

Do czego AI nie służy

Przewodnik dla nauczycieli
stworzony przez grupę roboczą ds. AI



IBE



INSTYTUT
BADAN
EDUKACYJNYCH

Zespół autorski:

dr Tomasz Łukawski, *Instytut Badań Edukacyjnych*

Adam Łukawski, *Leiden University i Orpheus Institute*

Mateusz Rafał, *Ministerstwo Edukacji i Nauki*

Współpraca:

dr Łukasz Grzęda, *Instytut Badań Edukacyjnych*

Ewa Kołodziejczyk, *Ministerstwo Edukacji i Nauki*

Dariusz Wronikowski, *Akademia Leona Koźmińskiego w Warszawie*

Agnieszka Halicka, *Szkoła Podstawowa Integracyjna nr 5 w Konstancinie-Jeziornie, EduTriki*

Krzysztof Kalinowski, *Szkoła Podstawowa nr 1 i m. Jana Pawła II w Czerwionce-Leszczynach i Zespół Szkół w Czerwionce-Leszczynach*

Paweł Czerwony, *Google Polska*

Barbara Michalska, *Microsoft Polska*

Właścicielem autorskich praw majątkowych jest Instytut Badań Edukacyjnych.

Utwór udostępniany jest na podstawie licencji Creative Commons:

Uznanie autorstwa-Użycie niekomercyjne 4.0 – licencja ta pozwala na kopiowanie, zmienianie, remiksowanie, rozprowadzanie, przedstawianie i wykonywanie utworu jedynie w celach niekomercyjnych. Warunek ten nie obejmuje jednak utworów wytworzonych przez Chat GPT (mogą zostać objęte inną licencją).

Wstęp	5
I. O kreatywności generatywnej sztucznej inteligencji.....	8
1.1. Generatywna sztuczna inteligencja (GenAI) a kreatywność.....	8
1.2. AI nie przetwarza informacji w taki sam sposób jak ludzie.....	11
1.3. Dla AI nasze pytania to tylko zestaw znaków i wzorów.....	12
1.4. Inżynieria podpowiedzi (ang. prompt engineering).....	13
1.5. Przykłady promptów do wykorzystania w pracy z uczniami.....	14
1.6. Czy sztuczna inteligencja jest kreatywna.....	15
1.7. Tezy przeczące kreatywności AI.....	16
1.8. Do czego nie służy AI w kontekście kreatywności.....	18
1.9. Wybrane sytuacje, w których nie należy korzystać z modeli językowych (LLM).....	19
II. Etyka i moralność w kontekście sztucznej inteligencji.....	22
2.1. Ogólnie o etyce w kontekście AI.....	22
Jak korzystać z AI etycznie i odpowiedzialnie.....	22
Etyczne ramy sztucznej inteligencji wg Komisji Europejskiej.....	23
Kiedy sztuczna inteligencja jest godna zaufania.....	24
2.2. Zdolności maszyn a zdolności ludzi w kontekście etyki i moralności.....	25
AI posiada wiedzę o etyce i moralności, ale czy potrafi ją zastosować?.....	25
Jeżeli AI nie posiada wiedzy, to ją „wymyśli”.....	26
III. Nowe narzędzia i funkcje modeli językowych.....	27
3.1. Nowości w ChatGPT przed konferencją OpenAI DevDay 2023 i po niej.....	27
Nowości na OpenAI DevDay 2023.....	29
GPTs – indywidualnie dostosowane modele ChatGPT: możliwość tworzenia przez użytkowników spersonalizowanych wersji ChatGPT.....	29
GPT Store – platforma dystrybucji GPTs: platforma umożliwiająca udostępnianie i komercjalizację własnych wersji ChatGPT.....	30
Multimodalność ChatGPT – obsługa różnych typów treści: ChatGPT może teraz przetwarzać teksty i obrazy.....	31
Konwersacja głosowa ChatGPT – interakcje głosowe: nowa funkcja ChatGPT, która umożliwia generowanie mowy na podstawie tekstu.....	32
Dall-E – generowanie obrazów: narzędzie do tworzenia obrazów na podstawie opisów tekstowych.....	33
Ulepszona multimodalność w jednym oknie czatu: integracja różnych typów mediów i interakcji w pojedynczym interfejsie użytkownika.....	33
Podsumowanie.....	34
3.2. Wybrane narzędzia AI, których nie opisano w poprzedniej broszurze.....	35
Google Bard.....	35
Bing Chat.....	35
xAI Grok.....	36
Elephas AI.....	36

Hugging Face.....	37
3.3. Pozytywne oraz negatywne aspekty nowych funkcji modeli językowych.....	38
IV. Do czego AI nie służy – szkolne przykłady działań niepożądanych.....	40
4.1. Przykłady niewłaściwego wykorzystywania narzędzi AI przez uczniów, nauczycieli i dyrektorów szkół.....	40
4.2. Zagrożenia przy wdrażaniu rozwiązań AI w edukacji.....	41
Ochrona danych.....	42
Uczeń a AI.....	42
Zdrowie fizyczne i psychiczne.....	45
Nauczyciel a AI.....	46
Etyka i transparentność w korzystaniu z AI.....	48
Dostępność technologii.....	49
Koszty i dostępność.....	49
Zależność od dostawców technologii.....	49
Podsumowanie.....	50
O autorach.....	52

Wstęp

Niniejszą publikację traktujemy jako kontynuację broszury *Chat GPT w szkole*.

*Szanse i zagrożenia*¹, skierowanej do nauczycieli i nauczycielek polskich szkół. Czytelnicy mogli tam uzyskać odpowiedzi na pytania, co to jest sztuczna inteligencja oraz gdzie się ją wykorzystuje. Autor przedstawił wybrane narzędzia wykorzystujące sztuczną inteligencję oraz opisał, czym jest ChatGPT i jak wpływa na edukację. Wskazał również, na co należy zwracać uwagę podczas pracy pedagogicznej z wykorzystaniem bota. Dlaczego skupił się właśnie na tym narzędziu? Publikacja ukazała się w lipcu 2023 roku, kiedy społeczność akademicka i szkolna była poruszona dysrupcyjną zmianą, którą spowodował ChatGPT. Nie tylko zresztą społeczność polska, dowody tego poruszenia (w formie różnorodnych publikacji) można było bowiem znaleźć i w zagranicznych środowiskach związanych z edukacją. Wspomniana broszura miała być więc formą wsparcia ze strony Ministerstwa Edukacji i Nauki, które zanim wydało tekst, otrzymało szereg zapytań, jak sprawdzić, czy autorem pracy jest uczeń, czy ChatGPT. w związku z tym w publikacji nie zabrakło także wskazówek w tej kwestii. Odniesiono się również do opinii naukowych na temat bota i wskazano, jak model ten może wesprzeć osoby ze specjalnymi potrzebami.

Od lipca 2023 roku wiele się zmieniło w związku z dynamicznym rozwojem zaawansowanych modeli językowych, stąd decyzja o wydaniu kolejnej broszury skierowanej do nauczycieli i nauczycielek polskich szkół. Pomimo że jej tytuł brzmi: *Do czego AI nie służy*, skupiamy się w niej wyłącznie na opisie modeli językowych. Ten zabieg jest celowy. Wprowadźcie pojęcie *sztuczna inteligencja* jest znacznie szersze niż termin *duży model językowy* (ang. *large language model* – LLM), o czym piszemy szczegółowo w rozdziale] pierwszym, lecz rewolucja rozwoju generatywnej sztucznej inteligencji, zapoczątkowana eksponencjalnym wzrostem popularności modelu ChatGPT po 30 listopada 2022 roku², spowodowała, że **w powszechnym odbiorze i mowie potocznej AI oraz tzw. chatboty stały się niemal synonimami**. A ponieważ naszym celem jest dotarcie do jak największej grupy odbiorców, to zgodnie z tym trendem **rozumiemy tytułową AI wąsko – jako LLM, mając jednak pełną świadomość, że jest to uproszczenie**.

¹ Machura, M. (2023): *Chat GPT w szkole. Szanse i zagrożenia*, Instytut Badań Edukacyjnych, <https://www.gov.pl/web/ai/chat-gpt-w-szkole-szanse-i-zagrozenia> [dostęp: 16.11.2023].

² ChatGPT zyskał ponad milion użytkowników w ciągu zaledwie 5 dni. Zdobycie 100 milionów użytkowników zajęło mu z kolei dwa miesiące (*Sztuczna inteligencja w mainstreamie. ChatGPT bije rekordy popularności*, CyberDefence24.pl, 9.12.2022 [dostęp: 16.11.2023]).

W niniejszym dokumencie zastanawiamy się nad kreatywnością dużych modeli językowych oraz rozważamy kwestie etyki i moralności w kontekście pracy z nimi. Prezentujemy wskazówki i przestrogi dla nauczycieli i nauczycielek pracujących z modelami, a także – w ostatnim rozdziale – podajemy przykłady niepożądanych zachowań użytkowników AI w szkole oraz definiujemy zagrożenia przy wdrażaniu rozwiązań AI w edukacji.

Jako ciekawostkę traktujemy podsumowanie dokumentu, które zleciliśmy do wygenerowania modelowi językowemu zgodnie z przygotowanym przez nas promptem (wyjaśnienie znajduje się w rozdziale pierwszym).

Tematyka, którą poruszamy, została dobrana tak, aby stanowić odpowiedź na nieustanną zmienność czasów, w których żyjemy. Można pokusić się o stwierdzenie, że jedyna zmienna, która jest i, jak się zdaje, będzie niezmienna, to właśnie **nieustanna zmiana**. Obecnie kompetencje, które nabywamy w szkole i życiu dorosłym, starzeją się szybciej niż kiedykolwiek wcześniej. Dzieje się tak z powodu gigantycznego przyrostu nowych informacji w krótkim czasie. Eksperci od rekrutacji zawodowej okres ważności większości nowo nabytych kompetencji szacują dziś na od sześciu miesięcy do pięciu lat. **Najważniejszym wyzwaniem współczesnej szkoły staje się więc efektywna nauka samej umiejętności uczenia się oraz uczenia się przez całe życie.**

Umiejętność uczenia się pojawia się w wielu definicjach inteligencji. Aaron Sloman, filozof, kognitywista i badacz sztucznej inteligencji, w swoim wypracowaniu *The structure of the space of possible minds* (w tłumaczeniu: *Struktura przestrzeni możliwych umysłów*) zdefiniował inteligencję jako zdolność danej jednostki do osiągnięcia jej celów w danym środowisku³. Interpretując tę definicję, można więc założyć, że **inteligencja to zdolność uczenia się: im lepiej się uczysz, tym inteligentniejszy jesteś.**

Według modelu 4K wyróżnia się cztery umiejętności centralne: **kreatywność, komunikację, krytyczne myślenie i kooperację**⁴. Tę pierwszą sir Kenneth Robinson, lider w dziedzinie rozwoju edukacji innowacyjności i zasobów ludzkich, uznał za tak samo ważną jak nauka pisania i czytania⁵. Do niedawna kreatywność była uważana za kompetencję typowo ludzką, której nie dało się zautomatyzować z pomocą komputerów. Jednak wraz z rozwojem

³ Sloman, A. (1984): *The structure of the space of possible minds*, w: S. Torrance (red.), *The Mind and the Machine: philosophical aspects of Artificial Intelligence* (s. 35-42), Ellis Horwood, <http://www.cs.bham.ac.uk/research/projects/cogaff/sloman-space-of-minds-84.pdf> [dostęp: 23.10.2023].

⁴ Lamri, J. (2021): *Kompetencje XXI wieku. Kreatywność, komunikacja, krytyczne myślenie, kooperacja*, Wolters Kluwer, s. 128.

⁵ Robinson, K., Aronica, L. (2020): *Kreatywne szkoły. Oddolna rewolucja, która zmienia edukację*, Element, s. 43.

technologii i badań nad tzw. generatywną sztuczną inteligencją (szerzej o tym w rozdziale pierwszym) **definicja kreatywności ponownie zaczęła być przedmiotem debat** zarówno środowisk artystycznych, jak i naukowych. W tym kontekście rozważania nad nową kreatywnością i możliwościami sztucznej inteligencji w tym zakresie muszą być przedmiotem pogłębionej analizy również w edukacji.

Tomasz Łukawski

I. O kreatywności generatywnej sztucznej inteligencji

Adam Łukawski

1.1. Generatywna sztuczna inteligencja (GenAI) a kreatywność

Generatywna sztuczna inteligencja (ang. *generative artificial intelligence*), znana również jako GenAI, to jeden z najbardziej fascynujących i dynamicznie rozwijających się obszarów w dziedzinie sztucznej inteligencji. Jest to typ AI, który wyróżnia się zdolnością do kreowania nowych danych, takich jak tekst, obrazy, muzyka czy nawet filmy. GenAI, oparta na zaawansowanych technikach uczenia maszynowego, przekracza granice tradycyjnych modeli, ograniczających się do uczenia na podstawie istniejących wzorców i podejmowania decyzji na tej podstawie. GenAI idzie o krok dalej, nie tylko ucząc się przewidywania podobnych danych, ale także tworząc nowe instancje danych, które naśladują lub reinterpreterują właściwości danych wejściowych.

