołowego dostawcy procesorów graficznych wykorzystywanych w systemach SI – znajdują się na Tajwanie. A bez nich systemy SI nie będą działać.

Innym istotnym czynnikiem wzrostu lub ograniczania rozwoju systemów SI są przygotowywane regulacje zmuszające dostawców systemów SI do samoograniczania się przez konieczność ponoszenia znacząco wyższych kosztów ich wdrażania.

### Podsumowanie

Dyrektor generalny Google Sundar Pichai twierdzi, że sztuczna inteligencja wpłynie bardziej na rozwój ludzkości niż wykorzystanie ognia czy wynalazek elektryczności. To jest ta dobra strona medalu.

Ale równocześnie czytamy, iż Władimir Putin powiedział, że dla Rosji sztuczna inteligencja jest tak samo ważna jak atomowe rakiety w ZSRR [30]. I to jest ta gorsza strona medalu...

Obserwujmy więc bacznie to, co się dzieje, bo przyszłe pokolenia będą się na lekcji historii uczyły o dzisiejszych czasach jako o czasach doniosłego przełomu technicznego i społecznego!

### Bibliografia

- [1] Crawford K.: *The atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, 2021.
- [2] Tadeusiewicz R.: *Introduction to Intelligent Systems*, chapter No 1 in book: Wilamowski B.M., Irvin J.D. (red.): *The Industrial Electronics Handbook – Intelligent Systems*, CRC Press, Boca Raton, 2011, pp. 1–1 ÷ 1–12 (ISBN: 9781439802830 DOI: 10.1201/b10604-3)
- [3] Grabowski J.: 2023 – rok sztucznej inteligencji. Domena – pismo PTI, 2023, [https://kopia.pti.org.pl/wp-content/uploads/2023/03/2\\_2023-rok-sztucznej-inteligencji.pdf](https://kopia.pti.org.pl/wp-content/uploads/2023/03/2_2023-rok-sztucznej-inteligencji.pdf)

