Sztuczna inteligencja – o co gramy?

*Tomasz KAZANA**

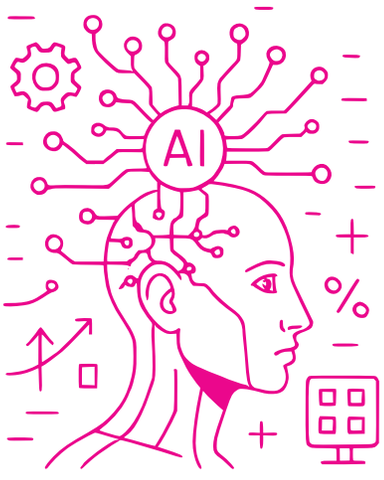
* Osoba fizyczna

Nie da się nie zauważyć ogromnego postępu, jaki dokonał się w świecie sztucznej inteligencji w ostatnim czasie. I chyba nie da się również nie zadać sobie najważniejszych pytań: Dokąd to zmierza? Jak daleko to zajdzie?

Co zastanawiające, bardzo wielu cenionych ludzi zdaje się mieć wyraźne i jednoznaczne zdanie na ten temat. I co równie ciekawe, mocno wzajemnie spolaryzowane.

AI to skrót od angielskiego *artificial intelligence*, czyli sztuczna inteligencja. Mimo, że w tekście po polsku chciałoby się pisać SI, to nie będziemy szli pod prąd i – zgodnie z przyjętą w tekstach polskojęzycznych niepisaną konwencją – będziemy używać skrótu AI.

Z jednej bowiem strony Elon Musk jasno stwierdza, że AI już lada dzień po prostu „przewyższy człowieka praktycznie we wszystkim". Wtórują mu takie osobistości jak Sam Altman („AI rozwiąże problemy, które dziś wydają się niemożliwe do rozwiązania") czy Andrew Ng („AI jest jak elektryczność: będzie fundamentem dla każdej dziedziny życia"). Demis Hassabis nie ma wątpliwości, że „AI przyspieszy odkrycia naukowe i pomoże rozwiązać fundamentalne pytania natury" – innymi słowy, AI nie tylko zastąpi nas w działaniach rutynowych, ale i tych wymagających skrajnie twórczej i kreatywnej aktywności. Zresztą już znacznie wcześniej Ray Kurzweil przewidywał, że do roku 2045 osiągniemy *osobliwość* (zob. Δ^{11}_{18} na stronie 20), a AI przewyższy inteligencję zbiorową całej ludzkości.



Z drugiej strony, Bill Gates stwierdził, że AI nie zastąpi nawet samych programistów w ciągu 100 lat – a to dlatego, że „pisanie programów wymaga kreatywności, osądu i głębokiego myślenia. A sztuczna inteligencja – choć może pomagać w prostszych zadaniach, jak szukanie błędów, to »prawdziwa« twórcza praca (kompletny projekt, podejmowanie złożonych decyzji) pozostaje domeną ludzkiego umysłu". Podobne zdanie wyrażają Gary Marcus („AI nie rozumie świata, tylko dopasowuje wzorce; obecne sukcesy są powierzchowne"), Hubert Dreyfus („Inteligencja wymaga cielesnego doświadczenia; komputerom zawsze będzie tego brakować") czy Noam Chomsky („AI manipuluje statystykami, ale nie rozumie języka ani znaczenia, więc nie dorówna człowiekowi"). Najbardziej skrajnym sceptykiem jest jednak Roger Penrose (laureat Nagrody Nobla z fizyki z roku 2020), który uważa, że „AI nigdy nie osiągnie ludzkiego poziomu rozumienia, a świadomości nie da się sprowadzić do algorytmów, bo po prostu ludzki umysł nie działa jak komputer", czy bardziej poetycko: „człowiek to więcej niż mokry komputer". Jego przemyślenia idą zresztą znacznie dalej. Penrose twierdzi, że „świadomość ma głębsze podstawy fizyczne, być może związane z procesami kwantowymi w neuronach" (tzw. hipoteza Orch-OR, rozwinięta wspólnie ze Stuartem Hameroffem). Mówi też wprost, że maszyny nigdy „nie zdobędą prawdziwego rozumienia ani intuicji matematycznej, które cechują ludzi".

