Aleksander Piecuch
https://orcid.org/0000-0001-5889-9643

DOI: 10.34866/0584-df51

Sztuczna inteligencja w perspektywie społecznej

Artificial intelligence in a social perspective

Key words: expert system, artificial intelligence, ChatGPT-3, deepfake, voice cloning, terrorism.

Abstract: We welcome and embrace every new scientific and technological development that somehow makes our lives a little easier. Often, we do not reflect at all, are not aware of, or relegate to a distant perspective, the negative consequences of "technical innovations". We all recognize the positives of general-purpose artificial intelligence, but do we also recognise its negative aspects? This article only hints at some selected negative aspects of its use. The purpose of this study is to draw the reader's attention to the dangers potentially posed by the use of artificial intelligence.

Słowa kluczowe: system ekspertowy, sztuczna inteligencja, ChatGPT-3, deepfake, klonowanie głosu, terroryzm.

Streszczenie: Cieszymy się i przyjmujemy każde nowe osiągnięcie naukowo-techniczne, które w jakiś sposób czyni nasze życie nieco łatwiejszym. Często w ogóle nie zastanawiamy się, nie mamy świadomości, albo odsuwamy na dalszą perspektywę negatywne konsekwencje „nowinek technicznych". Pozytywy sztucznej inteligencji ogólnego przeznaczenia dostrzegamy wszyscy, ale czy dostrzegamy również jej negatywne aspekty? W artykule zasygnalizowano tylko niektóre wybrane negatywne aspekty jej wykorzystania. Celem niniejszego opracowania jest zwrócenie uwagi czytelnika na niebezpieczeństwa, jakie potencjalnie stwarza wykorzystanie sztucznej inteligencji.

Wstęp

Przyzwyczailiśmy się do stwierdzeń, że przełom wieków XX i XXI przyniósł na niespotykaną skalę rozwój nauki i techniki, a w tym również technik informatycznych i informacyjnych. Poglądu tego nie można podważać, niemniej jednak nad technologiami, także tymi związanymi bezpośrednio z informatyką, pracowano od lat 40. ubiegłego wieku. W obszarze zainteresowań ówczesnych naukowców znalazła się również sztuczna inteligencja. W roku 1956 John McCarthy (amerykański informatyk) na konferencji w Dortmouth po raz pierwszy do obiegu naukowego wprowadził termin sztuczna inteligencja (ang. *Artificial Intelligence – AI*), który z powodzeniem funkcjonuje po dzień dzisiejszy.

Problemy definicyjne

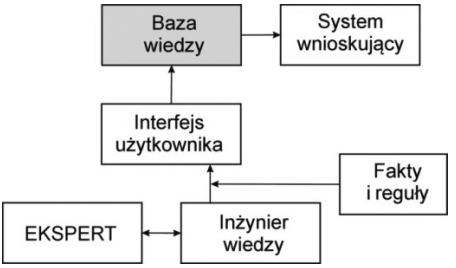
Pomimo upływu czasu, jaki minął od zaistnienia w przestrzeni naukowej terminu sztuczna inteligencja, nadal brak jest jednoznacznej i ogólnie akceptowalnej definicji zarówno w przemyśle, jak i środowisku naukowym. W literaturze przedmiotu odnajdziemy różnie brzmiące definicje. Dla przykładu przytoczmy kilka z nich. „Stuart Russell i Peter Norvig opisują ją tak: »Dziedzina badań nad agentami, które przyjmują bodźce ze środowiska i wykonują działania«, natomiast John McCarthy stosuje następującą definicję: »Dziedzina nauki i inżynierii związana z budowaniem inteligentnych maszyn, a zwłaszcza inteligentnych programów komputerowych (...) Inteligencja jest obliczeniowym aspektem możliwości realizowania celów w świecie«" (za: McIlwraith, Marmanis, Babenko, 2017, s. 27). Według pioniera badań nad sztuczną inteligencją Nilsa J. Nilssona „sztuczna inteligencja to działalność poświęcona uczynieniu maszyn inteligentnymi, a inteligencja to taka jakość, która umożliwia podmiotowi właściwe i przewidywalne funkcjonowanie w swoim otoczeniu" (za: Nowak-Nova, 2021).

Włodzisław Duch sztuczną inteligencję definiuje następująco: „to dziedzina nauki, zajmująca się rozwiązywaniem problemów efektywnie niealgorytmizowalnych, w oparciu o modele wiedzy (Duch, 1997, s. 54). R. Tadeusiewicz do kwestii definicyjnej odnosi się w słowach „ze sztuczną inteligencją mamy do czynienia wtedy, gdy maszyna (komputer albo elektronicznie sterowane urządzenie: robot, autonomiczny pojazd, samoorganizująca się sieć połączeń) przejawia zachowania, które obserwowane u człowieka powodowałyby, że bylibyśmy skłonni je uznać za skutek jego inteligencji (Tadeusiewicz, 2020, s. 27).

