

# 2023 – rok sztucznej inteligencji



Kolaż autora tekstu wykonany z grafik wygenerowanych przez sztuczną inteligencję

Wszystko wskazuje na to, że bieżący rok będzie pierwszym w historii świata, w którym maszynowa „inteligencja” zacznie zastępować ludzi w pisaniu programów, ekspertyz czy też artykułów do gazet. To już się dzieje – od momentu uruchomienia przez organizację OpenAI usługi internetowej ChatGPT, czyli bota zdolnego do nawiązania dialogu z człowiekiem i generowania sensownych tekstów napisanych zgrabnym, literackim językiem, znalazło się już wiele firm, które zaczęły komercyjnie wykorzystywać jego zdolności.



**Jacek Grabowski**

z wykształcenia specjalista gazownictwa i górnictwa naftowego, przygodę z informatyką rozpoczął w końcu lat 80. XX wieku od współpracy z wydawnictwem „Lupus”, gdzie publikował teksty głównie w dwutygodniku „PCkurier” i miesięczniku „Enter”. Współtwórca pierwszego w Polsce informatycznego czasopisma B2B „MRK” (1997). Był redaktorem naczelnym miesięcznika „Reset”, współpracownikiem wielu innych tytułów (magazyn „WWW”, „IT Reseller”, „Komputer Świat”). Obecnie freelancer, współpracuje m.in. z warszawską komunikacją miejską.



ChatGPT może „wygryźć” słabego programistę, a także dziennikarza. Dyrektor generalny koncernu Axel Springer zapowiedział już, że AI może w niedalekiej przyszłości zastąpić pracowników jego firmy. Czy mamy czego się bać?

Minęły 73 lata od chwili, kiedy Alan Turing wymyślił test inteligencji maszynowej, który nazwał „imitation game”, czyli grą w udawanie. Test polegał na tym, że maszyna miała udawać człowieka w rozmowie z innymi ludźmi. Jeśli sędziowie testu nie byłiby zdolni do odróżnienia wypowiedzi maszyny od wypowiedzi ludzi, to test byłby zaliczony. Oczywiście dla ówczesnych komputerów taki test był niemożliwy do przejścia, Turing zakładał jednak, że do 2000 r. powinna pojawić się maszyna, która oszuka co najmniej 1/3 sędziów biorących udział w teście.

### Trudne narodziny sztucznej inteligencji

Kiedy Turing zaproponował swój test, termin „sztuczna inteligencja” jeszcze nie istniał. Przyjmuje się, że po raz pierwszy użył go John McCarthy w 1955 r. podczas konferencji w Dartmouth, definiując SI jako maszynę, której działanie można określić jako podobne do ludzkiej inteligencji. Jednak mimo powstania testu i definicji SI, droga do zbudowania komputera i oprogramowania spełniającego warunki stawiane przez Turinga i McCarthy’ego była długa, trudna i wyboista, miała wiele etapów i punktów, z których trzeba było zawracać.

Już w 1966 r. powstał pierwszy działający chatbot o imieniu ELIZA, odpowiadający na proste pytania – jego zadaniem było wspomaganie psychoterapii. ELIZA daleka była jednak do wygrania testu Turinga (pacjentów informowano, że rozmawiają z maszyną), podobnie jak i późniejsze podobne systemy, np. MYCIN, powstały w 1973 r. system wspomagający diagnostykę lekarską. Brak większych sukcesów spowodował nawet spadek zainteresowania sztuczną inteligencją, nazywaną „zimą SI”, który trwał przez drugą połowę lat 70.

Uczenie maszynowe zgrane z coraz szybszym wzrostem mocy obliczeniowej pozwoliło w latach 90. osiągnąć kolejne sukcesy SI: w 1997 r. człowiek – arcymistrz Garri Kasparow – po raz pierwszy przegrał w szachy z komputerem i oprogramowaniem IBM o nazwie Deep Blue. Był to wówczas wielki triumf algorytmów uczenia maszynowego. Kasparow uważał zresztą, że przegrał z człowiekiem udającym maszynę, twierdząc, że sposób gry Deep Blue był „zbyt ludzki”.