Poza zarysowanym sporem o potencjał AI równolegle prowadzona jest jeszcze jedna dyskusja. Otóż wciąż nie potrafimy ocenić, jak realne ryzyko niesie ze sobą oddanie kontroli nad ogromną częścią świata maszynom zaprogramowanym za pomocą AI. Czy nieuchronnie musi się to skończyć jak w każdym dobrym filmie o buncie maszyn?

My, redakcja *Delty*, nie udajemy, że wiemy, co z tego wyniknie. Zresztą nie ma wśród nas ekspertów z dziedziny AI, nie chcemy więc też obstawiać, kto ostatecznie okaże się mieć rację – Roger Penrose czy wizjonerzy z Doliny Krzemowej. Obiecujemy za to, że nie będziemy kibicować żadnej ze stron, lecz rzetelnie relacjonować rozwój wypadków. Spróbujemy potraktować to, co się dzieje, jako jeden wielki eksperyment przeprowadzany przez ważnych i możnych tego świata. I jak w każdym uczciwym eksperymencie – nie będziemy zakładać z góry jego wyniku!

Model LLM (Large Language Model) to sztuczna sieć neuronowa wytrenowana na ogromnych zbiorach tekstu, zdolna do generowania i rozumienia języka naturalnego w sposób zbliżony do ludzkiego.

Oczywiście świadomi jesteśmy stawki tych zawodów. Jest tam przecież i wymiar transcendentalno-religijny, i dotykający gigantycznych ego jednych, ale i egzystencjalnych lęków innych. Nie wskazujemy więc, kto nie ma pokory, kto jest naiwny, kto chciwy, a kto mądry czy głupi. Sami zresztą z nieukrywanymi wypiekami na twarzach będziemy czekać na rozstrzygnięcie, czy powstanie w końcu pierwszy dowód jakiegoś ważnego problemu matematycznego wygenerowany przez AI. A może i naszą redakcję uda się kiedyś zastąpić dobrze wyuczonym modelem LLM?

Plan działania

Po tym przydługim wstępie czas na konkrety. Obiecujemy zacząć pisać o AI regularnie. Rozpoczynamy cykl artykułów zatytułowany „Człowiek o AI” (niniejszy tekst można traktować jako zerową część tej serii). Kolejnych zapraszanych autorów będziemy zachęcać do zapoznania się z poprzednimi tekstami oraz do napisania czegoś nowego. Przy czym nie będziemy narzucać niczego konkretnego. Chcemy publikować i polemiki, i opinie filozoficzne; opinie entuzjastów, ale i głosy omawiające zagrożenia; teksty długie i króciutkie notatki; a przede wszystkim artykuły typowo techniczne: o samym procesie uczenia maszynowego, o modelach LLM itp. W tej serii swój udział mogą mieć także Czytelnicy – niniejszym serdecznie zachęcamy do nadsyłania swoich propozycji artykułów.

Oczywiście o AI w *Delcie* już nie raz pisaliśmy, w szczególności o samych jej fundamentach (Δ^{1}_{18}, Δ^{5}_{18}, Δ^{11}_{18}, Δ^{2}_{23}, Δ^{5}_{24}). Zachęcamy do sięgnięcia i po te teksty!

Zapewne wyjaśnić należałoby, co dokładnie rozumiemy przez *napisane/wykonane przez sztuczną inteligencję*. Jednakże nie będziemy tutaj bardzo ściśli – dla nas znaczy to po prostu tyle, że wybranemu silnikowi AI (jak ChatGPT, Grok, Gemini, DeepSeek, Claude, LLaMA czy Mistral) zostanie podane jakieś krótko opisane zadanie, a po jego wstępnym wykonaniu – być może kilka dodatkowych próśb o drobne poprawienie/ulepszenie wskazanego fragmentu przy pomocy dalszych krótkich instrukcji. Tych instrukcji (tzw. promptów) nie będziemy publikować ani szczegółowo omawiać.