Zastosowanie generatywnej sztucznej inteligencji jest szerokie i różnorodne. W dziedzinie tworzenia treści kreatywnych GenAI otwiera nowe możliwości dla artystów, pisarzy i muzyków, umożliwiając generowanie dzieł sztuki w stylach znanych artystów, tworzenie poezji czy komponowanie muzyki. **Największy potencjał GenAI zauważamy jednak w nowych rozwiązaniach, które nie tylko naśladują działalność człowieka, ale i całkowicie zmieniają sposób, w jaki treści mogą być generowane.** Na przykład aplikacja internetowa HeyGen umożliwia tworzenie spersonalizowanych cyfrowych reprezentacji użytkowników w postaci awatarów⁶. Dzięki temu można stworzyć film, w którym nasz własny awatar prezentuje dany tekst, działając jak cyfrowy sobowtór. To innowacyjne podejście pozwala na tworzenie wysoce spersonalizowanych treści wideo, w których użytkownicy mogą przedstawiać informacje, prowadzić szkolenia lub prezentować pomysły bez konieczności fizycznej obecności przed kamerą. W grach komputerowych wykorzystanie wielkich modeli językowych rewolucjonizuje sposób komunikowania się graczy z postaciami z gier, które to postaci, wyposażone w specjalnie trenowane modele językowe, nie są już ograniczone predefiniowanymi skryptami czy dialogami. Zamiast tego

⁶ HeyGen – AI Video Generator: <https://www.heygen.com/> [dostęp: 16.11.2023].

mogą prowadzić płynne i dynamiczne rozmowy, odpowiadając na pytania graczy, reagując na ich działania oraz uczestnicząc w tworzeniu głębokich, interaktywnych narracji. W dziedzinie muzyki niektóre modele AI, takie jak Lyria, stworzona przez Google DeepMind, już teraz umożliwiają wygenerowanie wysokiej jakości dźwięku na podstawie zapytania, które obejmuje melodię zaśpiewaną przez użytkownika wraz z tekstowym opisem tego, co użytkownik chce usłyszeć. Dzięki temu za pomocą jednego przycisku beatbox może stać się profesjonalnym nagraniem perkusyjnym, a prosta zaśpiewana melodia może zostać przekształcona w solowe nagranie saksofonu jazzowego, wykonanie orkiestry symfonicznej, lub dźwięk gitary elektrycznej⁷.

GenAI znajduje zastosowanie w automatyzacji codziennych zadań, nie tylko tych artystycznych czy związanych z ogólnie pojętym tworzeniem treści. Przykładem może być generowanie raportów, tłumaczenie języków czy pisanie e-maili. **W aspekcie rozwiązywania problemów GenAI staje się narzędziem zdolnym do generowania innowacyjnych pomysłów na nowe produkty, usługi czy strategie biznesowe, wskazując na potencjał AI w dziedzinach, gdzie tradycyjne metody dochodzenia do rozwiązań mogą być ograniczone. Przykładowe zastosowania GenAI obejmują reklamę** (tworzenie skuteczniejszych i bardziej angażujących kampanii reklamowych), **edukację** (personalizowane metody nauczania, dostosowane do indywidualnych potrzeb uczniów), **a także przemysł** (automatyzacja procesów produkcyjnych, optymalizacja wydajności i udział w tworzeniu nowych produktów).

Potencjał GenAI do transformacji wielu dziedzin życia jest ogromny, jednakże należy pamiętać, że jest to dość młoda technologia, wciąż w fazie rozwoju. GenAI wymaga więc odpowiedzialnego i etycznego stosowania, o czym więcej w następnym rozdziale. Istnieje ryzyko nadużyć i niewłaściwego wykorzystania tej technologii, co podkreśla potrzebę ustanowienia solidnych ram etycznych i prawnych, aby zarówno wykorzystywać korzyści, jak i minimalizować zagrożenia związane z tą technologią. W kontekście edukacji zrozumienie i właściwe wykorzystanie generatywnej sztucznej inteligencji przez nauczycieli i uczniów może odegrać kluczową rolę w kształtowaniu przyszłości zarówno samej edukacji, jak i społeczeństwa jako całości.

Zagadnienie kreatywności sztucznej inteligencji (AI) stanowi przedmiot intensywnych debat

⁷ Google DeepMind – Transforming the future of music creation: <https://deepmind.google/discover/blog/transforming-the-future-of-music-creation/> [dostęp: 16.11.2023].

i analiz. Tradycyjnie kreatywność, oznaczająca zdolność do oryginalnego myślenia, innowacji i twórczego rozwiązywania problemów, była postrzegana jako domena wyłącznie ludzka. **W kontekście AI zaś kreatywność jest często rozpatrywana w kategoriach zdolności do generowania nowych, wartościowych i nieoczekiwanych rozwiązań czy pomysłów.** Generatywna sztuczna inteligencja wykazuje pewne cechy kreatywności, ponieważ tworzy unikalne dzieła sztuki (w tym muzyczne) czy teksty, które mogą zaskakiwać swoją oryginalnością i głębią. **Jednakże ważne jest rozróżnienie między kreatywnością wynikającą z prawdziwej innowacyjności a kreatywnością będącą rezultatem przetwarzania i rekonfiguracji istniejących danych.** AI, choć potrafi wytwarzać treści, które wydają się nowe i oryginalne, w dużej mierze opiera się na analizie i syntezie istniejących wzorców. W związku z tym pojawia się pytanie: czy AI jest rzeczywiście kreatywna, czy jedynie naśladuje kreatywność ludzką, wykorzystując zaawansowane algorytmy?

Kontynuując formułę poprzedniej broszury⁸, nasze rozważanie zaczynamy od pytania zadanego samemu modelowi językowemu:

Czy sztuczna inteligencja może być kreatywna?

ChatGPT:

Tak, sztuczna inteligencja (AI) może wykazywać pewną formę kreatywności.

Sztuczna inteligencja jest zdolna do generowania nowych treści, rozwiązywania problemów i tworzenia czegoś, co jest uznawane za oryginalne.

Odpowiedź bota jest intrygująca. Przecież istnieje wiele opinii, że człowiek, chcąc konkurować ze sztuczną inteligencją, może to zrobić właśnie na polu kreatywności. Czy tak jest w rzeczywistości? Zobaczmy co dalej powiedział nam ChatGPT:

ChatGPT:

Oto kilka przykładów, jak AI może być kreatywna:

- **Generowanie treści artystycznej:** *Niektóre algorytmy generowania obrazów, muzyki i tekstu potrafią tworzyć dzieła sztuki, które są uznawane za kreatywne i oryginalne.*

⁸ Machura, M. (2023): *Chat GPT w szkole. Szanse i zagrożenia*, Instytut Badań Edukacyjnych, <https://www.gov.pl/web/ai/chat-gpt-w-szkole-szansy-i-zagrozenia> [dostęp: 16.11.2023].

- **Projektowanie:** AI może być wykorzystywana do projektowania różnych produktów, od ubrań po architekturę, tworząc unikalne rozwiązania.
- **Tworzenie tekstu:** Sztuczna inteligencja, w tym modele oparte na uczeniu maszynowym, może generować oryginalne teksty, artykuły, opowiadania lub poezję.
- **Optymalizacja:** AI jest wykorzystywana do optymalizacji procesów, planowania trasy w transporcie czy zarządzania zasobami w firmach, co może wymagać kreatywnego podejścia do rozwiązywania problemów.

Jednakże trzeba zaznaczyć, że **kreatywność AI jest często ograniczona do reguł i wzorców, które zostały wyuczone podczas treningu**. Sztuczna inteligencja nie posiada świadomości ani zdolności do kreatywnego myślenia w takim sensie, w jakim to robi człowiek. Wciąż istnieje wiele dyskusji i badań nad tym, w jaki sposób można polepszyć zdolności kreatywne AI oraz jak równoważyć kreatywność np. z etyką i moralnością.

Szczególnie interesująca z punktu widzenia edukacji jest uwaga, którą ChatGPT zawarł na końcu swojej „wypowiedzi”: **sztuczna inteligencja nie przetwarza informacji w taki sam sposób jak ludzie**.

1.2. AI nie przetwarza informacji w taki sam sposób jak ludzie

Duży model językowy⁹ (ang. *large language model* – LLM), taki jak np. ChatGPT, jest modelem statystycznym. Do jego wytrenowania została wykorzystana olbrzymia baza istniejących tekstów. Jednakże stwierdzenie, że tekst generowany przez modele językowe pochodzi bezpośrednio z bazy danych wykorzystanej do ich wytrenowania, jest zbytnim uproszczeniem i często powielanym mitem. W praktyce w procesie uczenia się modeli językowych architektura ucząca o nazwie *transformer*¹⁰ analizuje wielkie zbiory tekstów, aby stworzyć model, według którego różnego rodzaju kombinacje znaków zawarte w zbiorze danych zostają ułożone w olbrzymią przestrzeń połączeń (tzw. przestrzeń cech, ang. *feature space*). Ta przestrzeń określa nie tylko prawdopodobieństwo, z jakim pewne układy liter występują po innych, ale w pewnym sensie również znaczenie i kontekst tych różnych układów liter, słów, zdań i całych tekstów w ich różnych możliwych permutacjach.

Oznacza to, że taki **model może być wykorzystany do napisania całkowicie nowego**

⁹ https://en.wikipedia.org/wiki/Large_language_model [dostęp: 16.11.2023].

¹⁰ Vaswani, A., Shazeer, N.M., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I. (2017): *Attention is All you Need*. Neural Information Processing Systems.

tekstu, który nie był zawarty w bazie danych. Kiedy ktoś zadaje pytanie (pisze prompt), które nie było częścią tych wcześniejszych tekstów, model porównuje to pytanie z tym, czego się nauczył, i tworzy odpowiedź (prawdopodobną kontynuację promptu). W teorii sam wytrenowany model, jak każdy model oparty na sieciach neuronowych, jest deterministyczny: w wypadku danego wejścia i ustalonych wag modelu zawsze zwróci tę samą odpowiedź. Jednakże w praktycznych implementacjach modeli językowych, takich jak ChatGPT, często stosuje się różne techniki, takie jak próbkowanie temperatury (dostępne w bardziej zaawansowanej wersji ChatGPT dla programistów) oraz tzw. ziarna losowości (ang. *seed*), które wprowadzają element przypadku do generowanych odpowiedzi na samym poziomie architektury wytrenowanego modelu. Takim ziarnem losowości może być na przykład data i godzina zapytania użytkownika. W ChatGPT użytkownicy nie mają kontroli nad tym parametrem, dlatego w praktycznym zastosowaniu **model prawie zawsze zwróci różne odpowiedzi na to samo pytanie, mimo że został wytrenowany na teoretycznie skończonej** (choć olbrzymiej) **ilości danych.**

1.3. Dla AI nasze pytania to tylko zestaw znaków i wzorów

Ważne jest, aby pamiętać, że chociaż model językowy potrafi tworzyć prawdopodobne kontynuacje tekstów i nawet odpowiedzi na pytania, to **nie rozumie ich tak jak ludzie.**

Dla algorytmu to tylko zestaw znaków i wzorów. **Nie jest świadomy.** Korzysta z informacji, które wcześniej przyswoił. **Chociaż może się wydawać, że rozmawiamy z czymś**

świadomym, to tak nie jest. Maszyny mają zdolność uczenia się i dostosowywania do nowych informacji, ale nie rozumieją ich znaczenia w taki sam sposób jak my.

W rzeczywistości sam wytrenowany model językowy podczas rozmowy z nami nie ma nawet w pamięci tej rozmowy jako takiej. Pamięć krótkotrwała modelu jest symulowana w aplikacji, poprzez którą komunikujemy się z nim. Z każdym kolejnym zapytaniem użytkownika aplikacja tworzy podsumowanie całej dotychczasowej rozmowy. Następnie to podsumowanie w całości jest wysyłane do modelu, a ten ponownie wylicza prawdopodobieństwo wystąpienia znaków, które mogłyby stanowić kontynuację tej rozmowy.

Nowa wersja ChatGPT może posiadać dostęp do informacji nowszych niż te z czasu jego wytrenowania. W momencie wydania tej broszury model ChatGPT 4 posiada wiedzę na temat wydarzeń do kwietnia 2023 roku. Doszkalanie modeli AI może być jednak przeprowadzane przez specjalistów w ramach tzw. dotrenowania (ang. *fine-tuning*), na dodatkowej porcji danych/przykładów. Ponadto ChatGPT-4 posiada dostęp do informacji

nowszych niż te z czasu jego wytrenowania dzięki dodatkowym narzędziom zapewniającym dostęp do internetu i innych baz danych, takich jak na przykład zbiory prac naukowych (korzysta z trybu wyszukiwania w internecie oraz z tzw. pluginów). Proces pozyskiwania danych do interakcji z modelami językowymi ze źródeł zewnętrznych jest obecnie przedmiotem wielu badań w nowej dziedzinie nauki nazwanej **inżynierią podpowiedzi**, która zajmuje się tworzeniem metod efektywnej komunikacji z modelami językowymi.

1.4. Inżynieria podpowiedzi (ang. prompt engineering)

„Promptowanie” to umiejętność zadawania pytań modelom językowym w taki sposób, aby otrzymać odpowiedzi jak najbardziej zbliżone do naszych oczekiwań. Jest to dynamicznie rozwijająca się dziedzina nauki o sztucznej inteligencji oraz kompetencja coraz częściej poszukiwana przez pracodawców – „inżynier podpowiedzi” (ang. *prompt engineer*) to oficjalna nazwa stanowiska na rynku pracy, odzwierciedlająca rosnące zapotrzebowanie na specjalistów zdolnych do efektywnego komunikowania się z zaawansowanymi systemami AI. Ta nowa rola zawodowa obejmuje zarówno rozwijanie umiejętności tworzenia precyzyjnych i skutecznych promptów, jak i zrozumienie mechanizmów działania modeli językowych, co umożliwi generowanie odpowiedzi maksymalnie dopasowanych do specyficznych potrzeb i oczekiwań użytkownika.