Fachowa literatura dostarczyłaby jeszcze większej liczby definicji. Wydaje się, że te przywołane powyżej oddają sens znaczenia sztucznej inteligencji. Na marginesie dodajmy, że samo pojęcie jest oksymoronem. Z naukowego punktu widzenia „Inteligencja to konstrukt teoretyczny odnoszący się do względnie stałych warunków wewnętrznych człowieka, determinujących efektywność działań wymagających udziału typowo ludzkich procesów poznawczych. Warunki te kształtują się w wyniku interakcji genotypu, środowiska i własnej aktywności. (...) Inteligencja, która właściwa jest człowiekowi, uzewnętrznia się w tzw. inteligentnym zachowaniu. Nie jest ona zdeterminowana wyłącznie przez genotyp, podobnie jak nie jest wytworem tylko środowiska. Inteligencja jest wynikiem interakcji między tymi czynnikami. Szczególną rolę w procesie tej interakcji odgrywa własna aktywność jednostki. Ona determinuje w znacznej mierze, z jakim środowiskiem i w jaki sposób człowiek wchodzi w interakcję, a więc w konsekwencji, jakie wpływy środowiska zostają zinterioryzowane („uwewnętrznione") przez człowieka. Na szczególną uwagę zasługuje to, że dzięki własnej aktywności warunki wewnętrzne człowieka, do których odnosi się ów konstrukt teoretyczny zwany inteligencją, zmieniają się" (Strelau, 1987, s. 15). Czy zatem można przypisać programowi komputerowemu inteligencję skoro ten pozbawiony jest samoświadomości, moralności i etyki, a nadto nie rozróżnia prawdy od fałszu.

Systemy ekspertowe

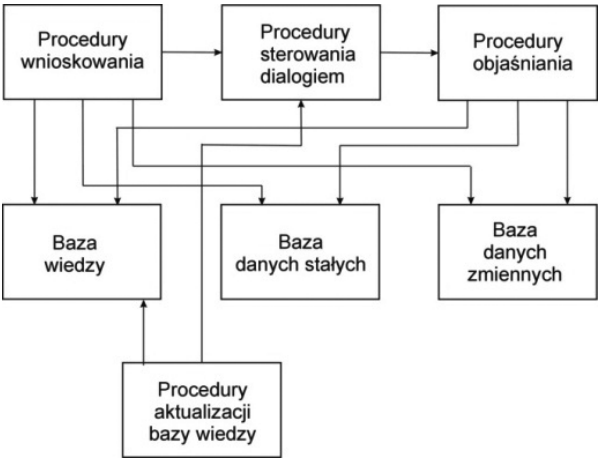
Zanim zajmiemy się tytułową problematyką *General Purpose AI* zasadne jest przywołanie i krótkie omówienie zagadnień związanych z systemem ekspertowym. „System ekspertowy jest programem komputerowym, który wykonuje złożone zadania o dużych wymaganiach intelektualnych i robi to tak dobrze jak człowiek będący ekspertem w tej dziedzinie. Określenie »system ekspertowy« może być zastosowane do dowolnego programu komputerowego, który na podstawie szczegółowej wiedzy może wyciągać wnioski i podejmować decyzje, działając w sposób zbliżony do procesu rozumowania człowieka" (Mulawka, 1996, s. 20). Ogólną architekturę systemu ekspertowego pokazano na rys. 1.



Rys. 1. Podstawowa architektura systemu ekspertowego

Źródło: opracowanie własne na podstawie: Mulawka, 1996, Tadeusiewicz, 2020.

Uszczegóławiając architekturę z rys. 1 otrzymuje się (rys. 2):



Rys. 2. Główne komponenty systemu ekspertowego

Źródło: Mulawka, 1996, s. 23.

Prezentowana na rys. 1 i rys 2 architektura systemu ekspertowego na pierwszy rzut oka dostarcza bardzo istotnej informacji o zamkniętej strukturze systemu. Oznacza

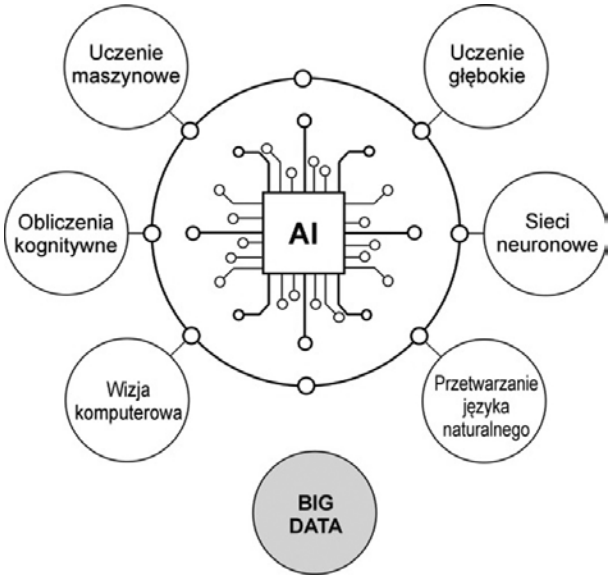
to, że do bazy wiedzy nie mogą i nie trafiają przypadkowe informacje. Wszystkie one pochodzą ze „znanego i wiarygodnego źródła" – od eksperta. Sposób działania systemu wyjaśniono poniżej.