Lata 80. przyniosły triumfalny powrót sztucznej inteligencji, do czego przyczyniły się zwłaszcza szybko rozwijające

się systemy ekspertowe dla biznesu. Doskonalenie tych systemów doprowadziło w końcu do rozwoju uczenia maszynowego (*machine learning*), czyli algorytmów uczących się automatycznie na podstawie wprowadzanych danych. W 1989 r. za sprawą Yanna LeCuna pojawiła się także konwolucyjna, wielowarstwowa sieć neuronowa wzorowana na siatkówce ludzkiego oka, którą wykorzystano do maszynowego rozpoznawania obrazów.

### XXI wiek – czas głębokiego uczenia

Po sukcesie Deep Blue rozpoczął się żywiołowy rozwój różnych systemów i technik sztucznej inteligencji. Był to wynik postępu technicznego w dziedzinie przetwarzania danych, a także rosnącej błyskawicznie ilości danych cyfrowych, którymi można było „faszerować” systemy SI. W 2012 r. Geoffrey Hinton z Uniwersytetu w Toronto razem z dwoma studentami stworzył głęboką, wielowarstwową sieć neuronową „AlexNet”, która zwyciężyła konkurs rozpoznawania obrazów. Od tego momentu zauważono korzyści płynące z zastosowania głębokich sieci neuronowych w uczeniu maszynowym, nazywając tę technikę „głębokim uczeniem” (*deep learning*).

7 czerwca 2014 r., na imprezie zorganizowanej z okazji 60. rocznicy śmierci Alana Turinga, udający 13-letniego ukraińskiego chłopca chatbot „Eugene Goostman”, stworzony przez rosyjskich i ukraińskich programistów, oszukał jedną trzecią sędziów i tym samym spełnił warunek zdania testu Turinga. Sam Turing, jak pamiętamy, przewidywał, że stanie się to w roku 2000, więc niewiele się pomylił. Zauważono jednak, że wygranie „gry w udawanie” ułatwiło chatbotowi imitowanie osoby niepełnoletniej, dzięki czemu mógł łatwiej ukrywać swoje niedostatki.

Kolejnym przełomowym momentem, pokazującym efektywność głębokiego uczenia z zastosowaniem sieci neuronowych, była w 2015 r. wygrana programu „AlphaGo” z mistrzem gry w Go, Fan Hui. Warto zauważyć, że Go to gra trudniejsza dla sztucznej inteligencji od szachów. Innym ważnym wydarzeniem była wygrana programu Libratus w pokera z czterema zawodowymi graczami – oznaczało to, że zdolności Libratusa pozwalają mu rozumować na podstawie niepełnych albo wręcz fałszywych informacji („blef”). Z kolei w roku 2017 chiński robot medyczny XiaoYi zdał oficjalny egzamin lekarski (China’s National Medical Licensing Exam) z bardzo dobrym wynikiem 456 punktów na 600 możliwych, przy minimum wymaganym do zdania 360 pkt.

## Model GPT – mózg najnowszych chatbotów

I w ten oto sposób drobnymi kroczkami dotarliśmy do punktu, kiedy w 2018 r. pojawiły się pierwsze próby z wykorzystaniem opracowanego w organizacji OpenAI modelu GPT. GPT (ang. *Generative Pre-trained Transformer*) to rodzaj modelu uczenia maszynowego, używanego do rozpoznawania, interpretacji i generowania tekstu w języku naturalnym. Algorytmy GPT są wstępnie wyszkolone na dużym zbiorze danych tekstowych (ostatnia wersja GPT-3 na 570 GB tekstów, głównie z angielskiej Wikipedii i dwóch korpusów książek).