Pierwszym przykładem efektywnego promptowania może być użycie konkretnych i precyzyjnych pytań, które prowadzą bezpośrednio do pożądanej odpowiedzi.

Na przykład zamiast pytania „Jak wygląda sztuczna inteligencja?”, dość ogólnego, lepiej zadać bardziej precyzyjne: „Jakie są główne cechy generatywnej sztucznej inteligencji w kontekście tworzenia treści?”. **Drugim przykładem jest dostosowywanie języka i tonu promptu do kontekstu i odbiorcy, co może znacząco wpływać na charakter i formę odpowiedzi.** W biznesie prompt może brzmieć bardziej formalnie i zawierać specyficzne terminy branżowe, podczas gdy w edukacji może być bardziej przystępny i zrozumiały dla młodszych użytkowników.

Obecnie na całym świecie obserwuje się wzrost zainteresowania **inżynierią podpowiedzi**. W niniejszym opracowaniu celowo zwracamy uwagę na ten trend, aby spopularyzować go wśród nauczycieli. Pragniemy, **aby zdawali sobie sprawę ze znaczenia tej nowej metody pracy z technologiami AI w szkole.**

W kontekście edukacji inżynieria podpowiedzi przydaje się w uczeniu opartym na zapytaniach (ang. *prompt-based learning*). Wiedza zdobyta przez wstępnie wytrenowane modele językowe na dużej ilości danych tekstowych jest wykorzystywana do rozwiązywania różnych typów zadań, takich jak klasyfikacja tekstu, tłumaczenie maszynowe, wykrywanie nazwanych elementów, streszczanie tekstu itp.¹¹ Pod względem metodycznym inżynieria podpowiedzi jest opisywana jako metoda „zadaj pytanie lub zestaw pytań oraz dodaj kontekst”¹². Koncepcja ta polega na optymalizacji opracowywania i dostosowywania pytań, poleceń lub sugestii wejściowych (zwanymi „promptami”) w celu uzyskania precyzyjniejszych, użyteczniejszych i zgodnych z oczekiwaniami wyników.

Mając na uwadze powyższe twierdzenia, **nauczyciele powinni uczyć zadawania pytań modelom językowym w taki sposób, aby uczeń otrzymał odpowiedzi jak najlepiej rozwiązujące problem, nad którym aktualnie pracuje się z użyciem AI.** Przyczyni się to do nabywania przez uczniów umiejętności aktualnie wymaganych w pracy z modelami językowymi.

1.5. Przykłady promptów do wykorzystania w pracy z uczniami

Poniżej przedstawiamy kilka zapytań („promptów”) opracowanych na potrzeby niniejszej broszury przez Dariusza Wronikowskiego (Akademia Leona Koźmińskiego) do wykorzystania w pracy z uczniami:

- Podaj strategię na efektywną organizację nauki i planowanie czasu podczas sesji egzaminacyjnej.
- Jakie są najlepsze metody zapamiętywania materiału przed ważnym egzaminem?
- Jak mogę efektywnie organizować moje notatki z różnych przedmiotów, aby łatwo do nich wracać? Jakie narzędzia informatyczne są dostępne na rynku?
- Jak mogę skutecznie przygotować się do prezentacji, aby wypaść profesjonalnie i pokonać treść przed wystąpieniem publicznym? Uwzględnij fakt, że jestem uczniem klasy 4 liceum ogólnokształcącego.
- Jestem uczniem szkoły średniej. Pomóż mi ułożyć plan prezentacji dotyczącej tematu: <temat pracy>.

¹¹ Na podstawie: https://pl.wikipedia.org/wiki/In%C5%BCynieria_podpowiedzi [dostęp: 16.11.2023].

¹² Radford A. (2019): *Language Models are Unsupervised Multitask Learners* [dostęp: 16.11.2023].

- Czy możesz mi podać przydatne źródła do pracy z przedmiotu w zakresie: <temat pracy>?
- Jak mogę napisać poprawnie pracę związaną z tematem: <temat pracy>?
- Jak mogę rozwijać i pogłębiać swoje umiejętności w zakresie: <zakres zainteresowań>?
- Czy mogę uzyskać przykłady dobrych wprowadzeń do mojej pracy: <temat pracy>?
- Napisz post dotyczący produktu, uwzględniając fakt, że tworzysz treść dla osób, np. uczniów szkół średnich.
- Napisz post na Instagram/Facebook/LinkedIn promujący konkretny produkt, markę lub usługę. Uwzględnij fakt, że jestem menedżerem marki i odpowiadam za ten produkt.

Rozmawiając z modelem językowym, za każdym razem należy **dokładnie opisywać kontekst**, który nas interesuje. Należy przy tym wyraźnie i szczegółowo **podawać modelowi nasze oczekiwania**, następnie sprawdzić, czy wygenerowane informacje są prawdziwe, jeśli jest to kluczowe dla naszego zapytania (o tym więcej w podpunkcie 1.8). Zachęcamy do przetestowania powyższych promptów w pracy z uczniami, aby mogli nauczyć się sztuki konwersacji z AI. Niech nauczyciele pomagają im przygotować prompty początkowe, a następnie kontynuować wymianę informacji z botem zgodnie z metodologią tworzenia promptów.

1.6. Czy sztuczna inteligencja jest kreatywna

Poszczególne modele sztucznej inteligencji potrafią generować różnego rodzaju dane. Te ciągi tekstu, grafiki, nut itp. my, ludzie interpretujemy jako kreatywne wytwory sztucznej inteligencji. **Generatywna sztuczna inteligencja może produkować sztukę, muzykę, literaturę i inne wytwory, które są nie do odróżnienia od dzieł stworzonych przez ludzi.** AI to jednak wciąż algorytmy, które funkcjonują w oparciu o modele wykształcone na podstawie danych. Te dane powstały w wyniku działań ludzi, toteż model symuluje działanie ludzkie i jego wytwory **mogą sprawiać wrażenie, jakby były wytworami ludzkimi, mimo że bezpośrednio nimi nie są.** AI może badać i analizować ogromne ilości danych, aby znaleźć nowe rozwiązania problemów, co można postrzegać jako pewnego rodzaju formę „kreatywności”. AI może być też partnerem współpracującym z artystami, stymulując

ich proces twórczy poprzez sugerowanie pomysłów, doskonalenie koncepcji i przyspieszanie pracy nad dziełem. Obrazy generowane przez AI są już sprzedawane na aukcjach, a muzyka skomponowana przez AI jest wykorzystywana komercyjnie. **Są to wybrane argumenty za tym, że sztuczna inteligencja może być kreatywna.**

1.7. Tezy przeczące kreatywności AI

Istnieją również opinie przeciwne tezie o kreatywności maszyn, oparte na założeniu, że **kreatywność jest nieodłącznie związana ze świadomością, emocjami i subiektywnymi doświadczeniami**. Ponieważ sztuczna inteligencja ich nie ma, nie może być, zdaniem zwolenników tej tezy, naprawdę kreatywna – ona tylko naśladuje wzory, na których została przeszkolona. Remiksuje je i łączy istniejące pomysły. Jednak prawdziwa kreatywność, jak twierdzą przeciwnicy tezy o kreatywności maszyn, wymaga generowania naprawdę nowych pomysłów, a nie tylko permutacji rozwiązań już istniejących. Jeżeli model AI nauczył się, że pewne sekwencje znaków występują w języku w różnych sytuacjach z większym prawdopodobieństwem niż inne, to liczba kombinacji, które model może wygenerować, jest – choć olbrzymia – teoretycznie wciąż skończona. Tak więc **możliwości interpretacyjne modelu w określonej sytuacji, tak samo jak liczba permutacji wyuczonych cech, są teoretycznie ograniczone**. Choć model jest w stanie przekazać nam tę samą informację na nieskończoną liczbę sposobów, to liczba pojęć, na których został wytrenowany, jest konkretna i skończona. Tymczasem według Noama Chomskiego ludzie podejmują decyzje inaczej. Zauważa on (tłumaczenie: ChatGPT):

Ludzki umysł nie jest, jak ChatGPT i podobne mu, ociężałym statystycznym silnikiem do dopasowywania wzorców, pochłaniającym setki terabajtów danych i ekstrapolującym najbardziej prawdopodobną odpowiedź w rozmowie czy najbardziej prawdopodobne rozwiązanie naukowego problemu. Wręcz przeciwnie, ludzki umysł to zadziwiająco wydajny, a nawet elegancki system, który działa na niewielkich ilościach informacji; nie szuka on brutalnych korelacji między punktami danych, ale stara się tworzyć wyjaśnienia¹³.

Podczas gdy ludzie często tworzą intencjonalnie i pod wpływem emocji, AI nie towarzyszą wówczas motywacje osobiste. Za pomocą algorytmów wytwarza ona wyniki

¹³ Chomsky, N.: *The False Promise of ChatGPT*, The New York Times, <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html> [8.03.2023, dostęp: 16.11.2023].

oparte na instrukcjach, bez jakiegokolwiek „intencji” tworzenia. Kontekst, w którym powstaje dzieło sztuki lub inny wytwór kreatywności, jest kluczowy, aby to dzieło zinterpretować i docenić. **AI nie może naprawdę zrozumieć kontekstu kulturowego, historycznego czy emocjonalnego tak jak ludzie**, choć może to zrozumienie symulować, wykorzystując wielkie zbiory danych. Rozwój sztucznej inteligencji spowodował, że tzw. sprawczość (ang. *agency*) jest obecnie przedmiotem pogłębionych debat. Jak zauważył Martin Zeilinger, naukowiec zajmujący się sztuką cyfrową i sztuczną inteligencją:

Sprawczość [ang. agency] jest najogólniej rozumiana jako przejaw zdolności do działania. Ale koncepcja ta jest również ściśle powiązana z kwestiami samostanowienia, autonomii, wolności słowa i możliwości posiadania własności. Jak zatem konstytuuje się sprawczość w dobie sztucznej inteligencji? Kto lub co jest teraz [sprawczym] agentem?¹⁴

W związku z tym sporem zadajemy kolejne pytanie modelowi językowemu ChatGPT:

Czy w powyższym sporze można znaleźć złoty środek?

ChatGPT:

*Niektórzy myśliciele proponują złoty środek, sugerując, że **AI ma formę kreatywności, ale różni się od ludzkiej kreatywności**. Chodzi bardziej o rozpoznawanie wzorów, optymalizację i eksplorację w ramach wcześniej określonych granic. **Natomiast ludzka kreatywność jest związana z emocją, świadomością i kontekstem kulturowym**.*

Bez względu na stanowisko w tej debacie, niezaprzeczalne jest, że rola AI w dziedzinach twórczych rośnie, a jego wyniki stają się bardziej zaawansowane. Współpraca między AI a artystami może w przyszłości przededefiniować nasze rozumienie kreatywności.

Wiemy już, że **generatywna sztuczna inteligencja** jest w stanie symulować ludzkie procesy kreatywne dzięki uczeniu się na dużej liczbie przykładów. Warto więc zapytać, **do czego nie służy sztuczna inteligencja** w kontekście kreatywności.

¹⁴ Zeilinger, M. (2021): *Tactical entanglements: AI art, creative agency, and the limits of intellectual property*, <https://doi.org/10.14619/1839>, s. 15 (tłumaczenie: Tłumacz Google).

1.8. Do czego nie służy AI w kontekście kreatywności

Z pewnością nie ma potrzeby korzystania z AI, gdy znasz rozwiązanie, tj. wiesz, jak postąpić algorytmicznie. Nie używaj AI, jeśli wiesz, jak postępować z danymi, aby za każdym razem osiągnąć taki sam cel. Spróbujmy to wyjaśnić na przykładzie. Skoro istnieje kalkulator, który algorytmicznie zawsze dokładnie doda 2 do 2 i poda wynik: 4, to nie ma sensu pytać o to ChatGPT, który według rachunku prawdopodobieństwa może czasami podać wynik: 5. Piątka jest prawdopodobnym wynikiem takiego zapytania, mimo że nie jest wynikiem prawdziwym. **ChatGPT nie jest jednak narzędziem, które specjalizuje się w wynikach prawdziwych. To narzędzie, które specjalizuje się w wynikach prawdopodobnych, a w kontekście prawdopodobnej rozmowy ktoś przecież mógłby na ww. pytanie odpowiedzieć błędnie.** Jeśli więc chcesz coś zautomatyzować i wiesz, jak to zrobić, użyj gotowego algorytmu. Jeśli zaś coś ma wymiar ludzki, tzn. potrafimy to zrobić, ale nie umiemy tego zautomatyzować za pomocą algorytmu, wtedy z pomocą przychodzi nam AI, czyli symulacja ludzkiego zachowania. AI nie istnieje po to, aby zautomatyzować zadania, które wiemy, jak zautomatyzować. **AI możemy wykorzystać do zautomatyzowania zadania, które nie wiemy, jak zautomatyzować.** John von Neumann¹⁵ stwierdził: „Upierasz się, że jest coś, czego maszyna nie może zrobić. Jeśli powiesz mi dokładnie, czego maszyna nie potrafi, zawsze będę w stanie stworzyć maszynę, która właśnie to zrobi”. Neumann powiedział, że jeżeli wiesz dokładnie, z czego składa się Twoje zadanie, to wiesz też, jak je zautomatyzować maszynowo. Wtedy nie potrzebujesz AI.