System ekspertowy działa dwuetapowo. Pierwszy etap polega na dostarczeniu do bazy wiedzy specjalistycznych informacji. Te można pozyskać jedynie od wysokiej klasy specjalistów, tzw. ekspertów dziedzinowych. Problem polega jednak na tym, że te nie mogą zostać wprowadzone do bazy wiedzy w sposób dowolny. Specjalistyczne informacje pochodzące od eksperta dziedzinowego należy wprowadzić do bazy poprzez interfejs użytkownika w postaci faktów oraz reguł i tym problemem zajmuje się inżynier wiedzy. W ten sposób gromadzone są specjalistyczne informacje w bazie wiedzy. Oprócz nich do systemu należy dostarczyć również informacje natury ogólnej (baza danych stałych), czyli takie, które dla człowieka są czymś oczywistym, ale dla komputera takimi nie są. Baza danych stałych zabezpiecza w ten sposób system wnioskowania przed takimi rozwiązaniami postawionego problemu, które byłyby nie akceptowalne przez człowieka (np. natury moralnej, etycznej). Po tych czynnościach system gotowy jest do drugiego etapu, którym jest jego eksploatacja.

Użytkownik systemu ekspertowego wprowadza w języku naturalnym zapytanie do systemu. Jeśli problem został niewystarczająco precyzyjnie sformułowany przez użytkownika, następuje dialog systemu z użytkownikiem. W wyniku wnioskowania użytkownik może otrzymać rozwiązanie swojego problemu, przy czym partykuła „może" nie jest tu bez znaczenia. System może uznać, że problem jest nierozwiązywalny i taki komunikat skieruje do użytkownika. Jednocześnie zasugeruje dostarczenie większej liczby danych wejściowych lub zmianę pierwotnie określonych przez użytkownika kryteriów. Jeśli zaistnieje któraś z wymienionych sytuacji lub obie jednocześnie, problem zostanie rozwiązany. Bardzo ważnym z punktu widzenia użytkownika jest system objaśniający. To dzięki niemu użytkownik otrzymuje możliwość prześledzenia toku „rozumowania" (wnioskowania) przez system ekspertowy. Dodajmy jeszcze, że każdy nowy rozwiązany problem wzbogaca bazę wiedzy systemu. Poziom zaawansowania technologicznego współczesnych systemów ekspertowych jest już na tyle wysoki, że może posłużyć do rozwiązywania dowolnej klasy problemów. Stąd też istnieje możliwość nabycia tzw. systemu szkieletowego. Posiada on wszystkie funkcjonalności systemu ekspertowego i pustą bazę wiedzy.

Sztuczna inteligencja

Sztuczna inteligencja pod względem zastosowanych w niej rozwiązań programistycznych jest niezwykle skomplikowana. Jej strukturę można przedstawić jedynie w sposób uproszczony tak jak pokazano na rysunku 3. Wyraźnie trzeba zaznaczyć, że wskazanych na rys. 3 elementów AI nie można rozpatrywać jako wyizolowanych komponentów. Wszystkie jednocześnie tworzą strukturę AI i nawzajem wykorzystują swoje algorytmy.



Rys. 3. Kluczowe komponenty sztucznej inteligencji

Źródło: opracowanie własne na podstawie V. Kanade, 2022.

Funkcjonalność poszczególnych bloków AI definiuje się następująco:

- Uczenie maszynowe – (ang. *Machine Learning* – ML). To podzbiór AI, nie wymagający jawnego programowania. Uczy się automatycznie na podstawie uprzednich doświadczeń. Dokładność, z jaką przebiega uczenie maszynowe, rośnie wraz z czasem i ilością danych.
- Głębokie uczenie – (ang. *Deep Learning* – DL). Uczenie głębokie przetwarza dane za pomocą sztucznych sieci neuronowych. DL jest podzbiorem uczenia maszynowego ML. „Trening z dużą ilością danych konfiguruje neurony w sieci neuronowej. Wynikiem takiego treningu jest model uczenia głębokiego, który po przeszkoleniu przetwarza nowe dane. Modele uczenia głębokiego pobierają informacje z wielu źródeł danych i analizują te dane w czasie rzeczywistym, bez konieczności interwencji ze strony człowieka" (ORACLE). W procesie uczenia głębokiego procesory graficzne (GPU) są zoptymalizowane pod kątem modeli szkoleń, ponieważ mogą przetwarzać wiele obliczeń jednocześnie.
- Sieci neuronowe (ang. *Neural Network* – NN). „System przeznaczony do przetwarzania informacji, którego budowa i zasada działania są w pewnym stopniu wzorowane na funkcjonowaniu fragmentów rzeczywistego (biologicznego) systemu nerwowego. Na przesłankach biologicznych oparte są schematy sztucznych neuronów wchodzących w skład sieci oraz (w pewnym stopniu) jej struktura. Jednak schematy połączeń neuronów w sieci neuronowej są wybierane arbitralnie, a nie stanowią modelu rzeczywistych struktur nerwowych. Wyróżniającą cechą sieci neuronowej jako narzędzia informatycznego jest możliwość komputerowego rozwiązywania przy jej pomocy praktycznych problemów bez ich uprzedniej matematycznej formalizacji. Dalszą zaletą jest brak konieczności

odwoływania się przy stosowaniu sieci do jakichkolwiek teoretycznych założeń na temat rozwiązywanego problemu. Nawet założenie o przyczynowo-skutkowych zależnościach między wejściem a wyjściem nie musi być egzekwowane! Najbardziej znamienną cechą sieci neuronowej jest jej zdolność uczenia się na podstawie przykładów i możliwość automatycznego uogólniania zdobytej wiedzy (generalizacja)" (Tadeusiewicz, Szaleniec, 2015, s. 94).