” *W trakcie szkolenia model GPT uczy się przewidywać następane słowo we fragmencie tekstu, co pozwala mu dobrze sobie radzić z zadaniami nawet o ograniczonej ilości danych. Zastosowano przy tym uczenie nienadzorowane, odbywające się z minimalną ingerencją człowieka, gdzie dane są bez etykiet i nie jest znany pożądaný wynik. Zakłada się, że taki sposób najbardziej przypomina ludzkie uczenie się.*

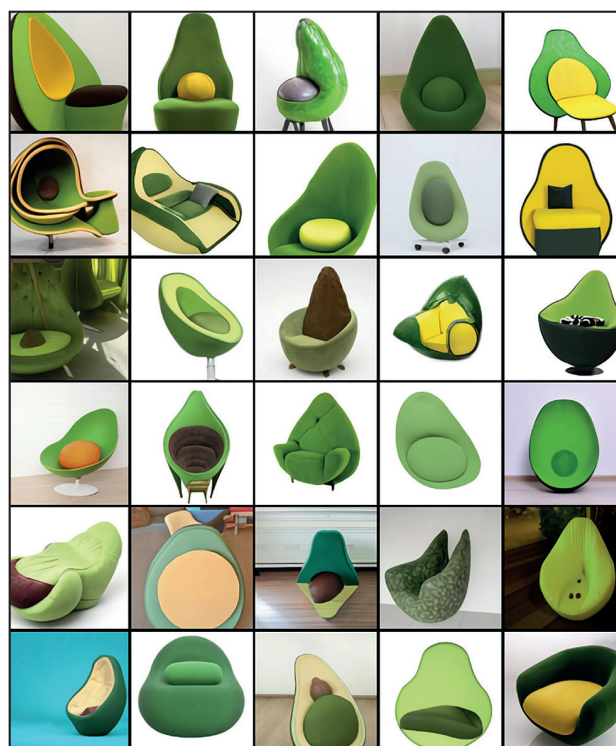
OpenAI rozwijała modele GPT przez kilka lat, udoskonalając ich działanie, ale przede wszystkim karmiąc je coraz większą ilością danych. Już w 2020 r. beta testerzy donosili o ciekawych możliwościach botów opartych na tym modelu, jednak na ich udostępnienie większej liczbie użytkowników trzeba było poczekać do 2022 r. Pod koniec września OpenAI zdecydowała się otworzyć szeroki dostęp do robota DALL-E 2, który tworzy obrazy na podstawie opisu wprowadzonego przez użytkownika.

## DALL-E: pierwsze mocne uderzenie GPT

Nazwa DALL-E to połączenie imienia animowanego robota WALL-E mającego zdolność okazywania uczuć i nazwiska hiszpańskiego artysty surrealisty Salvadora Dali. DALL-E 2 może generować obrazy w wielu stylach, w tym obrazy fotorealistyczne, obrazy imitujące malarstwo, obrazy z modelami 3D i emotikony. Może też manipulować i przestawiać obiekty na obrazach i zdjęciach. Wystarczy wprowadzić zapytanie (*prompt*), zawierające opis tego, co ma zawierać powstający obraz.

Siłą rzeczy jakość i sposób sformułowania zapytania wpływa na końcowy wynik działania DALL-E. Wykorzystali to

użytkownicy, tworząc abstrakcyjne zapytania typu: „fotel w kształcie avocado” i otrzymując niezwykle wyniki w postaci dość dziwacznych rysunków. Ten rodzaj zabawy technologią wpłynął dość mocno na opinię o robocie. Tymczasem przy konkretnym zapytaniu DALL-E rysuje profesjonalnie, a użytkownicy zwrócili również uwagę na realizm jego rysunków, bowiem np. potrafi on różnicować wygląd przedmiotów w czasie: telefon z 1999 r. na jego obrazach odpowiednio różni się od telefonu współczesnego.