Sztuczna inteligencja zrewolucjonizowała wiele dziedzin życia, pozwalając maszynom symulować ludzkie zachowania i procesy myślowe w sposób, który jeszcze kilka lat temu byłby uznany za niemożliwy. Jednakże mimo tych niewątpliwych postępów ważne jest zrozumienie, że AI nie jest wszechmogąca i istnieją obszary, w których jej wykorzystanie może być nie tyle niepotrzebne, ile wręcz niepożądane.

¹⁵ https://pl.wikipedia.org/wiki/John_von_Neumann [dostęp: 16.11.2023].

1.9. Wybrane sytuacje, w których nie należy korzystać z modeli językowych (LLM)

Poniżej przedstawiamy kilka sytuacji, w których zdecydowanie nie należy korzystać z modeli językowych lub podobnych narzędzi opartych na sztucznej inteligencji:

Leczenie i diagnoza. LLM nigdy nie powinien być wykorzystywany do uzyskiwania informacji na temat chorób i leczenia. Nie powinien być używany do diagnozowania lub zastępowania profesjonalnej opinii medycznej. Łatwo jest wyobrazić sobie niefortunną sytuację, w której model językowy, zapytany o dawkowanie leku, popełnia błąd, który może skończyć się dla człowieka uszczerbkiem na zdrowiu lub nawet śmiercią.

Badania naukowe. Chociaż LLM może być użytecznym narzędziem do uzyskiwania informacji ogólnych, to badacze powinni polegać na źródłach pierwotnych i dokładnych metodach badawczych. LLM jest dobrym generatorem tekstu, ale tekst ten nie musi być prawdziwy.

Porady prawne. Chociaż LLM może dostarczać ogólnych informacji na temat prawa, to nie zastępuje prawnika i nie weźmie na siebie odpowiedzialności za wydane opinie prawne.

Tłumaczenia, w których duże znaczenie ma dokładność. Chociaż LLM jest zdolny do przekładania tekstu na różne języki, nie zastąpi profesjonalnych tłumaczy, zwłaszcza w przypadku tekstów skomplikowanych lub specjalistycznych. Błędne tłumaczenia, na przykład w polityce, mogłyby doprowadzić do skandali dyplomatycznych.

Informacje poufne. Nigdy nie należy udostępniać sztucznej inteligencji danych poufnych, takich jak numery kart kredytowych, hasła itp. Wprawdzie OpenAI deklaruje, że w przypadku dostępu do ChatGPT poprzez interfejs programistyczny API, nie przechowuje ani nie analizuje wpisów użytkowników, ale lepiej być ostrożnym.

Decyzje życiowe. Chociaż LLM może dostarczyć ogólnych informacji na różne tematy, to nie powinno się polegać na nim, podejmując ważne decyzje życiowe: medyczne, finansowe czy prawne.

Autentyczność emocji. W sytuacjach wymagających empatii, zrozumienia emocji i budowania głębokich relacji AI nie jest w stanie zastąpić człowieka i jego reakcji.

Kwestie etyczne i moralne. W ważnych kwestiach etycznych i moralnych lepiej polegać na ludzkiej intuicji bądź na opiniach ekspertów lub doradców niż na odpowiedziach generowanych przez AI.

Sytuacje awaryjne. W sytuacjach awaryjnych, takich jak katastrofy naturalne czy zagrożenia medyczne, lepiej polegać na oświadczeniach ekspertów lub służb ratunkowych.

Dane personalne. Nie powinno się zadawać pytań związanych z konkretnymi osobami. Choć LLM może dysponować informacjami na temat ważnych postaci historycznych, to jest wielce prawdopodobne, że jego odpowiedź okaże się nieprawdziwa.

Dane osobowe w szkole. W szkołach przechowywane są różnorodne dane osobowe, w tym tzw. dane osobowe szczególnych kategorii (dokumenty z poradni psychologiczno-pedagogicznych, np. orzeczenia o potrzebie kształcenia specjalnego czy nauczania indywidualnego). W szkole gromadzi się też protokoły rad pedagogicznych, zawierające informacje, które jeśli zostaną ujawnione, mogą negatywnie wpłynąć zarówno na życie uczniów, jak i nauczycieli. **Dlatego w przypadku korzystania z rozwiązań ogólnodostępnych, takich jak ChatGPT czy Bard, nie powinno się w rozmowie z AI przekazywać żadnych danych osobowych ani informacji objętych tajemnicą rady pedagogicznej. Pod żadnym pozorem nie wolno prosić otwartego modelu językowego (mowa o LLM w wersji ogólnodostępnej) o wygenerowanie na przykład indywidualnego programu edukacyjno-terapeutycznego na podstawie udostępnianych mu dokumentów¹⁶.**

Historia. LLM, zapytany o wydarzenia historyczne, może przekazać nam informacje nieprawdziwe, niezweryfikowane, a nawet stworzyć całkowicie zmyśloną opowieść, którą przedstawi jako prawdę historyczną.

Podsumowując, **generatywna sztuczna inteligencja** jest w stanie symulować ludzkie procesy kreatywne dzięki uczeniu się na dużej liczbie przykładów. **Nie powinna być jednak wykorzystywana w sytuacjach, w których duże znaczenie ma prawdziwość**

¹⁶ Więcej przykładów działań niepożądanych zawarliśmy w rozdziale 4, zatytułowanym *Do czego nie służy AI – szkolne przykłady działań niepożądanych*.

uzyskanych informacji. We współczesnym świecie, zdominowanym przez technologię, sztuczna inteligencja odgrywa coraz większą rolę, a co za tym idzie – kuszące jest korzystanie z narzędzi AI w każdej możliwej sytuacji. Jednak trzeba pamiętać, że mają one swoje ograniczenia. **Choć ChatGPT, Bard i podobne systemy okazały się rewolucyjne i oferują szeroką gamę zastosowań, to nie są pozbawione wad.** Warto, by użytkownicy AI mieli tego świadomość.

II. Etyka i moralność w kontekście sztucznej inteligencji

dr Tomasz Łukawski

2.1. Ogólnie o etyce w kontekście AI

Obecnie wpływ AI na uczniów i nauczycieli jest niewątpliwie znaczący. Trendy rozwoju narzędzi opartych na technologiach sztucznej inteligencji wskazują iż świat zmienia się na naszych oczach diametralnie. Jednym z głównych wyzwań stojących przed edukacją jest przygotowanie uczniów do życia w dorosłym świecie. Nie ma wątpliwości, że będzie on zdominowany przez wszechobecne systemy oparte na AI. W kręgach Komisji Europejskiej, obok nawoływań do monitorowania rozwoju systemów AI, coraz częściej słychać **apele o przyjęcie regulacji, które dotyczyć będą z jednej strony etycznego projektowania AI, a z drugiej etycznego korzystania z jej narzędzi**. Na łamach niniejszej broszury kierujemy ten apel do użytkowników będących członkami ekosystemu edukacji, czyli nauczycieli oraz zastanawiamy się, **co można zrobić, aby korzystanie z AI oraz projektowanie narzędzi opartych na AI było etyczne i odpowiedzialne**.

Jak korzystać z AI etycznie i odpowiedzialnie

Etyka odgrywa istotną rolę w różnych dziedzinach życia, pomagając ludziom dokonywać wyborów moralnych, rozstrzygać konflikty wartości i rozwijać zasady współżycia społecznego. To dziedzina filozofii zajmująca się badaniem moralności, wartości i norm postępowania. Jej celem jest analiza i opracowywanie zasad moralnych, które kierują ludzkim postępowaniem. Etyka stara się odpowiedzieć na pytania dotyczące tego, co jest moralnie dobre, złe czy sprawiedliwe. W jej ramach można wyróżnić rozmaite podejścia i teorie:

- **etyka normatywna** – zajmuje się ustalaniem norm moralnych, czyli określaniem, co powinniśmy robić, a czego unikać;
- **etyka opisowa** – bada faktyczne zachowania moralne ludzi, analizując rzeczywiste normy i wartości obowiązujące w społeczeństwie;
- **metaetyka** – skupia się na naturze samej etyki, szukając odpowiedzi na pytanie, co oznacza bycie dobrym lub złym oraz jakie są podstawy moralności;
- **etyka stosowana** – zajmuje się stosowaniem ogólnych zasad etycznych

w konkretnych sytuacjach (etyka medyczna, biznesowa czy środowiskowa).

Etyka sztucznej inteligencji to dziedzina, która zajmuje się badaniem moralnych aspektów związanych z rozwojem, implementacją i użytkowaniem systemów sztucznej inteligencji. W miarę postępu technologii, zwłaszcza w obszarze sztucznej inteligencji, pojawiają się różnorodne kwestie etyczne, które dotyczą zarówno projektantów, programistów, jak i społeczeństwa jako całości. Grupa ekspertów wysokiego szczebla Unii Europejskiej ds. sztucznej inteligencji opublikowała pierwszy projekt wytycznych w zakresie etyki AI w grudniu 2018 roku. Po konsultacjach z zainteresowanymi i spotkaniach z przedstawicielami państw członkowskich, wymieniona grupa przedstawiła Komisji zmieniony dokument w marcu 2019 roku¹⁷. Znajdują się w nim konkretne wskazówki etyczne dla twórców, dostawców i użytkowników sztucznej inteligencji.

Etyczne ramy sztucznej inteligencji wg Komisji Europejskiej

W przywołanym dokumencie znajdziemy między innymi zalecenie, by systemy AI były opracowywane, wdrażane i wykorzystywane zgodnie z takimi zasadami etycznymi, jak: **poszanowanie autonomii człowieka, zapobieganie szkodom, sprawiedliwość i możliwość wyjaśnienia**. Trzeba sobie jednak zdawać sprawę z możliwości wystąpienia konfliktów między tymi zasadami i podejmować w tym zakresie stosowne działania. Przenosząc te rozważania na grunt szkolny, można stwierdzić, iż nauczyciele, zwłaszcza podczas pracy z AI, powinni zwracać szczególną uwagę na poszanowanie autonomii uczniów oraz na zagadnienia etyczno-moralne w sytuacjach wywierających wpływ na dzieci. Mówiąc wprost: **nie należy zostawiać dziecka sam na sam z technologią AI**, gdyż nie można przewidzieć, jaką odpowiedź otrzyma ono po zadaniu modelowi językowemu konkretnego pytania. Znane są przypadki, w których sztuczna inteligencja prezentowała w odpowiedzi wyuczone schematy, które są sprzeczne z ogólnie przyjętymi kanonami moralnymi i etycznymi. **Obecność nauczyciela jest zatem niezbędną**, aby na bieżąco reagował i prostował te odpowiedzi AI, które są etycznie wątpliwe.

Systemy AI przynoszą poszczególnym osobom oraz społeczeństwu znaczne korzyści, ale **korzystanie z nich może się również wiązać z określonym ryzykiem i wywoływać**

¹⁷ <https://digital-strategy.ec.europa.eu/en/library/communication-building-trust-human-centric-artificial-intelligence> [dostęp: 16.11.2023].

negatywne skutki, w tym skutki, które mogą okazać się trudne do przewidzenia, zidentyfikowania lub zmierzenia (...) ¹⁸.

Powyższe przestrogi nie oznaczają, aby nie korzystać z AI. Wręcz przeciwnie: **współczesne zdobycze sztucznej inteligencji to dla nauczyciela wyjątkowa okazja**, aby wykorzystał AI jako narzędzie wspierające **rozwijanie wśród uczniów umiejętności krytycznego myślenia**. Zalecamy jedynie, aby proces ten odbywał się **pod kontrolą świadomego pedagoga**.

Kiedy sztuczna inteligencja jest godna zaufania

Najnowsze modele językowe w wersji komercyjnej (płatna subskrypcja) umożliwiają ¹⁹ tworzenie własnych chatbotów GPT (piszemy o tym szerzej w rozdziale trzecim).

Nie możemy więc wykluczyć, że część nauczycieli i uczniów już teraz z tej opcji korzysta.

Wówczas osoby te **stają się nie tylko użytkownikami systemów AI, ale i projektantami własnych chatbotów**, opartych na personalnych i specjalistycznych danych i komunikujących się z innymi. Dlatego też wytyczne dotyczące godnej zaufania sztucznej inteligencji, zawarte w przytoczonym dokumencie UE, **odnoszą się również do tej części uczniów i nauczycieli jako twórców systemów**.

W wytycznych sformułowano siedem wymogów:

- przewodnia i nadzorczą rolę człowieka,
- solidność techniczna i bezpieczeństwo,
- ochrona prywatności i zarządzanie danymi,
- przejrzystość,
- różnorodność, niedyskryminacja i sprawiedliwość,
- dobrostan społeczny i środowiskowy,
- odpowiedzialność.

Tym samym podstawowym celem regulacji ²⁰ staje się **wdrożenie godnej zaufania sztucznej inteligencji**. Pojedynczy uczniowie i nauczyciele oraz całe społeczności szkolne będą mogli zaufać technologii tylko wtedy, gdy ustanowione zostaną przejrzyste

¹⁸ Tamże.

¹⁹ W wypadku ChatGPT-4 możliwość ta istnieje od listopada 2023 roku.

²⁰ Niezależna grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji, *Wytyczne w zakresie etyki dotyczące godnej zaufania sztucznej inteligencji*, s. 3-8.

i kompleksowe ramy zapewniające możliwość praktycznego zagwarantowania, że technologia ta jest godna zaufania²¹.