- Obliczenia kognitywne (ang. *Cognitive Computing* – CC). Obliczenia kognitywne symulują ludzkie procesy myślenia. Istota obliczeń kognitywnych sprowadza się do uczenia się, rozumienia zadań i interpretacji danych. Do tej grupy zaliczane jest rozpoznawanie obrazów i przetwarzanie języka naturalnego. Obliczenia kognitywne wykorzystują przede wszystkim sztuczne sieci neuronowe.
- Przetwarzanie języka naturalnego (ang. *Natural Language Processing* – NLP). Jest narzędziem pozwalającym na komunikację z człowiekiem. Polega na rozpoznawaniu, rozumieniu, interpretowaniu języka naturalnego. Komunikacja może odbywać się w formie tekstowej i generowaniu mowy.
- Wizja komputerowa (ang. *Computer Vision* – CV). Przetwarzanie wizyjne wykorzystuje proces głębokiego uczenia. Pozwala na identyfikację zróżnicowanych wzorców obrazowych np.: tabele, obrazy, wykresy, wideo.

Podsumowując, tutaj ujawnia się różnica pomiędzy systemem ekspertowym a *General Purpose AI* (sztuczną inteligencją ogólnego przeznaczenia). System sztucznej inteligencji jest otwarty, a jego efektywne działanie zależy od dostępu do Big Data (źródeł internetowych). Uprawnione zatem będzie stwierdzenie, że jakość dialogu AI z człowiekiem będzie funkcją jakości informacji, z których ona korzysta. Być może nie jest to problem dnia dzisiejszego ani jutrzejszego, ale nie mamy pewności, że w przyszłości internet celowo (np. farma trolli) nie zostanie zarzucany gigantyczną ilością fałszywych informacji, które następnie wykorzysta AI.

Jak złożona i skomplikowana jest struktura sztucznej inteligencji niech świadczy przykład GPT-3. „GPT-3, czyli Generative Pre-trained Transformer 3 (GPT-3) jest autoregresyjnym modelem językowym, który wykorzystuje głębokie uczenie do tworzenia tekstu łudząco podobnego do napisanego przez człowieka. Jest to model predykcyjny trzeciej generacji w serii GPT-n (i następca równie słynnego GPT-2), stworzony przez jedną z najbardziej pionierskich w obszarze sztucznej inteligencji firm, czyli OpenAI. Pełna wersja GPT-3 ma pojemność 175 miliardów parametrów uczenia maszynowego. GPT-3, który został wprowadzony w maju 2020 r., a od lipca 2020 r. znajdował się w fazie testów beta, wpisuje się w trend systemów przetwarzania i rozumienia języka naturalnego (*Natural Language Processing* – NLP i *Natural Language Understanding* – NLU), polegający na wstępnie wytrenowanych reprezentacjach językowych. Przed wydaniem GPT-3 największym modelem językowym był Turing NLG Microsoftu, wprowadzony w lutym 2020 r., o pojemności 17 miliardów parametrów – mniej niż jedna dziesiąta pojemności GPT-3" (Przegalińska, 2022, s. 17)

Sztuczna inteligencja w perspektywie społecznej

Zasady, którymi kierują się współczesne światowe gospodarki, są w zasadzie dość proste. Dążyć do jak najszybszego rozwoju, minimalizując koszty i maksymalizując zyski, a tym samym uzyskać przewagę technologiczną nad konkurencyjnymi gospodarkami. Taki status osiąga się rozwijając badania naukowo-techniczne. W wyścigu technologicznym kluczową rolę odgrywa czas. Im jest on krótszy, tym większe szanse daje na osiągnięcie zakładanych celów. Sztuczna inteligencja jest tym narzędziem, które ów czas dochodzenia do zakładanych celów może wydatnie skrócić. W tym wymiarze odnosi się do rozwoju gospodarczego, a ściślej mówiąc do przyspieszenia rozwoju kluczowych technologii przemysłowych. Warto dodać, że AI nie pozostaje bez wpływu na politykę i społeczeństwo. Spektrum oddziaływania jest tak duże, że nie sposób odnieść się do wszystkich wątków. Stąd w dalszej części zostaną zasygnalizowane wybrane zagadnienia.

Rosnące zainteresowanie AI wynika bezpośrednio z szerokiego spektrum możliwości jej implementacji „od przemysłu i medycyny do rozwiązań społecznych i edukacji" (Rzeźnik, 2023, s. 5). W ślad za wspomnianymi możliwościami podążają strategie innowacyjności. „Założeniem aktualnej polityki rozwoju jest odchodzenie od dotychczasowego wspierania wszystkich sektorów/branż i skupienie się na wspieraniu sektorów strategicznych, mogących stać się motorami polskiej gospodarki" (https://smart.gov.pl). Zgodnie z tym założeniem wyróżniono 13 obszarów specjalizacji KIS[1] bezpośrednio lub pośrednio związanych z rozwojem sztucznej inteligencji. Należą do nich (https://smart.gov.pl/pl/):

- KIS-1: Zdrowe społeczeństwo.
- KIS-2: Nowoczesne rolnictwo, leśnictwo i żywność.
- KIS-3: Zrównoważone(Bio)produkty. (Bio)procesy i środowisko.
- KIS-4: Zrównoważona energia.
- KIS-5: Inteligentne budownictwo zeroemisyjne.
- KIS-6: Transport przyjazny środowisku.
- KIS-7: Gospodarka o obiegu zamkniętym.
- KIS-8: Zaawansowane materiały i nanotechnologie.
- KIS-9: Elektronika i fotonika
- KIS-10: Technologie informacyjne, komunikacyjne oraz geoinformacyjne.
- KIS-11: Automatyzacja i robotyka.
- KIS-12: Przemysły kreatywne.
- KIS-13: Technologie morskie.