Źródło: <https://openai.com/research/dall-e>

Udostępnienie DALL-E wywołało sporo problemów. Chatbot jest tak wytrenowany, żeby np. omijać tematy gwałtów i seksu, jednak zauważono, że wpływa to na „myślenie” algorytmu o płciach i ograniczenie liczby kobiet na rysunkach, co prawdopodobnie wynika z większej seksualizacji płci żeńskiej w danych używanych do szkolenia robota. Obawiano się również prób używania robota do tworzenia deepfake’ów i dezinformacji, dlatego też wprowadzane zapytania są odpowiednio „cenzurowane”. Okazało się szybko, że użytkownicy jednak dość łatwo znajdują zamienniki cenzurowanych terminów i omijają część zabezpieczeń. Jest to zresztą ogólny problem wszystkich współczesnych chatbotów, dotyczący także tekstowego ChatGPT.

OpenAI pod koniec zeszłego roku udostępniła API do DALL-E 2, więc usługa ta pojawiła się także w innych miejscach, np. Microsoft umieścił ją w wyszukiwarce Bing. Poza tym w internecie jest jeszcze kilka podobnych, lecz niez-

leżnie rozwijanych botów generujących obraz z tekstu, np. Midjourney, uważany za najlepszy tego typu. Za jego pomocą m.in. wygenerowano obrazek na okładkę jednego z zeszłorocznych numerów „The Economist”. Większość usług tego typu jest płatna, sam DALL-E jest bezpłatny do wyczerpania limitu darmowych obrazków. Płatność zależy często od rozdzielczości generowanego obrazu.

## ChatGPT – nokaut?

Naturalność używanego przez ChatGPT języka i łatwość przetwarzania informacji z internetu (opisywania „własnymi zdaniami”) sprawiają, że entuzjaści sztucznej inteligencji wieszczą rewolucję wszędzie tam, gdzie występuje tekst w formie pisanej.

- Zasadniczą częścią bazy wiedzy ChatGPT jest Common Crawl – zestaw materiałów zgromadzonych na przestrzeni 12 lat z ponad 60 mln domen internetowych. Stanowi on 60 proc. danych używanych do szkolenia. Wikipedia – 3 proc.
- System wykorzystuje 175 mld parametrów.
- Do treningu systemu w języku polskim wykorzystano m.in. materiały „Gazety Wyborczej”, „Rzeczpospolitej”, „Polityki”, „Newsweeka”, „Wprost”, Onetu, Interii, WP, TVN24, RMF FM, czy TOK FM.
- Jednym z twórców OpenAI pracującym nad następcami GPT-3 oraz sztuczną inteligencją tworzącą oprogramowanie jest Polak, Wojciech Zaremba.

W listopadzie 2022 r. OpenAI udostępniła szerokiej publiczności (po zarejestrowaniu) chatbota ze sztuczną inteligencją, który od razu zwrócił na siebie uwagę klarownością i elegancją generowanego tekstu, a także szerokim zakresem wiedzy z wielu dziedzin. Oparty na dodatkowo dotrenowanym modelu GPT-3 robot jest zdolny nie tylko do prowadzenia naturalnego dialogu z człowiekiem, lecz także do pisania np. wypracowań szkolnych czy nawet dłuższych esejów. Stworzy wiersze, teksty piosenek, a nawet scenariusze filmowe. Potrafi także pisać kod stron internetowych i aplikacji w różnych językach programowania, takich jak JavaScript, Python i React. Może również zdebugować wygenerowany przez siebie kod. Bot ma zdolność przyznawania się do swoich błędów, odpowiadania na dodatkowe pytania, kwestionowania błędnych

przesłanek i odrzucania nieodpowiednich próśb. Rozmawia w różnych językach, w tym polskim, może też dokonywać tłumaczeń.

W pierwszym okresie po udostępnieniu ChatGPT internet oszalał na jego punkcie. Jednak dość szybko pojawiły się pierwsze wątpliwości. ChatGPT nie tylko potrafił się paskudnie mylić, ale nawet... zmyślał. Są udokumentowane przypadki, kiedy robot podawał fałszywą odpowiedź, a w dalszym ciągu konwersacji powoływał się na nieistniejące w rzeczywistości prace naukowe, które wymyślił! Co ciekawe, na bezpośrednie pytanie, czy zmyślił te prace, odpowiedział szczerze, że tak. W mediach społecznościowych opublikowano w formie zrzutów ekranu wiele dowodów takiego i podobnego zaskakującego działania bota.