2.2. Zdolności maszyn a zdolności ludzi w kontekście etyki i moralności

W kontekście naszych rozważań o etyce i AI dochodzimy do pytania: **czy zdolności maszyn są w stanie przekroczyć zdolności ludzkie?**

Jeżeli zapytasz model językowy o zasady etyki czy moralności, uzyskasz bardzo szczegółową odpowiedź, która prawdopodobnie doprowadzi Cię do wniosku, że model posiada znaczną wiedzę w tej dziedzinie. Jednakże po głębszej dyskusji zdecydowanie stwierdzisz, że pomimo znaczącej wiedzy **AI nie potrafi stosować tych zasad w konkretnej sytuacji.**

Mając na względzie wcześniejsze spostrzeżenie, że jest to dla maszyny wyłącznie ciąg znaków wygenerowany na podstawie modeli, według których te znaki powiązała, **nie jest zasadne pytać LLM o kwestie dotyczące etyki, moralności czy religii.** Należy mieć bowiem świadomość, że maszyna nie posiada żadnej moralności, żadnego światopoglądu ani żadnego doświadczenia w stosowaniu zasad etycznych.

AI posiada wiedzę o etyce i moralności, ale czy potrafi ją zastosować?

Posiadanie wiedzy (AI posiada jakąś część wyuczonej wiedzy) **do jej stosowania** ma się tak, jak religioznawstwo do religijności, relacje służbowe do przyjaźni, wiedza o świecie do ciekawości poznawczej, oczekiwanie na konkretną godzinę do tęsknoty czy w końcu zdjęcie osoby do miłości do drugiego człowieka. Przenosząc tę analogię na grunt szkolny, to tak, jakby wystawić nauczycielowi ocenę pracy na papierze a zachwycać się jego osiągnięciami przed zgromadzonym audytorium pedagogicznym.

Ludzie często podejmują decyzje intuicyjnie, na podstawie swoich upodobań, które nie muszą być ani racjonalne, ani pragmatyczne, i w oparciu o pamięć sensoryczną. Znamy takie przykłady z własnego życia. Coraz częściej więc w kontekście projektowania

²¹ Obecnie projektowane rozporządzenie Parlamentu Europejskiego i Rady ustanawiające zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji) i zmieniające niektóre akty ustawodawcze Unii przewiduje wprowadzenie swoistych ram certyfikacji dla określonych rodzajów systemów sztucznej inteligencji. Jego projekt powstał na podstawie ww. wskazówek dotyczących godnej zaufania AI.

narzędzi czy modeli AI mówi się o tzw. **teorii ucieleśnionego poznania**²². Przykładem może być nasze emocjonalne nastawienie do melodii, którą śpiewał nam rodzic, gdy w pokoju unosił się zapach lawendy. Gdy po kilkudziesięciu latach słyszymy podobną melodię, intuicyjnie ją lubimy, choć do końca nie wiemy, dlaczego i nie potrafimy tego racjonalnie wytłumaczyć. Tymczasem **modele sztucznej inteligencji nie są w stanie analizować danych związanych ze źródłem uczuciowości ani intuicyjnie i empatycznie odnieść się do wrażliwości drugiego człowieka tak samo jak inny człowiek, który potrafi współczuć przez wzgląd na swoje własne ucieleśnione wspomnienia.**

Jeżeli AI nie posiada wiedzy, to ją „wymyśli”

Inną kwestią jest to, że model sztucznej inteligencji, jeżeli nie zna odpowiedzi na zadane pytanie, **zaczyna wymyślać rzeczywistość**. Zjawisko to nazywane jest „halucynowaniem”. Na domiar złego, gdy złapiemy go na tym „kłamstwie”, uprzejmie przeprosi, po czym... uraczy nas kolejnym „kłamstwem”. Przykład takiego postępowania zaprezentowano w broszurze *Chat GPT w szkole. Szanse i zagrożenia*.

A skoro już wspomnieliśmy o poprzedniej publikacji, to pora wskazać, co zmieniło się w dziedzinie sztucznej inteligencji od czasu publikacji *Chat GPT w szkole. Szanse i zagrożenia*. Powiemy o tym w następnym rozdziale.

²² Poznanie ucieleśnione (z ang. *embodied cognition*) – określenie stosowane w filozofii poznania, psychologii poznawczej, kognitywistyce i neuronauce na taki typ poznania, na którego wpływ mają nie tylko procesy przetwarzania informacji w mózgu, ale i w całym ciele. Poprzez to teoria ta stoi w opozycji do zakładanego w kognitywistyce symbolicznej poglądu o obliczeniowej teorii umysłu (komputacjonizmu), mówiącego, że mózg jako operujący symbolicznymi obliczeniami organ nadrzędnie przyczynia się do poznania świata. Podejście to stoi także w opozycji do dualizmu kartezjańskiego. Zakłada, że nasza wiedza o świecie jest nierozzerwalnie związana z naszymi ciałami, językiem i społecznym usytuowaniem, czyli ucieleśnieniem (Francisco Varela, *The Embodied Mind: Cognitive Science and Human Experience*, 1991). Zob. https://pl.wikipedia.org/wiki/Poznanie_uciele%C5%9Bnione [dostęp: 16.11.2023].

III. Nowe narzędzia i funkcje modeli językowych

dr Tomasz Łukawski

3.1. Nowości w ChatGPT przed konferencją OpenAI DevDay 2023 i po niej

Od dnia publikacji poprzedniej broszury minęło ponad pół roku. W tym okresie udostępniona została rozszerzona (płatna) wersja GPT-4. Nauczyciele jednak powszechnie korzystają z bezpłatnej wersji konsumenckiej. Powinni przy tym pamiętać, że **nie wolno używać podstawowych wersji modeli językowych, jeśli zamierza się pracować na danych osobowych, zwłaszcza tych należących do kategorii wymienionych w art. 9 RODO²³ (np. orzeczenia, opinie poradni pedagogiczno-psychologicznych, zalecenia itp.), lub przetwarzać informacje, którymi nie wolno dzielić się z osobami trzecimi (np. notatki z posiedzeń rady pedagogicznej)**. Należy mieć na uwadze, że ChatGPT, podobnie jak inne chatboty, może „uczyć się” na podstawie zadawanych mu pytań i wprowadzanych do niego informacji. Co więcej: kierowane do niego pytanie będzie widoczne dla dostawcy ChatGPT, czyli firmy OpenAI. **Dlatego najpierw należy zapoznać się z warunkami użytkowania narzędzia oraz polityką prywatności jego twórcy.**

Najnowsza wersja GPT-4 pozwala na wyszukiwanie informacji w internecie oraz generowanie obrazów. Co ciekawe: możliwa jest też konwersacja o obrazach (można przykładowo zapytać, co się znajduje na grafice i do czego służy). **Wersja mobilna aplikacji pozwala ponadto na rozmowę głosową z ChatGPT.**

Aplikacja zawiera **zestaw pluginów** (tzw. wtyczek), które GPT-4 może wykorzystywać do pozyskiwania specjalistycznych informacji, rozmowy na temat plików o rozszerzeniu .pdf, przeszukiwania znanych baz danych czy prac naukowych.

Aplikacja wyposażona jest również w **funkcję Advanced Data Analysis**, która pozwala na edytowanie arkuszy kalkulacyjnych oraz analizowanie i tworzenie kodu komputerowego. Dzięki niej ChatGPT potrafi liczyć i rozwiązywać zadania matematyczne z większą precyzją

²³ Chodzi o dane osobowe ujawniające pochodzenie rasowe lub etniczne, poglądy polityczne, przekonania religijne lub światopoglądowe, przynależność do związków zawodowych oraz dane genetyczne, dane biometryczne w celu jednoznacznego zidentyfikowania osoby fizycznej oraz dane dotyczące zdrowia, seksualności lub orientacji seksualnej (*Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych)*, Dziennik Urzędowy Unii Europejskiej L 119/1).

niż dotychczas. Korzystając z tej funkcji, można przesłać chatbotowi dowolny plik i poprosić o jego analizę.

Z uwagi na ograniczoną objętość niniejszej broszury zainteresowanych tym, jak używać dużych modeli językowych w szkole, **odsyłamy do polskiej wersji oficjalnego dokumentu *Nauczanie z wykorzystaniem sztucznej inteligencji*²⁴**, znajdującego się na stronie producenta Open AI. Jest to przewodnik dla nauczycieli korzystających z ChatGPT, zawierający sugerowane odpowiedzi oraz informacje, jak działa model językowy i jakie są jego ograniczenia. Twórcy aplikacji dzielą się w nim kilkoma przykładami wykorzystywania ChatGPT na lekcjach w szkole oraz odpowiedziami, które pomogą nauczycielom rozpocząć korzystanie z tego narzędzia. Zainteresowani znajdą tam również odpowiedzi na najczęściej zadawane pytania.

Oto kilka przykładów wprost ze strony OpenAI²⁵:

- Doktor Helen Crompton, profesor Uniwersytetu Old Dominion, zachęca swoich studentów, aby używali ChatGPT jako partnera w dyskusji, który wskaże luki w ich argumentacji, jako rekrutera przeprowadzającego z nimi rozmowę rekrutacyjną lub jako nowego szefa, który może przekazać im informację zwrotną w określony sposób.
- Z kolei Fran Bellas, profesor Universidade da Coruña w Hiszpanii, zaleca nauczycielom używanie ChatGPT jako asystenta pomagającego w tworzeniu quizów, egzaminów i scenariuszy zajęć. Sugeruje, aby najpierw udostępnić treść programu nauczania ChatGPT, a następnie poprosić go o pomysły na quizy i plany lekcji. Bellas twierdzi: „Jeśli poprosisz ChatGPT o stworzenie egzaminów składających się z 5 pytań na temat obwodów elektrycznych, otrzymasz nowe pomysły, które możesz uczynić własnymi”.
- Dr Anthony Kaziboni, kierownik działu badań na Uniwersytecie w Johannesburgu, uczy studentów, którzy poza zajęciami w większości nie mówią po angielsku. Kaziboni wierzy, że znajomość języka angielskiego jest ogromną zaletą i że nierozumienie subtelności angielskiej gramatyki może utrudniać studentom rozwój. Zachęca ich do wykorzystywania ChatGPT jako narzędzia pomagającego w tłumaczeniu, doskonaleniu pisania w języku angielskim i ćwiczeniu konwersacji.

²⁴ <https://openai.com/blog/teaching-with-ai> [dostęp: 16.11.2023].

²⁵ <https://openai.com/blog/teaching-with-ai> [dostęp: 16.11.2023].

- Geetha Venugopal, nauczycielka informatyki w American International School w Ćennaju w Indiach, porównuje nauczanie uczniów o narzędziach sztucznej inteligencji do nauczania ich tego, jak odpowiedzialnie korzystać z internetu. Na zajęciach radzi im, aby pamiętali, że odpowiedzi udzielane przez ChatGPT nie zawsze muszą być wiarygodne i dokładne, oraz aby zastanowili się, czy w danej sytuacji powinni zaufać botowi, a następnie potwierdzili uzyskane informacje w innych źródłach. Celem jest pomóc im „zrozumieć znaczenie ciągłej pracy nad oryginalnym krytycznym myśleniem, umiejętnościami rozwiązywania problemów i kreatywnością”.

Nowości na OpenAI DevDay 2023

W dynamicznie rozwijającym się świecie technologii sztuczna inteligencja coraz śmieiej wkracza do przestrzeni edukacyjnej, oferując nie tylko wsparcie dydaktyczne, ale i otwierając nowe ścieżki kreatywnego rozwoju. Podczas wydarzenia DevDay, które odbyło się 7 listopada 2023 roku w San Francisco z inicjatywy OpenAI, zespół projektowy ChatGPT przedstawił podsumowanie dotychczasowych osiągnięć oraz zaprezentował nowe funkcje modelu. Ostatnie ogłoszenia OpenAI w kontekście modelu ChatGPT wskazują, że przed nauczycielami i uczniami otwiera się szereg możliwości, które mogą znacząco wzbogacić proces nauczania i uczenia się. Oceniając potencjał nowych narzędzi, warto zastanowić się nad ich wpływem na przyszłość edukacji i nad wyzwaniem, przed którym staje człowiek korzystający z tych innowacji. Dlatego też w tej sekcji zajmiemy się omówieniem najnowszych rozwiązań zaproponowanych przez OpenAI i ich znaczeniem dla szkoły jutra.

GPTs – indywidualnie dostosowane modele ChatGPT: możliwość tworzenia przez użytkowników spersonalizowanych wersji ChatGPT

Nowa funkcja GPTs pozwala użytkownikom na tworzenie indywidualnie dostosowanych wersji ChatGPT bez konieczności posiadania zaawansowanej wiedzy programistycznej. To innowacja, która demokratyzuje dostęp do technologii AI, umożliwiając personalizację modelu językowego z uwzględnieniem potrzeb konkretnego użytkownika. Proces ten wykorzystuje intuicyjne interfejsy, pozwalając na wprowadzenie i dostosowanie danych, które ma wykorzystać model. W praktyce oznacza to, że każdy może „nauczyć” ChatGPT języka branżowego bądź przekazać mu specyficzne informacje, co zaowocuje stworzeniem

optymalnego modelu nadającego się do konkretnych zastosowań: od edukacji po obsługę klienta.

W kontekście edukacyjnym GPTs otwiera drzwi do stworzenia niemal nieograniczonej liczby specjalistycznych narzędzi wspomagających proces nauczania. Nauczyciele mogą skonfigurować ChatGPT tak, aby odpowiadał na potrzeby edukacyjne ich uczniów, np. poprzez włączenie materiałów dydaktycznych z danego przedmiotu lub dostosowanie dialogu do wieku uczniów. Personalizowane modele mogą również wesprzeć uczniów z indywidualnymi potrzebami edukacyjnymi, takimi jak nauka drugiego języka lub zapewnienie pomocy w zrozumieniu złożonych koncepcji naukowych.