Ze względu na ograniczoną objętość opracowania pozostajemy tylko przy wymienieniu obszarów KIS. Zainteresowani szczegółowy opis Krajowych Inteligentnych Specjalizacji znajdą w Raporcie PARP-2023 (red. G. Rzeźnik) lub na stronie internetowej https://smart.gov.pl/pl/. Działania podejmowane w ramach KIS – w praktyce mają również realizować unijną strategię wzrostu *Europa 2020,* której cele odnoszą

[1] KIS – Krajowa Inteligentna Specjalizacja.

się do: zatrudnienia, innowacji, edukacji, włączenia społecznego oraz zmian klimatu i polityki energetycznej. Wskazane obszary inteligentnych specjalizacji pozostają poza polemiką. Jeśli celem jest rozwój w szerokim tego słowa znaczeniu, to jedynie słuszną drogą do jego osiągnięcia jest wykorzystywanie najnowszych zdobyczy naukowo-technicznych. Te wprawdzie rozwiązują aktualne potrzeby gospodarki (społeczeństwa, człowieka), ale też niosą z sobą nieprzewidywalne ryzyko. Nie mylił się więc Jacques Ellul wskazując, iż:

1. „Wszelki postęp techniczny powoduje zarówno zyski, jak i straty; gdy coś dodaje, to zawsze coś ujmuje.
2. Wszelki postęp techniczny stwarza więcej problemów, aniżeli rozwiązuje; skłania nas do postrzegania tych problemów jako technicznych ze swej natury i popycha do szukania rozwiązań technicznych.
3. Negatywne aspekty technicznych innowacji są nierozłącznie związane z aspektami pozytywnymi. Naiwnością jest sąd, że technika jest neutralna, iż może być używana dla dobrych albo dla złych celów; w rzeczywistości dobre i złe konsekwencje są równoczesne i nieodłączne.
4. Wszystkie wynalazki techniczne mają nieprzewidywalne konsekwencje" (za: Goban-Klas, 2007, s. 49–50).

Poglądy Ellula znajdywały jak do tej pory potwierdzenie. Znajdują także w odniesieniu do sztucznej inteligencji. Oprócz oczywistych udogodnień, z których korzystamy codziennie, np. wyszukiwarki internetowe, poczta, chatboty itp., ujawniają się również negatywne skutki jej wdrażania i eksploatacji. Przypomnijmy tylko ważniejsze fakty, które na naszych oczach tworzą historię sztucznej inteligencji. W 2016 r. samochód Tesli jadący w trybie autopilota po autostradzie na Florydzie „nie zauważył" kilkunastotonowej ciężarówki, kierowca zginął na miejscu. W marcu 2018 r. pojazd autonomiczny testowany przez Ubera na jednej z dróg amerykańskiego stanu Arizona wziął udział w wypadku, w którego wyniku zginęła kobieta. O wiele bardziej tragiczne w skutkach okazały się dwie katastrofy lotnicze Boeninga 737 Max, w których łącznie zginęło 346 osób, a wśród nich dwóch obywateli Polski. Mowa tu o katastrofie indonezyjskich linii lotniczych Lion Air z 29 października 2018 r. i etiopskich linii lotniczych (Ethiopian Airlines) z 10 marca 2019 r. (zob: Szulczewski, 2019, s. 21). Zdaniem ekspertów od wypadków lotniczych odpowiedzialny za katastrofy okazał się być innowacyjny system MCAS (ang. *Maneuvering Characteristics Augmentation System* – system poprawy charakterystyki manewrowej), mający zapobiegać tzw. przeciągnięciu. „Jeśli system zdecyduje o skierowaniu nosa 737 MAX w dół, to pilot nie może zatrzymać tego opadania poprzez zwyczajne pociągnięcie urządzenia sterowego do siebie!" (zob.: Gałabuda, 2019). Zawiodła sztuczna inteligencja, a za jej błędne decyzje konsekwencje poniosły konkretne osoby znane z imienia i nazwiska – ofiary katastrof.

Pomimo stosunkowo krótkiego czasu funkcjonowania ChatGPT-3 odnotowano już próby jego wykorzystania do tworzenia złośliwych narzędzi. Analitycy Check Point Research potwierdzili w 2023 r. fakt przydatności sztucznej inteligencji do przepro-

wadzania ataków phishingowych. AI pozwala na automatyczne generowanie bardzo naturalnie brzmiących wiadomości, które trudno odróżnić od tych napisanych przez człowieka. Można przewidywać, że sukcesywnie będzie wzrastało zagrożenie dla przedsiębiorstw i indywidualnych użytkowników. Jak dodają „W ostatnich tygodniach widzimy dowody na to, że hakerzy zaczęli używać ChatGPA do pisania złośliwego kodu. ChatGPT może przyspieszyć proces hakerów, dając im dobry punkt wyjścia" (zob.: https://managerplus.pl/).