W tym miejscu wyraźnie zaznaczmy: nie traktujemy bynajmniej artykułów z serii „AI we własnej osobie” jako równoprawnych tekstów, które uważamy za dorównujące jakością ludzkim wytworom. Wręcz odwrotnie: chcemy tylko bezpośrednio pokazywać, co obecnie daje się zrobić. Prosimy więc do tej serii podchodzić przedmiotowo, a nie podmiotowo. Oraz raczej z okiem przymrużonym niż skupionym.

Dodatkowo głos oddamy też samej AI. Będziemy bowiem publikować co jakiś czas artykuły w pełni napisane oraz rysunki w pełni wykonane przez sztuczną inteligencję. Takie artykuły będziemy oznaczać jako „AI we własnej osobie”.

W tym numerze zapraszamy do lektury artykułów człowieka-raczej-entuzjasty (Piotr Sankowski, *Nadchodzi rewolucja w prowadzeniu badań*, str. 8), człowieka-raczej-sceptyka (Jerzy Marcinkowski, *Osiem krótkich uwag z pamiętnika*, str. 11), człowieka-obserwatora (Marek Kordos, *W szponach metodologicznej poprawności*, str. 14), maszyny piszącej (ChatGPT, *Małe twierdzenie Fermata – o pięknie prostych idei*, str. 6) oraz maszyny rysującej (Grok, okładka).

Miłej lektury!

PS. Dziękuję Chatowi GPT za wydatną pomoc (kwerenda, drobna korekta) w pisaniu tego artykułu.

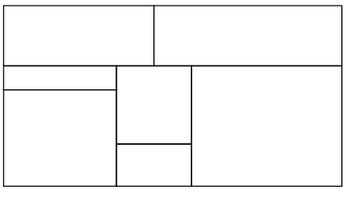
Zadania

Przygotował Patryk MICHALSKI

F 1139. Ołówek o masie m i promieniu r podparto przy końcach ustawionymi prostopadle do niego palcami. Współczynniki tarcia kinetycznego i statycznego między palcem a ołówkiem wynoszą odpowiednio μ i μ_{s}. W pewnym momencie palce zaczęto jednostajnie zbliżać do siebie wzdłuż osi ołówka. Czy zbliżając palce wystarczająco wolno, można sprawić, że ołówek w czasie ruchu nie straci równowagi?

F 1140. Walcowy korek o promieniu r wyciągany jest z prędkością v pionowo z szyjki butelki. Jaką dodatkowo prędkość kątową ω należy nadać korkowi, żeby siła tarcia działająca wzdłuż osi szyjki zmniejszyła się dwukrotnie?

Przygotował Dominik BUREK



M 1846. Na prostokątnym arkuszu papieru poprowadzono kilka odcinków równoległych do jego boków. Odcinki te podzieliły arkusz na prostokąty, wewnątrz których nie ma już narysowanych linii. Wanda chce w każdym z tych prostokątów poprowadzić jedną przekątną, dzieląc go na dwa trójkąty, a następnie pokolorować każdy trójkąt na czarno albo na biało. Czy jest prawdą, że zawsze można to zrobić w taki sposób, aby żadne dwa trójkąty tego samego koloru nie miały wspólnego odcinka brzegowego?

M 1847. W przestrzeni dane są odcinki AA_1, BB_1 oraz CC_1 o wspólnym środku M. Okazało się, że sfera ω, opisana na czworościanie $MA_1B_1C_1$, jest styczna do płaszczyzny ABC w punkcie D. Niech O będzie środkiem okręgu opisanego na trójkącie ABC. Udowodnić, że $MO = MD$.

M 1848. Dane są takie liczby dodatnie a, b, c, że z odcinków o długościach a^{2026}, b^{2026}, c^{2026} można zbudować trójkąt. Udowodnić, że jedną z liczb a, b lub c można podzielić przez 2026, otrzymując liczby a', b', c', tak że z odcinków o długościach a', b', c' również można zbudować trójkąt.

Rozwiązania na str. 24