Być może wprowadzenie GPTs w edukacji będzie krokiem milowym na drodze do stworzenia bardziej elastycznych i dostosowanych do użytkownika środowisk nauczania. Może to znacząco zmienić rolę nauczyciela, który stanie się nie tylko przewodnikiem po świecie wiedzy, ale również projektantem i kreatorem narzędzi edukacyjnych. Z kolei uczniowie zyskają dostęp do bardziej spersonalizowanego nauczania, co może przyczynić się do ich większego zaangażowania w proces uczenia się i lepszego zrozumienia przez nich materiału. Personalizacja edukacji dzięki GPTs może przynieść rewolucję w sposobie, w jaki młode pokolenia będą przyswajając wiedzę, czyniąc proces nauczania bardziej interaktywnym, indywidualizowanym i dostosowanym do potrzeb każdego ucznia.

GPT Store – platforma dystrybucji GPTs: platforma umożliwiająca udostępnianie i komercjalizację własnych wersji ChatGPT

GPT Store to nowatorska platforma OpenAI, która umożliwia udostępnianie oraz komercjalizację spersonalizowanych modeli ChatGPT stworzonych przez użytkowników. Działa podobnie do aplikacji sklepowych, w których indywidualnie stworzone modele mogą być przeglądane, pobierane, a nawet kupowane przez innych. Użytkownicy mają możliwość dzielenia się swoimi własnymi wersjami ChatGPT, które zostały dostosowane do specyficznych potrzeb, tematów lub branż. To otwiera zupełnie nowy wymiar w dystrybucji i wymianie wiedzy opartej na AI.

W edukacji GPT Store może przekształcić sposób, w jaki nauczyciele i instytucje dzielą się materiałami dydaktycznymi. Nauczyciele różnych specjalności będą mogli nie tylko czerpać z pracy swoich kolegów, ale również dostarczać własne, unikatowe treści do wykorzystania w klasach na całym świecie. Dzięki temu nauczyciele zyskają dodatkowe źródło zarobku,

a sektor edukacji zyska dostęp do szerokiej gamy specjalistycznych narzędzi, które są ciągle aktualizowane i rozwijane przez nauczycieli praktyków.

GPT Store w kontekście edukacyjnym uwydatnia potencjał tej platformy do tworzenia globalnej społeczności edukacyjnej, w której najlepsze praktyki i innowacyjne metody nauczania mogą być łatwo udostępniane i adaptowane przez różne systemy edukacyjne. Może to także zainspirować nauczycieli, świadomych, że ich praca może być doceniona i wykorzystana na szerszą skalę, do tworzenia coraz bardziej innowacyjnych materiałów edukacyjnych. Ponadto GPT Store sprzyja większej kolaboracji i profesjonalizacji wśród edukatorów, promując ideę ciągłego rozwoju i doskonalenia narzędzi edukacyjnych. W dłuższej perspektywie platforma może zaś zrewolucjonizować ekonomię wiedzy, przekształcając sposób, w jaki zasoby edukacyjne są tworzone, dystrybuowane i wykorzystywane na całym świecie.

Multimodalność ChatGPT – obsługa różnych typów treści: ChatGPT może teraz przetwarzać teksty i obrazy

Multimodalność ChatGPT oznacza, że model ten zyskał zdolność przetwarzania nie tylko tekstów, ale i obrazów. Dzięki temu użytkownicy mogą interaktywnie współpracować z ChatGPT, używając zarówno słów, jak i treści wizualnych. To rozszerzenie możliwości ChatGPT sprzyja bardziej złożonym i bogatszym formom komunikacji, otwierając drogę do nowych zastosowań narzędzia: edukacyjnych, biznesowych i osobistych.

W edukacji ta nowa funkcja może być wykorzystana do nauczania poprzez wizualizację. Uczniowie mogą na przykład przysyłać zdjęcia artefaktów, diagramów naukowych czy prac artystycznych, a ChatGPT – dostarczać dodatkowe informacje, zadawać pytania kontrolne lub sugerować powiązane tematy do dalszej eksploracji. To umożliwi uczniom głębsze zrozumienie materiału dzięki interaktywnym i angażującym doświadczeniom edukacyjnym, łączącym słowa z obrazami.

Rozbudowana refleksja nad multimodalnością ChatGPT wskazuje na potencjał tej technologii do tworzenia bardziej angażujących i interaktywnych środowisk edukacyjnych. Może to przyczynić się do rozwoju nowych metod nauczania, które lepiej odpowiadają na potrzeby uczniów wizualnych i kinestetycznych, a także tych, którzy najlepiej przyswajają wiedzę przez praktyczne działanie i eksperymentowanie. Ponadto funkcja ta może sprzyjać rozwojowi – niezbędnej w dzisiejszym, zdominowanym przez media świecie – umiejętności

krytycznego myślenia i analizy obrazów. W kontekście globalnym multimodalność ChatGPT może także zniwelować bariery językowe, umożliwiając komunikację i uczenie się z wykorzystaniem uniwersalnego języka obrazów.

Konwersacja głosowa ChatGPT – interakcje głosowe: nowa funkcja ChatGPT, która umożliwia generowanie mowy na podstawie tekstu

Nowa funkcja ChatGPT, konwersacja głosowa, to przełom w interakcji człowieka z technologią AI. ChatGPT może teraz generować mowę na podstawie tekstu, co przekształca go w interaktywnego asystenta, zdolnego nie tylko do prowadzenia rozmów pisemnych, ale również do uczestniczenia w dialogach głosowych. Ta funkcja otwiera drzwi do nowych form komunikacji z maszynami, form, w których interfejs głosowy staje się bardziej naturalny i dostępny dla szerszego grona użytkowników.

W świecie edukacji konwersacja głosowa z ChatGPT może znacząco ułatwić naukę, zwłaszcza osobom mającym trudności z czytaniem lub pisaniem, jak również tym, którzy uczą się języka obcego. Umożliwienie uczniom komunikacji z AI poprzez mowę może poprawić ich wymowę, intonację i płynność mówienia, a także pomóc im w zdobyciu większej pewności siebie podczas mówienia. Ponadto wykorzystanie asystentów głosowych w klasie może zaktywizować uczniów, umożliwiając prowadzenie bardziej interaktywnych i angażujących zajęć.

Refleksja nad wpływem konwersacji głosowej ChatGPT na edukację ukazuje potencjalne przemiany w nauczaniu i uczeniu się. Uczniowie mogą korzystać z głosowych wersji ChatGPT do przeprowadzania dyskusji na różne tematy, ćwiczenia umiejętności słuchania lub do odczytywania tekstu na głos, co może być szczególnie pomocne w nauce języków. Konwersacja głosowa z ChatGPT może również służyć jako pomoc dla uczniów preferujących różne style uczenia się, oferując im alternatywne środki przyswajania wiedzy. W szerszym kontekście wprowadzenie mowy jako formy interakcji z AI może także przyczynić się do zwiększenia dostępności edukacji, otwierając nowe możliwości dla osób z niepełnosprawnościami lub tych, którzy preferują naukę przez słuchanie.

Dall-E – generowanie obrazów: narzędzie do tworzenia obrazów na podstawie opisów tekstowych

Dall-E 3 to innowacyjne narzędzie OpenAI, które współpracuje z ChatGPT, pozwalając na tworzenie obrazów na podstawie opisów tekstowych. Ta integracja pozwala użytkownikom na zilustrowanie abstrakcyjnych koncepcji idei lub scenariuszy, które wcześniej wymagałyby skomplikowanego opisu słownego lub zaawansowanych umiejętności graficznych.

W kontekście edukacyjnym możliwości Dall-E 3 mogą być wykorzystywane do pobudzania wyobraźni i kreatywności uczniów. Na przykład podczas lekcji kreatywnego pisania uczniowie mogą generować wizualne reprezentacje swoich opowiadań, co pomoże im w głębszym zrozumieniu tekstów i rozwijaniu narracji. Podobnie na lekcjach historii czy literatury Dall-E 3 może wizualizować abstrakcyjne i metaforyczne opisy, pomagając uczniom lepiej zrozumieć i zinterpretować materiał.

Dall-E 3 może służyć także jako narzędzie wspierające rozwój myślenia analitycznego i krytycznego. Przez analizę i porównywanie generowanych obrazów z ich tekstowymi opisami uczniowie uczą się rozumieć subtelności języka i wpływ konkretnych słów na interpretację wizualną. Jest to przydatne nie tylko w kontekście sztuki i literatury, ale również w edukacji medialnej, gdzie umiejętność krytycznej analizy treści wizualnych jest kluczowa.

Za pomocą Dall-E 3 nauczyciele mogą zachęcać uczniów do eksperymentowania z różnymi formami wyrazu, wspierając tym samym rozwój umiejętności takich jak adaptacyjność myślenia i innowacyjność. Narzędzie to otwiera przed uczniami nowe możliwości wyrażania pomysłów i uczuć, przekształcając abstrakcyjne słowa w konkretne obrazy, co wzbogaca proces edukacyjny i czyni go bardziej angażującym.

Ulepszona multimodalność w jednym oknie czatu: integracja różnych typów mediów i interakcji w pojedynczym interfejsie użytkownika

Ulepszona multimodalność w jednym oknie czatu to kolejny etap rozwoju interfejsów użytkownika, umożliwiający integrację różnych typów mediów i interakcji. Użytkownicy mają teraz możliwość prowadzenia zaawansowanych konwersacji z ChatGPT: korzystania z jego zdolności do wyszukiwania informacji w internecie, generowania obrazów za pomocą Dall-E 3, przesyłania i analizowania własnych obrazów oraz wrzucania różnorodnych plików do analizy – wszystko w ramach jednego okna czatu. To umożliwi bardziej płynną

i wszechstronną interakcję, która przekracza tradycyjne ograniczenia czatu tekstowego.

W praktyce funkcja ta umożliwia ChatGPT odgrywanie roli zaawansowanego asystenta, który może w czasie rzeczywistym przeszukiwać internet w celu znalezienia najnowszych informacji, analizować przesłane obrazy i dokumenty, a także tworzyć treści wizualne pasujące do kontekstu rozmowy. W obszarze edukacji otwiera to nowe drzwi do świata interaktywnej nauki: nauczyciel lub uczeń może przekształcić rozmowę w bogatą lekcję multimedialną (uczniowie mogą na przykład przysyłać zdjęcia do analizy historycznej, mapy, wykresy przydatne podczas dyskusji o danych statystycznych lub tworzyć obrazy, które pomagają w zrozumieniu pojęć abstrakcyjnych).

Taka integracja różnych mediów sprawia, że nauka staje się bardziej interaktywna i angażująca. Uczniowie mogą natychmiast otrzymać informację zwrotną, a nauczyciele – tworzyć bardziej dynamiczne materiały edukacyjne. W rezultacie cały proces nauczania i uczenia się staje się bardziej przystępny i dostosowany do indywidualnych potrzeb uczących się. Jest to szczególnie ważne w kontekście rozwijania umiejętności cyfrowych, krytycznego myślenia i kreatywności – kompetencji, które są kluczowe w XXI wieku.

Podsumowanie

Nowe funkcje ChatGPT otwierają przed edukacją perspektywy, które jeszcze do niedawna wydawały się odległe. Mogą znacząco przyczynić się do transformacji procesu nauczania i uczenia się, uczynić go bardziej interaktywnym, dostosowanym do indywidualnych potrzeb i bogatszym w doświadczenia multimedialne. Trzeba jednak pamiętać, że dostęp do pełni możliwości ChatGPT, takich jak ulepszona multimodalność w jednym oknie czatu, konwersacje głosowe, generowanie obrazów Dall-E 3 czy dostosowywanie i dystrybucja modeli GPTs, jest obecnie ograniczony (tylko w ramach płatnej subskrypcji).

Darmowa wersja oprogramowania oferuje podstawowe funkcje, które mimo wszystko stanowią solidną platformę do rozwijania umiejętności i zdobywania wiedzy, jednak bez pełnego spektrum interakcji multimodalnej. Warto mieć nadzieję, że z czasem i te zaawansowane opcje staną się bardziej dostępne, otwierając jeszcze szersze możliwości przed nauczycielami i uczniami na całym świecie.

3.2. Wybrane narzędzia AI, których nie opisano w poprzedniej broszurze²⁶

Google Bard

Google Bard to duży model językowy (LLM) opracowany przez Google AI. Chatbot jest nadal w fazie rozwoju, ale potrafi już wykonywać wiele zadań, w tym:

- **odpowiada na pytania** w sposób kompleksowy i informacyjny, nawet jeśli są otwarte, trudne lub dziwne;
- **tworzy różnego rodzaju treści**: wiersze, kody, skrypty, utwory muzyczne, e-maile, listy itp.;
- **tłumaczy teksty na inne języki**;
- **prowadzi zabawne i ciekawe rozmowy** z użytkownikiem.

Bard jest dostępny w wielu językach, w tym w języku polskim. Można z niego korzystać za pośrednictwem strony internetowej bard.google.com.