W kolejnym przykładzie pokazujemy przewagę AI nad człowiekiem. Prawdopodobnie wszyscy znają nagrodzoną na prestiżowym konkursie fotograficznym Sony World Photography Awards pracę berlińskiego artysty Borisa Eldagssena wygenerowaną przez sztuczną inteligencję. Autor nagrodzonej pracy nagrody nie odebrał i ujawnił źródło jej pochodzenia. Można postawić pytanie retoryczne: skoro oceniający pracę członkowie jury, z bogatym doświadczeniem w tej branży, ulegli manipulacji, to jakie szanse w konfrontacji z AI ma statystyczny Kowalski i Nowak?

Kolejnym przykładem manipulacji z wykorzystaniem AI jest *deepfake*. „Samo słowo *deepfake* pochodzi od dwóch angielskich zwrotów: *deep learning* (głębokie uczenie) oraz *fake* (fałsz, podróbka). I już to dobrze tłumaczy, czym jest deepfake – **obróbką dźwięku i obrazu, która ma na celu utworzenie fałszywych obrazów i dźwięków przy użyciu technik z zakresu sztucznej inteligencji.** W założeniu pozwala to na stworzenie materiałów, które będą trudne lub niemożliwe do odróżnienia od filmów lub zdjęć, które zostały zrealizowane w tradycyjny sposób – z udziałem żywych osób" (Kulas, 2019). Sposobów na niezgodne z prawem wykorzystanie wizerunków i mowy osób jest wiele. Wskazać możemy na: fake news – tworzenie fałszywych materiałów z udziałem osób publicznych, pornografia – podmiana wizerunku osoby występującej w filmach[2], logowanie i autoryzacja – oszukiwanie systemów autoryzujących przy pomocy wizerunku i mowy, ataki finansowe – podszywanie się pod osoby decyzyjne podczas rozmów wideo. Istnieją już udokumentowane przypadki z wykorzystaniem klonowania głosu (ang. *voice cloning)*. „W 2019 roku, który zdaniem wielu ekspertów jest pierwszym udokumentowanym przypadkiem takiego ataku, oszuści wykorzystali konwersję głosu, by podszyć się pod prezesa i poprosić o pilny przelew środków (na ich konta). Rok później, w 2020 roku, kolejna grupa oszustów wykorzystała tę technologię do naśladowania głosu klienta, aby przekonać menedżera banku do przekazania 35 milionów dolarów na pokrycie »przejęcia«" (https://prnews.pl/, 2022). Na całym świecie odnotowuje się coraz więcej przypadków oszustw z wykorzystaniem technik klonowania głosu. Celem

[2] Za przykład może posłużyć wykorzystany w filmie pornograficznym wizerunek prezenterki i youtuberki, znanej jako Sunpi. Z badań firmy Deeptrace, zajmującej się cyberbezpieczeństwem, wynika, że aż 96 procent wszystkich dostępnych w sieci deepfake'owych wideo to filmy pornograficzne. Zgromadzone dotychczas dane wskazują, że w zdecydowanej większości do ich stworzenia wykorzystywane są wizerunki kobiet – mężczyźni stanowią około czterech procent wszystkich poszkodowanych. [źródło: https://tvn24.pl/ciekawostki/deepfake-sunpi-odkryla-w-internecie-deepfakeowe-tresci-pornograficzne-ze-swoja-twarza-musiala-zaplacic-za-ich-usuniecie-6895292] (dostęp: 3.09.2023).

ataków stały się już nie tylko różnorakie firmy, ale zwykli obywatele. Szczególnie bolesne są te przypadki, kiedy ofiarę ataku przekonuje się np. o porwaniu dziecka przedstawiając w rozmowie telefonicznej jako dowód wygenerowany przez AI głos uprowadzonego dziecka. Dodajmy, że do spreparowania fałszywego dowodu wystarczająca jest 3-sekundowa próbka głosu (zob.: Tałach, 2023). Jak twierdzą autorzy badania przeprowadzonego na University College London w 2020 roku, „fałszywe treści audio lub wideo zostały uznane przez ekspertów za najbardziej niepokojące wykorzystanie sztucznej inteligencji pod względem jej potencjalnych zastosowań w przestępczości lub terroryzmie (...). Deep fake'i mogą być również orężem w operacjach informacyjnych i wojnie informacyjnej" (za: Olech, Lis, 2021, s. 97–98).

Rozmiar oddziaływania AI na społeczeństwo jest tak duży, że nie sposób nie odnieść się w tym miejscu do aspektów bezpieczeństwa państwa. Korzystając z różnych narzędzi ICT na ogół dostrzegamy tylko jej pozytywne aspekty, nie zastanawiając się nad szerszym kontekstem, wykraczającym poza nasze potrzeby. Piotr Sienkiewicz przytacza wypowiedź Alvina Tofflera po wojnie w Zatoce Perskiej: „wojnę w Zatoce Perskiej wygrała inteligencja ukryta w mikroprocesorach systemów uzbrojenia oraz systemach dowodzenia, łączności i rozpoznania" (Sienkiewicz, 1999, s. 60). Od tego czasu upłynęło ponad 30. lat, a współczesna elektronika, technika informatyczna i wojskowa są już na zupełnie innym i nieporównywalnym poziomie. Sztuczna inteligencja znalazła się w „wykazie technologii priorytetowych dla NATO, kluczowych dla rozwoju zdolności militarnych, które powinny zostać rozwinięte w państwach członkowskich w perspektywie średnio- i długookresowej. Rozwój technologii autonomicznych ma obejmować SI, jej systemy wykorzystywane do wykonywania misji oraz szukanie rozwiązań umożliwiających optymalne współdziałanie człowieka i systemów autonomicznych. W 2017 r. SI została również uznana przez Europejską Agencję Obrony za przełomową technologię, oddziałującą w najbliższych latach na rozwój sprzętu wojskowego" (Fehler i in., 2021, s. 282).