Na przykład na pytanie: „gdzie mieszka Murzynek Bambo?” chatbot odpowiedział: „Murzynek Bambo to postać z polskiej literatury dziecięcej stworzona przez Marię Konopnicką. Jest to postać fikcyjna, więc nie mieszka on nigdzie konkretnie (...)”. Po doprecyzowaniu: „czy jesteś pewien, że tę postać stworzyła Maria Konopnicka?” odpowiedź brzmiała: „Przepraszam za moją pomyłkę. Rzeczywiście postać Murzynka Bambo pochodzi z wiersza (...) Władysława Bełzy, który opublikowano po raz pierwszy w 1902 roku”. Gdy ponownie zaskoczony odpowiedział użytkownik spytał: „A to nie jest wierszyk napisany przez Juliana Tuwima?”, chatbot uparł się i odpisał: „Nie, wiersz „Murzynek Bambo” został napisany przez Władysława Bełzę w 1902 roku, nie przez Juliana Tuwima (...)”. Skriny o takiej treści opublikował na Twitterze użytkownik o nicku „Przewodzik”.

” *Grupa polskich naukowców z Wydziału Informatyki i Telekomunikacji Politechniki Wrocławskiej w ramach projektu CLARIN sprawdziła wszechstronność wiedzy ChatGPT i precyzję jego odpowiedzi. Zadali chatbotowi ponad 38 tys. pytań w 25 kategoriach. Wyniki testu dobrze oddaje tytuł ich 40-stronicowej pracy: „Jack of all trades, master of none” („Spec od wszystkiego, mistrz niczego”).*

Przykład może być błahy, ale system udziela również takich błędnych odpowiedzi, które mogą przyczynić się do realnej krzywdy człowieka. Chodzi tu m.in. o nonsensowne porady medyczne czy niedozwolone metody optymalizacji podatkowej. ChatGPT potrafi się również pomylić przy rozwiązywaniu najprostszych zadań matematycznych. Dociekliwi testerzy odkryli również, że model ten posiada wbudowany mechanizm politycznej poprawności. Twórcy systemu

mogą mieć również kłopoty z powodu wykorzystania do treningu materiałów chronionych prawem autorskim<sup>1</sup>.

Czy jednak oznacza to, że ChatGPT to porażka? Wcale nie, to oprogramowanie przecież cały czas się... uczy. Chatbot każdego dnia jest trochę lepszy. Obecnie może np. napisać program na poziomie słabego amatora i nie poradzi sobie ze złożonym kodem wymaganym w profesjonalnych zastosowaniach, jednak programiści, którzy analizowali wyniki jego działań, są zgodni co do tego, że przy tym tempie udoskonalania najdalej za pięć lat bot będzie mógł generować programy na pełnym poziomie profesjonalnym.