Bing Chat

Bing Chat to usługa zintegrowana z przeglądarką Microsoft Edge, oferująca bardziej spersonalizowany sposób wyszukiwania odpowiedzi, łącząca funkcje asystenta naukowego, osobistego planisty i kreatywnego partnera. Dzięki temu zestawowi funkcji wspomaganych przez sztuczną inteligencję **możesz wykonywać następujące czynności**:

- **Zadać pytanie.** Gdy zadajesz złożone pytania, Bing udziela szczegółowych odpowiedzi.
- **Uzyskać kompleksową odpowiedź.** Bing przegląda wyniki wyszukiwania w internecie i przedstawia uwzględniającą je odpowiedź.
- **Wykazać się kreatywnością.** Gdy potrzebujesz inspiracji, Bing pomoże Ci napisać wiersz lub opowiadanie, a nawet stworzy dla Ciebie zupełnie nowy obraz. Na czacie możesz też prosić o uzupełnienie odpowiedzi, np. „Czy możesz to wyjaśnić w prostszy sposób?” lub „Podaj mi więcej opcji”, aby uzyskać inne, jeszcze bardziej

²⁶ Została ona wydana w lipcu 2023 roku (Machura, M. (2023): *Chat GPT w szkole. Szanse i zagrożenia*, Instytut Badań Edukacyjnych).

szczegółowe dane. Specyfika narzędzia powoduje jednak, że na konwersację na czacie składa się ograniczona liczba interakcji.

- **Zadawać pytania tak jak człowiekowi.** Czyli ze szczegółami, prosząc o wyjaśnienie lub dodatkowe informacje oraz tłumacząc usłudze Bing, jak mogłaby Ci bardziej pomóc.

Usługa Bing stara się, aby odpowiedzi były oparte na faktach, ale ponieważ wciąż jest udoskonalana, **może wyświetlać nieoczekiwane lub niedokładne wyniki na podstawie danych zebranych w internecie. Najlepiej więc kieruj się własnym osądem.**

W trybie twórczym na czacie możesz **zwrócić się do usługi Bing o stworzenie obrazu.** Wystarczy napisać: „Stwórz obraz” lub „Narysuj obraz”, a otrzymasz dokładnie to, czego oczekujesz. Przygotowując odpowiedź, Bing AI analizuje również źródła w internecie.

xAI Grok

Grok to zaawansowany model języka stworzony przez xAI, firmę AI założoną przez Elona Muska. Model ten, nazywany Grok, został zaprojektowany tak, aby umożliwiać użytkownikom uzyskiwanie aktualnej wiedzy o świecie za pośrednictwem platformy X.

Grok ma na celu odpowiadanie na szeroki zakres pytań i, przede wszystkim, sugerowanie, jakie pytania należy zadać. Ta zdolność Groka ma na celu uczynienie go narzędziem wszechstronnym, zdolnym do obsługi licznych rodzajów zapytań i zadań.

O potencjale Groka stanowi Grok-1, nowoczesny model języka. Zapewnia on aktualną wiedzę o świecie i jest zdolny do obsługi różnorodnych zadań, w tym z zakresu matematyki, rozumowania i kodowania. Sprawia to, że Grok jest porównywalny z GPT-3.5 firmy OpenAI, modelem stojącym za początkową wersją ChatGPT, zwłaszcza w zakresie odpowiadania na pytania matematyczne i zdolności rozumowania.

Elephas AI

Elephas to zaawansowane narzędzie AI do transformacji treści przygotowywanych przez użytkowników: postów, artykułów blogowych czy e-maili, usprawniające cały proces pisania.

Jako narzędzie do zwiększania produktywności, Elephas ma pomagać zapracowanym profesjonalistom oszczędzać czas na nużących zadaniach pisarskich. Oferuje funkcje takie

jak ładowanie danych z popularnych źródeł, super pasek poleceń do wykonywania zadań, czat z danymi oraz klucz OpenAI do pełnego wykorzystania potencjału GPT-3.

Elephas jest również opisywany jako osobisty asystent do pisania AI, zaprojektowany, aby wspierać zapracowanych profesjonalistów. Wśród jego kluczowych funkcji wymienia się w tym kontekście m.in. narzędzie do podsumowywania artykułów online, które pomaga użytkownikom w ekstrakcji kluczowych informacji z dowolnego artykułu, podsumowując najważniejsze informacje.

Hugging Face

Hugging Face to organizacja zajmująca się sztuczną inteligencją, znana przede wszystkim z działalności w obszarze przetwarzania języka naturalnego (NLP). Oto niektóre z kluczowych cech i osiągnięć Hugging Face:

- **Open source i społeczność.** Hugging Face jest znane z tworzenia i utrzymywania szerokiej gamy narzędzi i bibliotek AI dostępnych w ramach licencji otwartej. Dzięki temu stało się hubem dla społeczności programistów i naukowców zajmujących się AI.
- **Modele językowe.** Organizacja jest szczególnie znana z biblioteki Transformer, zawierającej szereg zaawansowanych modeli językowych, takich jak BERT, GPT, RoBERTa i T5, które miały znaczący wpływ na rozwój technologii NLP.
- **Łatwość integracji.** Hugging Face koncentruje się na zapewnieniu łatwej integracji swoich narzędzi z różnymi projektami AI, oferując interfejsy programistyczne ułatwiające wykorzystanie zaawansowanych technologii językowych. Sprawia to, że Hugging Face jest dobrym miejscem zarówno dla początkujących, jak i zaawansowanych programistów zainteresowanych rozwojem nowych modeli sztucznej inteligencji.
- **Współpraca i integracja.** Firma współpracuje z wieloma innymi organizacjami i instytucjami w celu rozwoju i popularyzacji technologii AI, co czyni jej narzędzia szeroko stosowanymi w projektach badawczych zarówno w środowisku akademickim, jak i komercyjnym.
- **Demokratyzacja AI.** Hugging Face angażuje się w demokratyzację dostępu do zaawansowanych technologii AI, oferując je za darmo lub w ramach licencji otwartych, co umożliwia ich szersze wykorzystanie.

3.3. Pozytywne oraz negatywne aspekty nowych funkcji modeli językowych

Spróbujmy zastanowić się nad pozytywami i negatywami nowych (nieistniejących w momencie publikacji poprzedniej broszury) funkcji modeli językowych.

Wyszukiwanie danych w internecie umożliwia szybkie znalezienie aktualnych informacji, które mogą nie być zawarte w bazie wyuczonego modelu. Istnieje jednak ryzyko, że może to zachęcać uczniów do bezkrytycznego korzystania z dostępnych źródeł informacji. Dlatego też w pracy pedagogicznej z uczniami, którzy wykorzystują do wyszukiwania informacji boty, należałoby położyć nacisk na weryfikowanie danych oraz krytyczne myślenie.

Użytkownicy obecnych modeli językowych mogą **komunikować się z nimi w celu uzyskiwania obrazów**, co jest użyteczne na zajęciach plastycznych, przyrodniczych czy zajęciach projektowania graficznego i sztuki. Rozmowa z botem generującym obrazy pozwala na wizualizację koncepcji oraz otwiera przed użytkownikiem całe spektrum możliwości, wspierając go w procesie tworzenia dzieła. **Konieczne jest jednak zachowanie ostrożności w interpretacji generowanych obrazów, gdyż te nie zawsze muszą być wiarygodne.** To z kolei może negatywnie wpływać na wyobraźnię twórczą dziecka, ograniczając ją poprzez dostarczanie dziecku gotowych produktów. Profesor Jerzy Vetulani²⁷ – polski psychofarmakolog, neurobiolog, biochemik, profesor nauk przyrodniczych, członek Polskiej Akademii Nauk i Polskiej Akademii Umiejętności – często wspominał o środowisku wzbogaconym, czyli takim, w którym **ludzki mózg, aby się rozwijać, powinien rozwiązywać problemy**. Tymczasem podrzucanie człowiekowi gotowych rozwiązań ogranicza w nim procesy kreatywne. Z drugiej jednak strony prowadzenie konwersacji z botem, który tworzy grafiki, może też wspierać nauczanie (identyfikacja fotografowanych przedmiotów, przygotowywanie instrukcji obsługi urządzeń), choć należy pamiętać, że **system nie zawsze musi prawidłowo identyfikować przedmioty i obiekty**.

Rozmowa głosowa z ChatGPT w wersji mobilnej ułatwia naukę uczniom ze specjalnymi potrzebami edukacyjnymi, mającym trudności z pisaniem i czytaniem. Może też sprzyjać nauce języków obcych poprzez uczenie wymowy.

W nowych LLM uczniowie mogą używać **Advanced Data Analysis** do programowania

²⁷ https://pl.wikipedia.org/wiki/Jerzy_Vetulani [dostęp: 16.11.2023].

i rozwiązywania skomplikowanych problemów matematycznych oraz analizowania danych. Podobnie jednak jak w przypadku innych funkcji istnieje ryzyko, że uczniowie będą polegać na narzędziu zamiast rozwijać własne umiejętności.

Podsumowując, nowe funkcje zaawansowanych modeli językowych oferują szereg potencjalnych korzyści dla edukacji, ale **ważne jest, aby uczniowie korzystali z nich pod nadzorem (czy zgodnie ze wskazaniem) nauczyciela czy rodzica, którzy zawsze mogą zareagować, jeżeli stwierdzą, że istnieje którekolwiek z wyżej wymienionych ryzyk.**

IV. Do czego AI nie służy – szkolne przykłady działań niepożądanych

Mateusz Rafał

4.1. Przykłady niewłaściwego wykorzystywania narzędzi AI przez uczniów, nauczycieli i dyrektorów szkół

W nawiązaniu do wcześniej przedstawionych informacji możemy teraz wyróżnić rodzaje działań, które nie powinny być podejmowane przy wykorzystaniu narzędzi sztucznej inteligencji. Ta sekcja zostanie podzielona na działania podejmowane przez poszczególnych uczestników procesu edukacyjnego: ucznia, nauczyciela i dyrektora szkoły.

Przykłady niewłaściwego wykorzystywania narzędzi AI przez uczniów:

- Zastępowanie rozmowy z drugą osobą. AI nie zastąpi drugiego człowieka w dyskusji o problemach uczniów ani nie zmierzy się z realną oceną działań drugiego człowieka. Opinia AI nie może stanowić kryterium oceny drugiego człowieka.
- Zastępowanie konsultacji z lekarzem rozmową z AI. Wszystkie problemy zdrowotne, w szczególności ludzi młodych, powinny być konsultowane przez wykwalifikowanego lekarza.
- Uznawanie opinii AI za nieomyślne i zawsze poprawne. Młodzi ludzie często nie sprawdzają źródła ani poprawności informacji podanej przez AI. Tymczasem sztuczna inteligencja może się mylić. AI należy używać do pogłębiania wiedzy, a nie do tworzenia i kopiowania zadań i opracowań.
- Wykorzystywanie AI do ściągania na klasówkach lub plagiatowania prac.

Przykłady niewłaściwego wykorzystywania narzędzi AI przez nauczycieli:

- Ocenianie zachowań oraz osiągnięć uczniów.
- Wyciąganie opinii na temat oryginalności pracy uczniów na podstawie narzędzi AI. Jeżeli weryfikujesz samodzielność pracy ucznia za pomocą narzędzi AI i otrzymujesz informację, że z dużym prawdopodobieństwem praca nie została napisana samodzielnie, to zweryfikuj to, zadając uczniowi kilka pytań na temat pracy.

- Wyciąganie wniosków o użyciu AI po stwierdzeniu w pracy takich samych błędów, jakie popełnia sztuczna inteligencja. Czasami błędy mogą sugerować wykorzystanie sztucznej inteligencji w pracy ucznia, ale takie wnioski bywają pochopne.
- Wykorzystywanie bezpłatnych narzędzi AI do sporządzania notatek z kluczowych spotkań, na których podawane są dane osobowe uczniów. Pod żadnym pozorem nie wykorzystuj narzędzi AI (w tym modeli językowych) do pracy z takimi dokumentami, jak: protokoły posiedzeń rad pedagogicznych, korespondencja z PPP (np. opinie o dziecku wymagane przez instytucje takie jak sąd czy poradnia psychologiczno-pedagogiczna), wielospecjalistyczna ocena poziomu funkcjonowania ucznia (WOPFU). Takie działanie powoduje, że przekazujesz narzędziom AI dane, do których te nie powinny mieć dostępu.

Przykłady niewłaściwego wykorzystywania narzędzi AI przez dyrektorów szkół:

- Wykorzystywanie AI do merytorycznej oceny raportu przygotowanego przez pracownika. Nawet najlepsze narzędzia AI nie zastąpią wewnętrznej wiedzy menedżera placówki oświatowej, którym jest jej dyrektor. Wykorzystywanie ich w celu oceny pracownika może prowadzić do błędnych i krzywdzących wniosków.
- Wypełnianie wniosków/działania administracyjne. Narzędzia AI nie służą obecnie do zastępowania człowieka na polu działań administracyjnych, a efekt ich pracy jest często obarczony błędem. Nie należy więc przygotowywać dokumentacji urzędowej lub komunikacji formalnej (w tym decyzji administracyjnych, w których wskazuje się podstawę prawną) z wykorzystaniem modeli językowych AI.

4.2. Zagrożenia przy wdrażaniu rozwiązań AI w edukacji

Musimy pamiętać, że współczesne technologie, w tym sztuczna inteligencja (AI), zyskują na znaczeniu w edukacji, oferując nowe możliwości personalizacji nauki oraz oceny i wsparcia uczniów, **jednak zastosowanie AI w szkole wiąże się także z pewnymi niebezpieczeństwami**. Poniżej przedstawiamy główne zagrożenia, które powinny być brane pod uwagę przy wdrażaniu takich rozwiązań w środowisku edukacyjnym.

Zagrożenia zostały podzielone na następujące kategorie: ochrona danych, uczeń a AI, zdrowie fizyczne i psychiczne, nauczyciel a AI, etyka i transparentność w korzystaniu z AI, dostępność technologii.