„Panuje powszechne przekonanie, że organizacje terrorystyczne mogą wykorzystywać sztuczną inteligencję do przeprowadzanych przez siebie zamachów. Wśród organizacji, które dysponują wystarczającymi środkami finansowymi, by uzyskać dostęp do tak zaawansowanych technologii, na szczególną uwagę zasługują następujące: al-Kaida, ISIS, Hamas, Hezbollah, talibowie (Taliban), Partia Pracujących Kurdystanu (PKK), Palestyński Islamski Dżihad, Kata'ib Hezbollah, Lashkar-e-Tayyiba i Boko Haram. (...) Gdyby terroryści uzyskali dostęp do broni kontrolowanej przez sztuczną inteligencję lub wspomaganej jej algorytmami, znacznie zwiększyłoby to zagrożenie dla społeczności międzynarodowej. Po pierwsze, skutkowałoby to zwiększeniem skuteczności przeprowadzanych przez nich zamachów. Drugim efektem mogłoby być zmniejszenie »zapotrzebowania« na zamachowców-samobójców. Po trzecie, istnieje możliwość, że organizacjom terrorystycznym łatwiej byłoby uzyskać poufne informacje o siłach zbrojnych państw je zwalczających poprzez operacje hakerskie wspierane przez sztuczną inteligencję" (Olech, Lis, 2021, s. 90–91).

Wraz z rozwojem AI urzeczywistniły się Orwellowskie wizje z powieści pt. *1984*. Doskonałym tego przykładem są ChRL, które zbudowały system rozpoznawania twarzy. Według przedstawicieli firmy zajmującej się tą problematyką system z ponad 95% skutecznością potrafi zidentyfikować osoby noszące na twarzy maseczki (Yang, 2020). „System jest wystarczająco szybki, aby przeskanować populację Chin w ciągu zaledwie jednej sekundy, a przeskanowanie każdego mieszkańca planety zajmuje mu tylko dwie sekundy, z dokładnością sięgającą 99,8%, twierdzi jego główny programista Yuan Peijiang" (Jinping, 2018). Efektem tak daleko posuniętego nadzoru nad społeczeństwem było zablokowanie 11,14 mln lotów obywateli tego państwa, a 4,25 mln nie mogło przemieszczać się szybką koleją. Były zastępca dyrektora chińskiego centrum badań nad rozwojem przy Radzie Państwa, Hou Yunchun, powiedział, że krajowy system kredytów społecznych musi karać osoby nie wywiązujące się z zobowiązań (PYMNTS, 2018). Dodajmy, że w Stanach Zjednoczonych funkcjonuje bliźniaczy system rozpoznawania twarzy – Amazon Rekognition. „Z zasobów Rekognition już dziś korzystają tamtejsze organy ścigania (zwłaszcza w zakresie identyfikacji podejrzanych" (Stój, 2019).

Podsumowanie

Dla rozwoju społeczno-gospodarczego AI jest narzędziem, w którym tkwi potencjał o nieograniczonych możliwościach. Stwarzający możliwości rozwoju na skalę do tej pory niespotykaną. Skutkiem ubocznym jest pewien margines wykorzystywany do celów cyberprzestępczości. W tych jednak przypadkach sztuczna inteligencja okazuje się być „bronią" obosieczną, bowiem przy jej pomocy walczy się z przestępczością zarówno w świecie wirtualnym, jak i realnym. Jeszcze kilkadziesiąt lat temu to, co zagrażało społeczeństwom, miało charakter lokalny i ograniczało się do klęsk żywiołowych (powodzie, pożary, epidemie, itp.) i ewentualnych konfliktów zbrojnych. Wraz z rozwojem cywilizacyjnym spektrum zagrożeń znacznie się poszerzyło. Dziś jako jednostki społeczne i całe społeczeństwo funkcjonujemy w świecie komputerowych bitów, i to w skali globalnej. Współczesność wymaga od nas coraz większego stopnia uwagi, którą winniśmy skupiać na tym, co i w jaki sposób robimy posługując się narzędziami ICT. Teraz zakres uwagi musimy poszerzyć o sztuczną inteligencję. Trzeba będzie liczyć się z pojawiającą się manipulacją oraz coraz bardziej wyrafinowanymi metodami oszustw. Prof. M. Kosiński z Stanford University konkluduje „Bardzo ciężko jest stworzyć nową technologię, która będzie miała wpływ na świat, ale z drugiej strony – mieć pewność, że ten wpływ będzie tylko pozytywny" (Kosiński, 2021).

Bibliografia

1. *AI pomaga hakerom w cyberatakach – są pierwsze dowody,* (2023), https://managerplus.pl/ai-pomaga-hakerom-w-cyberatakach-sa-pierwsze-dowody-33876 (dostęp: 10.09.2023).
2. Duch, W. (1997). *Fascynujący świat komputerów*. Poznań: Wyd. Nakom.
3. Fehler, W., Araucz-Boruc, A., Dana A., Lasota-Kapczuk, A. (2021). Systemy sztucznej inteligencji jako wyzwanie dla sfery bezpieczeństwa i obronności RP. *Zeszyty Prawnicze Biura Analiz Sejmowych Kancelarii Sejmu*, nr 2(70).