- [4] Tadeusiewicz R.: *Archipelag Sztucznej Inteligencji*, Akademicka Oficyna Wydawnicza EXIT, 2021
- [5] Tadeusiewicz R.: *Sieci neuronowe*. Akademicka Oficyna Wydawnicza, Warszawa, 1993
- [6] Tadeusiewicz R.: *Rozpoznawanie obrazów*, Skrypt UJ nr 499, Kraków 1985 (ISBN 83-233-0054-2)  
[https://www.academia.edu/42059662/Rozpoznawanie\\_obrazow\\_-\\_zarys\\_teorii](https://www.academia.edu/42059662/Rozpoznawanie_obrazow_-_zarys_teorii)
- [7] Tadeusiewicz R.: *Rozpoznawanie obrazów w zastosowaniach ekonomicznych*, Skrypt AE, Kraków, 1985,  
[https://www.academia.edu/42124750/Rozpoznawanie\\_obrazow\\_w\\_zastosowaniach\\_ekonomicznych](https://www.academia.edu/42124750/Rozpoznawanie_obrazow_w_zastosowaniach_ekonomicznych)
- [8] Tadeusiewicz R.: *Jak nauczyliśmy maszyny liczyć i myśleć za nas? Część 45: Test Turinga*. Rzecz o Historii, Rzeczpospolita, 21.04.2023, str. J8  
<https://www.rp.pl/historia-swiata/art38360571-jak-nauczylismy-maszyny-liczyc-i-myslec-za-nas-czesc-45-test-turinga>,  
[https://www.academia.edu/104851697/Jak\\_nauczylismy\\_maszyny\\_liczyc\\_i\\_myslec\\_za\\_nas\\_Czesc\\_45\\_Test\\_Turinga](https://www.academia.edu/104851697/Jak_nauczylismy_maszyny_liczyc_i_myslec_za_nas_Czesc_45_Test_Turinga)
- [9] Tadeusiewicz R.: *Jak nauczyliśmy maszyny liczyć i myśleć za nas? Część 46: Jak 20 uczonych w 8 tygodni zrewolucjonizowało relacje ludzi i maszyn*, Rzecz o Historii, Rzeczpospolita, 12.05.2023, str. J8 PDF <https://www.rp.pl/historia/art38463891-jak-nauczylismy-maszyny-liczyc-i-myslec-za-nas-czesc-46-jak-20-uczonych-w-osiem-tygodni-zrewolucjonizowalo-relacje-ludzi-i-maszyn>, [https://www.academia.edu/104852097/Jak\\_nauczylismy\\_maszyny\\_liczyc\\_i\\_myslec\\_za\\_nas\\_Czesc\\_46\\_Jak\\_20\\_uczonych\\_w\\_8\\_tygodni\\_zrewolucjonizowalo\\_relacje\\_ludzi\\_i\\_maszyn](https://www.academia.edu/104852097/Jak_nauczylismy_maszyny_liczyc_i_myslec_za_nas_Czesc_46_Jak_20_uczonych_w_8_tygodni_zrewolucjonizowalo_relacje_ludzi_i_maszyn)
- [10] Holyński M.: *AI a sprawa polska*, Domena nr 3/2023,  
[https://portal.pti.org.pl/wp-content/uploads/2023/10/4\\_AI-a-sprawa-polska.pdf](https://portal.pti.org.pl/wp-content/uploads/2023/10/4_AI-a-sprawa-polska.pdf)
- [11] Tadeusiewicz R.: *Bioniczna koncepcja analizy mowy*. Materiały Konferencji: Metody Bezpośredniego Wprowadzania i Wyprowadzania Informacji Tekstowej i Obrazowej do Maszyny Cyfrowej. Jabłonna 1973, str. VIII/1–VIII/15.  
[https://www.academia.edu/28240917/Bioniczna\\_koncepcja\\_analazy\\_mowy](https://www.academia.edu/28240917/Bioniczna_koncepcja_analazy_mowy)
- [12] Tadeusiewicz R., Mysona-Byrska E., Korpala E.: *Ustalenie rodzaju potrzebnej terapii w przypadku raka sutka metodą rozpoznawania obrazów z zastosowaniem maszyny cyfrowej*. Materiały konferencji „Elektronika i automatyka w służbie medycyny”, Katowice 1974, str. 84–86,  
[https://www.academia.edu/37300005/Ustalenie\\_rodzaju\\_potrzebnej\\_terapii\\_w\\_przypadku\\_raka\\_sutka\\_metoda\\_rozpoznawania\\_obrazow\\_z\\_zastosowaniem\\_maszyny\\_cyfrowej](https://www.academia.edu/37300005/Ustalenie_rodzaju_potrzebnej_terapii_w_przypadku_raka_sutka_metoda_rozpoznawania_obrazow_z_zastosowaniem_maszyny_cyfrowej)
- [13] Tadeusiewicz R.: *Komputer w roli sędziego?* Gazeta Prawnicza, 23 (307), 16.12.1976, str. 12  
[https://www.academia.edu/34807362/Komputer\\_w\\_rolu\\_siedziego](https://www.academia.edu/34807362/Komputer_w_rolu_siedziego)
- [14] Mazon S., Tadeusiewicz R.: *Konwersacja człowieka z maszyną cyfrową w języku naturalnym*. Informatyka, nr 9, 1978, str. 16–19.  
[https://www.academia.edu/28247444/Konwersacja\\_czlowieka\\_z\\_maszyna\\_cyfrowa\\_w\\_jezyku\\_naturalnym](https://www.academia.edu/28247444/Konwersacja_czlowieka_z_maszyna_cyfrowa_w_jezyku_naturalnym)
- [15] Mazon S., Tadeusiewicz R.: *Próba opisu pewnej klasy systemów naturalnej konwersacji człowieka z maszyną cyfrową*. Archiwum Automatyki i Telemechaniki, Tom XXIV, nr 2, 1979, str. 293–295.  
[https://www.academia.edu/36357748/Prоба\\_opisu\\_pewnej\\_klasy\\_systemow\\_naturalnej\\_konwersacji\\_czlowieka\\_z\\_maszyna\\_cyfrowa](https://www.academia.edu/36357748/Prоба_opisu_pewnej_klasy_systemow_naturalnej_konwersacji_czlowieka_z_maszyna_cyfrowa);