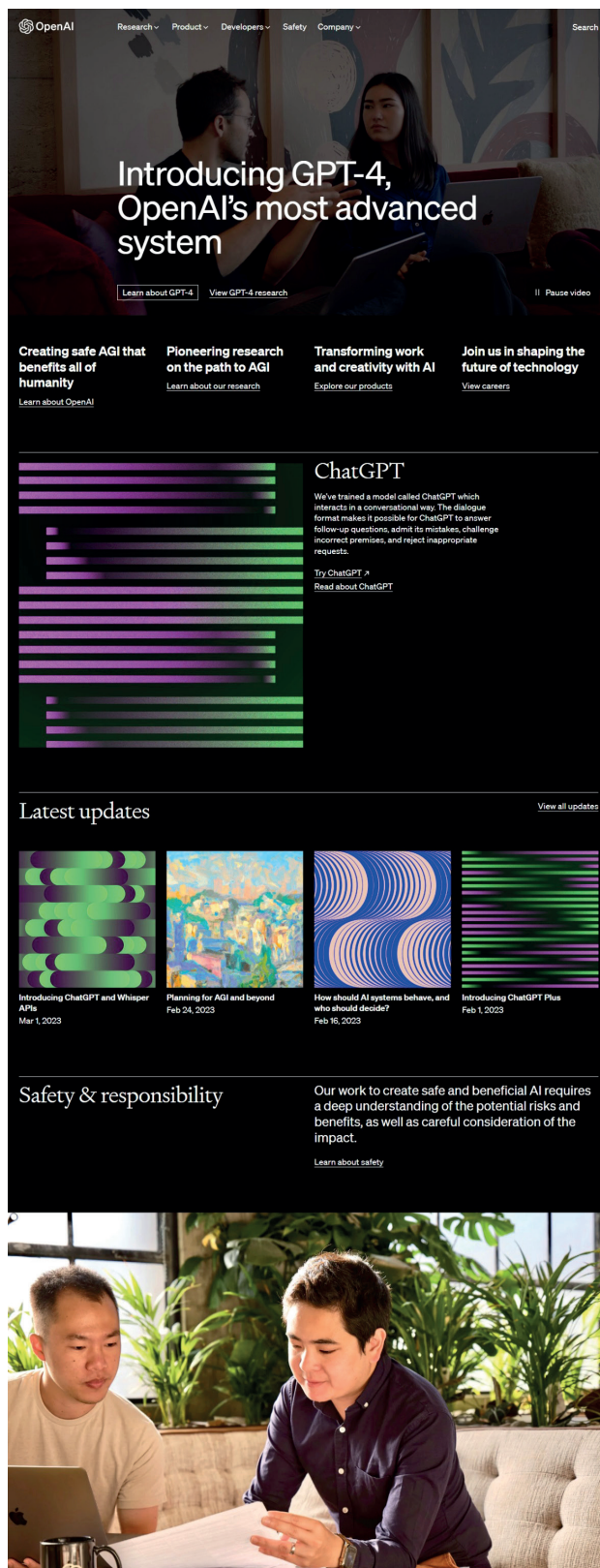
## Początek rewolucji

Bez żadnej wątpliwości można powiedzieć, że jesteśmy świadkami początku rewolucji sztucznej inteligencji. „Rewolucja” na razie nie oznacza walki ze zbuntowanymi inteligentnymi maszynami na ulicach. Póki co oznacza bardziej zmianę naszych przyzwyczajeń i sposobów wyszukiwania informacji w internecie. W dalszej perspektywie zapewne również zmianę sposobu komunikacji z urządzeniami na bardziej naturalny, za pomocą swobodnego dialogu.

” *Jednak czy będziemy chcieli coś czytać, rozmawiać o czymś, poznawać świat, jeśli o wszystko będziemy mogli zapytać swoją lodówkę?*

To nie do końca żart, bo usługa sztuczna inteligencja na pewno ma potencjał, żeby rozłeniwić nas intelektualnie. Na pewno też dysponuje możliwościami, by zastąpić nas w wielu miejscach pracy. Przy obecnym tempie rozwoju tej niby od dawna znanej, ale w obecnym kształcie nowej technologii SI, możemy przypuszczać, że za jakąś dekadę roboty zaczną zastępować nawet menedżerów niższego szczebla i wydawać polecenia pracownikom.

Trudno w tej chwili przewidywać, do czego dokładnie doprowadzi rewolucja sztucznej inteligencji. Jedynym, co może ją przyhamować, jest – paradoksalnie – ekologia i trudności na rynku energetycznym, bowiem są to rozwiązania podobnie energochłonne jak kopalnie kryptowalut, również wykorzystujące do obliczeń procesory graficzne. Jeśli jednak wszystko będzie rozwijało się stabilnie i bez przeszkód, to być może już za pięć lat nie poznamy się w lustrze.



<sup>1</sup> <https://www.traple.pl/chatgpt-w-praktyce-najwazniejsze-kwestie-prawne/>