4. Gałabuda, M. (2019). *Jak działa system MCAS w boeingach 737 MAX?,* https://www.pasazer.com/news/40893/jak,dziala,system,mcas,w,boeingach,737,max.html (dostęp: 3.09.2023).
5. Goban-Klas, T. (2007). Nadchodzące społeczeństwo medialne. *Chowanna*, t. 2 (29), Katowice: UŚ.
6. *Historia oszukiwania ludzi przez sztuczną inteligencję,* (2022). https://prnews.pl/historia-oszukiwania-ludzi-przez-sztuczna-inteligencje-467747
7. https://smart.gov.pl (dostęp: 9.09.2023).
8. Jinping X. (2018). *China's Surveillance System Can Scan Population 'in 1 Second',* https://nationalinterest.org/blog/the-buzz/chinas-surveillance-system-can-scan-population-%E2%80%98-1-second%E2%80%99-25743 (dostęp: 31.8.2023).
9. Kanade, V. (2022). *What Is Artificial Intelligence (AI)? Definition, Types, Goals, Challenges, and Trends in 2022,* https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ai/amp/ (dostęp: 3.09.2023).
10. Kosiński, M. (2021). *Jak daleko sięga władza algorytmów?,* https://www.kongresobywatelski.pl/wp-content/uploads/2021/12/ko-michal_kosinski-jak_daleko_siega_wladza_algorytmow.pdf (dostęp: 31.8.2023).
11. Kulas, T. (2019). *Co to jest deepfake?,* https://mitsmr.pl/b/co-to-jest-deepfake/PqAu1X2m1 (dostęp: 26.05.2023).
12. McIlwraith, D., Marmanis, H., Babenko, D. (2017). *Inteligentna sieć. Algorytmy przyszłości*. Gliwice: Wyd. Helion.
13. Mulawka, J. (1996). *Systemy ekspertowe.* Warszawa: Wyd. WNT.
14. Nowak-Nova, D. (2021). *Sztuczna inteligencja a obliczenia kognitywne,* https://nowak-nova.pl/sztuczna-inteligencja-a-obliczenia-kognitywne/ (dostęp: 1.09.2023).
15. Olech, A.K., Lis A. (2021). Wykorzystanie nowych technologii przez terrorystów na przykładzie dronów i deep fake'ów. *Wiedza Obronna*, Vol. 275 No. 2, Wyd.
16. ORACLE, *Czym jest uczenie głębokie?,* https://www.oracle.com/pl/artificial-intelligence/machine-learning/what-is-deep-learning/ (dostęp: 3.09.2023).
17. Przegalińska, A. (2022). Współpracująca sztuczna inteligencja. Przykład wirtualnych asystentów i konwersacyjnej AI. W: *Sztuczna inteligencja (AI) jako megatrend kształtujący edukację. Jak przygotowywać się na szanse i wyzwania społeczno-gospodarcze związane ze sztuczną inteligencją?,* red. J. Fazlagić. Warszawa: Wyd. IBE.
18. PYMNTS. (2018). *Former China Official Says Social Credit System Must Make People Bankrupt,* https://www.pymnts.com/news/international/2018/china-official-social-credit-system-bankrupt/ (dostęp: 31.8.2023).
19. Sienkiewicz, P. (1999). *Ewolucja informatyki i jej wojskowych zastosowań.* W: „Biuletyn Jubileuszowy (nr 2)". Warszawa: Wyd. Centrum Informatyki Sztabu Generalnego WP.
20. Stój, E. (2019). *Amazon w obronie Rekognition: technologii rozpoznawania twarzy,* https://www.purepc.pl/amazon-w-obronie-rekognition-technologii-rozpoznawania-twarzy (dostęp: 31.8.2023).
21. Strelau, J. (1987). *O inteligencji człowieka.* Warszawa: Wyd. Wiedza Powszechna.
22. Szulczewski, G. (2019). Sztuczna inteligencja a inteligencja moralna. Zagadnienia wstępne cybernetyki. *Annales. Ethics in Economic Life*, Vol. 22, No. 3. Łódź: Wyd. Uniwersytetu Łódzkiego.
23. Tadeusiewicz, R. (2020). Archipelag sztucznej inteligencji. *Część 1. Napędy i sterowanie*, Nr 12. Racibórz: Wyd. Druk-Art. S.C.
24. Tadeusiewicz, R., Szaleniec, M. (2015). *Leksykon sieci neuronowych.* Wrocław: Wyd. Fundacji „Projekt Nauka".

25. Tałach, S. (2023). "*Klonowanie głosu" w rękach oszustów. Tak wyłudzają pieniądze, mają nowy sposób,* https://biznes.interia.pl/gospodarka/news-klonowanie-glosu-w-rekach-oszustow--tak-wyludzaja-pieniadze-m,nId,7006885 (dostęp: 10.09.2023).
26. Yang, Y. (2020). *How China built facial recognition for people wearing masks,* https://arstechnica.com/tech-policy/2020/03/how-china-built-facial-recognition-for-people-wearing--masks/ (dostęp: 31.8.2023).

dr hab. Aleksander PIECUCH, prof. UR
Uniwersytet Rzeszowski, Kolegium Nauk Społecznych