- [https://www.researchgate.net/publication/265546583\\_An\\_attempt\\_at\\_describing\\_a\\_class\\_of\\_man-machine\\_dialogue\\_systems\\_based\\_on\\_a\\_natural\\_language](https://www.researchgate.net/publication/265546583_An_attempt_at_describing_a_class_of_man-machine_dialogue_systems_based_on_a_natural_language)
- [16] Stefanowicz B.: *22 Best AI Chatbots for 2023: ChatGPT & Alternatives*, 6.09.2023, Tidio, <https://www.tidio.com/blog/ai-chatbot/>
- [17] Iszkowski W.: *Wywiad ze sztuczną inteligencją OpenAI*. CRN, 2023-03-03. <https://www.crn.pl/wywiady-i-felietony/wywiad-ze-sztuczna-inteligencja-openai/>
- [18] *Pause Giant AI Experiments* – open letter, <https://www.futureoflife.org/open-letter/pause-giant-ai-experiments/>
- [19] Iszkowski W.: *Systemy SI: nowa wspaniała era czy koniec znanego świata?* CRN, 2023-06-05, <https://www.crn.pl/wywiady-i-felietony/systemy-si-nowa-wspaniala-era-czy-koniec-znanego-swiata/>
- [20] Tadeusiewicz R.: *Legendarna prehistoria robotów*, Gazeta Myślenicka nr 24, 30.06.2021, str. 6, <https://gazeta.myslenice.pl/legendarna-prehistoria-robotow;> [https://www.academia.edu/73927591/Legendarna\\_prehistoria\\_robot%C3%B3w](https://www.academia.edu/73927591/Legendarna_prehistoria_robot%C3%B3w)
- [21] Stern J.: *I Cloned Myself With AI. She Fooled My Bank and My Family*. *Wall Street Journal*, 28.04.2023. <https://www.wsj.com/articles/i-cloned-myself-with-ai-she-fooled-my-bank-and-my-family-356bd1a3>
- [22] *National AI Initiative Act of 2020* <https://www.ai.gov/#:~:text=The%20National%20AI%20Initiative%20Act,economic%20prosperity%20and%20national%20security>
- [23] The White House. *FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI*. 21.07.2023. <https://www.whitehouse.gov/briefing-room/statements-releases/2023/07/21/fact-sheet-biden-harris-administration-secures-voluntary-commitments-from-leading-artificial-intelligence-companies-to-manage-the-risks-posed-by-ai/>
- [24] Komisja Europejska: *Wniosek rozporządzenie Parlamentu Europejskiego i Rady ustanawiające zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji) zmieniające niektóre akty ustawodawcze unii*, 21.04.2021. COM(2021) 206 final. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>
- [25] Chustecki J.: *Polski UODO rozpatruje skargę dotyczącą działania bota ChatGPT*. 22.09. 023. ComputerWorld. <https://www.computerworld.pl/news/Polski-UODO-rozpatruje-skarge-dotyczaca-dzialania-bota-ChatGPT,446447.html>
- [26] European Parliament 2019–2024. *Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206–C9-0146/2021–2021/0106 (COD))*. 14.06.2023. [https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236\\_EN.pdf](https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf)
- [27] Iszkowski W.: *Nowa (unijna) definicja systemu SI*. CRN, 2023-09-11, <https://www.crn.pl/wywiady-i-felietony/nowa-unijna-definicja-systemu-si/>
- [28] Silicki K.: *Cyberbezpieczeństwo sztucznej inteligencji – na co musimy być przygotowani*. Cyberbezpieczeństwo Redefinicja zagrożeń, Wojskowa Akademia Techniczna, 2023
- [29] Chmielewski M.: *Strategiczny potencjał wykorzystania sztucznej inteligencji w budowie cyberodporności państwa*. Cyberbezpieczeństwo Redefinicja zagrożeń, Wojskowa Akademia Techniczna, 2023

- [30] Jędrzejczak Ł.: *Czy to już wojna? – cyberbezpieczeństwo oczami Szefa CSIRT MON*. Cyberbezpieczeństwo Redefinicja zagrożeń, Wojskowa Akademia Techniczna, 2023
- [31] Максим Куклин М.: *Путин раскрыл, что для России сейчас так же важно, как атомные ракеты в СССР 19.07.2023*. URA.RU, <https://ura.news/news/1052668256>

### Na marginesie dyskusji o sztucznej inteligencji

Artykuł przedstawia szereg uwag dotyczących sztucznej inteligencji (SI), a które nie są oczywiste dla osób, które tą dziedziną na co dzień się nie zajmują. Na początku jest dyskutowana sama nazwa „sztuczna inteligencja” oraz czym ona jest. Jest to potrzebne dla śledzenia dyskusji na temat tego, jaka jest owa SI. Potem opisano, jak SI powstawała – na świecie i w Polsce. Omówione są też konsekwencje pojawienia się programu Chat GPT, a także zasady jego działania. Na fali szerokiego zainteresowania SI pojawiło się też wiele dyskusyjnych wypowiedzi, często pochodzących od autorytetów naukowych z obszarów nauki odległych od informatyki. Polemika z takimi wypowiedziami jest w punkcie zatytułowanym „Chwasty do wyplewiania”. W dalszej części artykułu wykazano, że ludzie pragnęli SI już w starożytności oraz opisano, co potrafi zrobić inteligentny awatar. Zaprezentowano też, jakie prawne regulacje usiłuje się obecnie nałożyć na systemy SI – odpowiednio w Stanach Zjednoczonych oraz w Unii Europejskiej, a więc też w Polsce. Na koniec podjęto temat wzajemnych relacji między SI a cyberbezpieczeństwem.

**Słowa kluczowe:** sztuczna inteligencja, historia SI, systemy SI, GPT, definicja SI, cyberbezpieczeństwo

### Sideline discussion about artificial intelligence

The article presents a number of comments regarding artificial intelligence (AI) that are not obvious to people who do not deal with this field on a daily basis. At the beginning, the name “artificial intelligence” itself and what it is are discussed. This is needed to follow the discussion on what this AI is. Then it was described how AI was created – in the world and in Poland. The consequences of the appearance of the Chat GPT program, as well as the principles of its operation, are also discussed. Due to the widespread interest in AI, controversial statements have also appeared, often coming from scientific authorities from areas of science far from computer science. There is a polemic against such statements in the section entitled “Weeds to weed out”. The article goes on to show that people have wanted AI since ancient times and describe what an intelligent avatar can do. It also presents what legal regulations are currently being tried to impose on AI systems – in the United States and in the European Union, respectively, and therefore also in Poland. Finally, the topic of mutual relations between AI and cybersecurity was discussed.

**Key words:** artificial intelligence, history of AI, AI systems, GPT, AI definition, cybersecurity