Ochrona danych

W szkołach przechowywane są różnorodne dane, które obowiązek ochrony wynika z przepisów prawa lub standardów etycznych. Orzeczenia o potrzebie kształcenia specjalnego czy protokoły rad pedagogicznych zawierają informacje, które mogą wpłynąć na życie uczniów i ich rodzin, jeśli zostaną ujawnione. Jeżeli systemy oparte na AI nie są odpowiednio zabezpieczone, mogą stanowić łakomy kąsek dla cyberprzestępców. Nawet jeśli nie dojdzie do włamania, samo ryzyko ujawnienia takich informacji może powodować obawy rodziców i nauczycieli, dlatego w przypadku korzystania z ogólnodostępnych rozwiązań, takich jak ChatGPT czy Bard, **nie powinno się w rozmowie z AI przekazywać żadnych danych osobowych ani danych objętych tajemnicą rady pedagogicznej.**

W celu przetwarzania takich danych szkoła powinna zadbać o zamknięte i w pełni przez siebie kontrolowane środowisko dla danego modelu językowego (np. wstępnie wytrenowany model na lokalnym serwerze, „karmiony” danymi z zamkniętego środowiska edukacyjnego, dostępny tylko dla osób uprawnionych). **Jednak nawet wtedy istnieje ryzyko wycieku danych, dlatego należy właściwie zabezpieczyć system przed dostępem osób nieupoważnionych.**

Uczeń a AI

W dobie rewolucji związanej z upowszechnieniem narzędzi sztucznej inteligencji coraz częściej rodzi się obawa, że uczniowie mogą być zbyt zależni od technologii, przez co tracą zdolność do samodzielnego myślenia i rozwiązywania problemów. W związku z tym celem instytucji publicznych, w tym szkół i nauczycieli, jest zapewnienie takiego zrozumienia materii w środowisku szkolnym, aby sztuczna inteligencja była traktowana jako narzędzie wspierające, a nie zastępujące tradycyjne metody nauczania.

Dla przykładu uczniowie mogą:

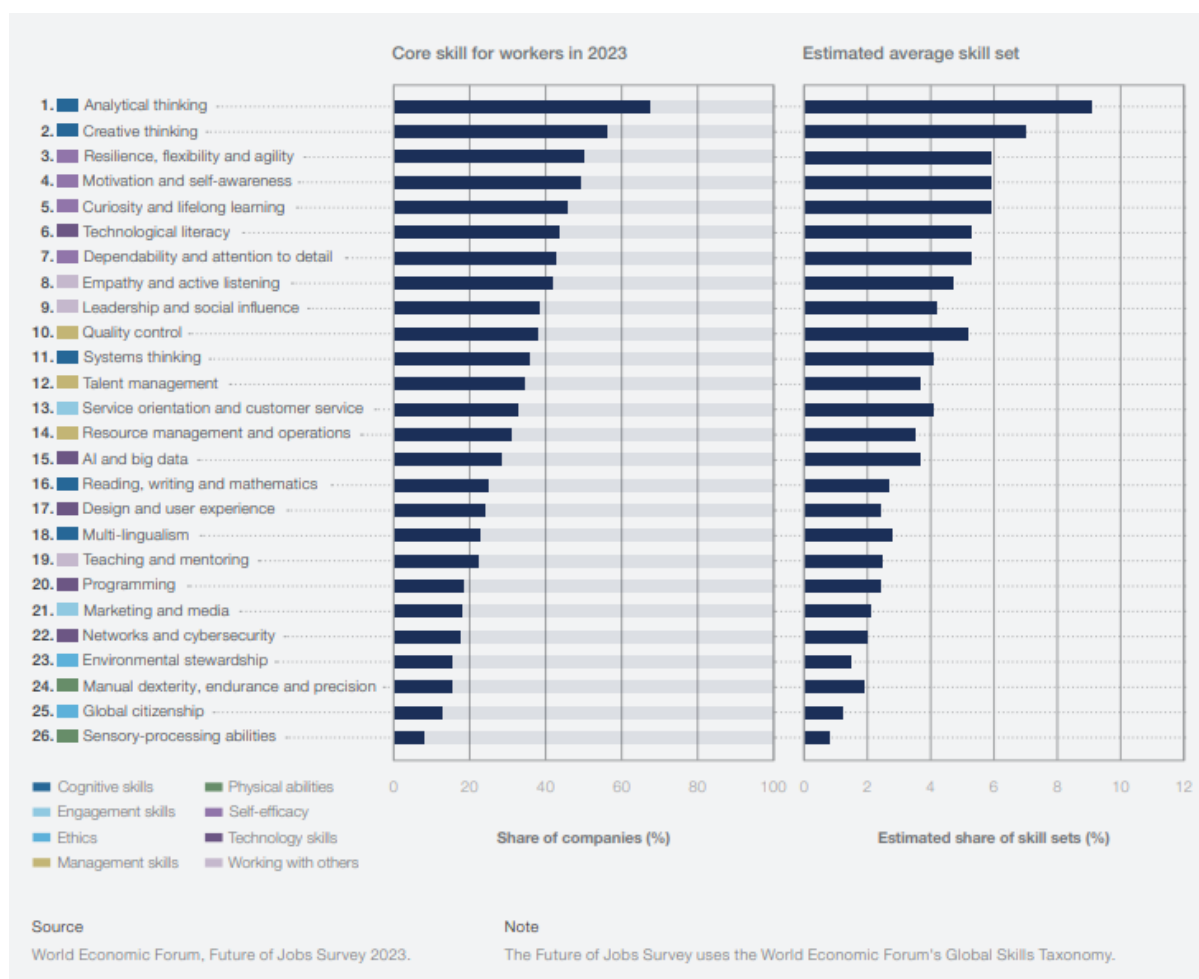
- polegać na narzędziach do tłumaczenia zamiast uczyć się nowych języków;
- korzystać z narzędzi AI, które umożliwiają rozwiązywanie skomplikowanych równań i zadań otwartych, zamiast rozwijać własne umiejętności matematyczne;
- poprosić AI o wygenerowanie wypracowania na danych temat, a nawet „nauczyć się” stylu pisania bota, by jeszcze bardziej uśpić czujność nauczyciela – to powoduje, że

uczniowie tracą możliwość zdobycia umiejętności sprawnego pisania, argumentowania, jasnego wyrażania myśli.

Należy pamiętać, że nadmierne poleganie na technologii może hamować rozwój krytycznego myślenia i innych kluczowych umiejętności. Poza tym uczniowie mogą stać się zbyt zależni od narzędzi AI i stracić zdolność do wykonywania zadań bez wsparcia technologicznego. Współczesny system edukacji powinien przygotowywać uczniów do funkcjonowania w świecie, w którym technologia jest wszechobecna, ale też uczyć ich samodzielności i umiejętności radzenia sobie bez niej.

Trzeba mieć również na uwadze, że bezkrytyczne i niekontrolowane korzystanie przez uczniów z narzędzi AI może prowadzić do ograniczenia ich kreatywności. Systemy oparte na AI, które generują wyczerpujące odpowiedzi, mogą nie zachęcać uczniów do wykraczania poza schematy myślowe czy do testowania różnych rozwiązań. Nadużywanie takich narzędzi minimalizuje więc przestrzeń eksploracji uczniów. W ten sposób zmarginalizowane zostaje również twórcze myślenie i podejmowanie ryzyka w procesie nauczania. Jeśli zaś AI zbyt upraszcza rozwiązywanie problemów lub podaje gotowe odpowiedzi, uczniowie mogą stracić szansę rozwijania kreatywności i analitycznego myślenia, a te dwie kompetencje są w dzisiejszym świecie bardzo istotne²⁸.

²⁸ Raport Światowego Forum Ekonomicznego z 2023 roku: <https://www.weforum.org/agenda/2023/05/future-of-jobs-2023-skills/>
https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf [dostęp: 16.11.2023].



Toczy się obecnie wiele dyskusji na temat błęd w procesie edukacyjnym i doniosłości praktycznej, która temu towarzyszy. Musimy zatem pamiętać, że jeżeli narzędzia AI natychmiast korygują błędy uczniów, ci mogą nie poznać wartości popełniania błędów jako części procesu nauki. Błędy są naturalną częścią tego procesu i pomagają uczniom zrozumieć, co robią źle. Jeżeli AI interweniuje zbyt szybko, uczniowie mogą nie mieć okazji do refleksji nad własnymi błędami²⁹.

W obecnych czasach, w epoce postępu sztucznej inteligencji, szczególne znaczenie przypisuje się umiejętności krytycznego myślenia. Pozwala ona na odróżnienie informacji rzetelnych od iluzji generowanych przez sztuczną inteligencję, a także – w sytuacjach bardziej niebezpiecznych – od informacji fałszywych, zmanipulowanych.

²⁹ Więcej o wartości błędów w procesie edukacyjnym:
<https://www.edunews.pl/system-edukacji/szkoly/6007-o-rol-i-bledu-w-uczeniu-i-relacjach>
<https://ceo.org.pl/jak-wykorzystac-potencjal-bledu-by-stal-sie-tworczy-a-nie-destrukcyjny/> [dostęp: 16.11.2023].

Korzystanie z narzędzi umożliwiających tworzenie treści powinno być połączone z odpowiedzialnością wynikającą z ich używania.

Powszechny dostęp do rozwiązań opartych na sztucznej inteligencji może prowadzić do zaniku krytycznego myślenia i samodzielnego wnioskowania wśród uczniów, którzy za bardzo polegają na tych technologiach. Wspieranie się technologią w procesie nauczania jest istotne, jednakże uczniowie powinni być świadomi konieczności rozwijania umiejętności samodzielnej oceny otaczającej ich rzeczywistości, zwłaszcza że nie wszystkie treści generowane przez AI odzwierciedlają stan faktyczny, przez co zarówno uczniowie, jak i nauczyciele są narażeni na dezinformację, która może wpłynąć na ich poglądy czy decyzje.

Innym **istotnym zagrożeniem**, które niesie za sobą sztuczna inteligencja, **jest rezygnacja z relacji międzyludzkich. Uczeń, który korzysta z osobistego asystenta AI w czasie nauki, może przestać rozwijać umiejętność komunikacji interpersonalnej**, koncentrując się wyłącznie na interakcji z maszyną. Sztuczna inteligencja może być wartościowym narzędziem wsparcia, ale nie zastąpi autentycznych relacji międzyludzkich, które są kluczowe dla emocjonalnego i społecznego rozwoju człowieka. Nadmierne wykorzystanie technologii może izolować uczniów ani utrudniać nabywania przez nich istotnych umiejętności społecznych.

Zdrowie fizyczne i psychiczne

Istnieje ryzyko, że powszechny dostęp do technologii opartych na sztucznej inteligencji skutkować będzie zwiększonym czasem spędzonym przed ekranem komputera lub smartfona. Osobami szczególnie narażonymi na negatywne skutki tego zjawiska są dzieci. Aby przeciwdziałać jego konsekwencjom, **szkoły powinny zadbać o balans między wykorzystaniem technologii a aktywnością fizyczną uczniów.**

Warto pamiętać, że brak odpowiednich środków ostrożności w stosowaniu technologii w życiu uczniów może prowadzić do nasilenia uczucia izolacji, wystąpienia lęku czy depresji. W momencie, gdy szkoły zaczną powszechnie korzystać ze sztucznej inteligencji, zauważalny będzie wzrost tempa pracy i produktywności. Jednakże nie wszyscy dostosują się łatwo do tych zmian. Nauczyciele będą w stanie efektywniej śledzić postępy uczniów i monitorować ich pracę, co może wywierać nadmierną presję na uczniów, zobowiązując ich do ciągłego osiągania lepszych wyników.

Należy zawsze uwzględniać czynnik ludzki. Mimo że nauczyciele potrafią identyfikować emocjonalne potrzeby uczniów i dostosowywać się do nich, systemy sztucznej inteligencji mogą nie być w stanie rozpoznać czy zrozumieć ludzkich uczuć i zachowań. W przypadku wykrycia podejrzanych sygnałów w zachowaniu dzieci i młodzieży to nauczyciele powinni reagować. **Tylko nauczyciel jest bowiem w stanie zauważyć różne subtelności i rozpoznać istotne zmiany w nastroju ucznia.**

Nauczyciel a AI

Nauczyciele zastanawiają się, czy używać sztucznej inteligencji, czy stanowi ona dla nich szansę, czy raczej zagrożenie dla ich pracy, a także czy AI może zwiększyć dystans między nimi a uczniami. **Z jednej strony technologia sztucznej inteligencji może zrewolucjonizować i ulepszyć proces edukacyjny oraz stać się pożytecznym narzędziem pracy, z drugiej zaś jej nieumiejętne wykorzystanie może doprowadzić do sytuacji konfliktowych oraz etycznie wątpliwych.** I to właśnie przed nauczycielami stoi trudne zadanie pokazania uczniom, jak korzystać ze sztucznej inteligencji w sposób odpowiedzialny i etyczny. Poniżej zostaną wymienione szczególne trudności, których doświadczają i mogą doświadczać nauczyciele.

Konieczne jest zajęcie się kwestiami etycznymi związanymi z wykorzystywaniem technologii sztucznej inteligencji. Choć algorytmy sztucznej inteligencji mają potencjał do bycia obiektywnymi i sprawiedliwymi, to istnieje ryzyko programowania ich w sposób, który skutkować będzie nieobiektywną oceną lub nawet dyskryminacją danego ucznia. To zjawisko mogłoby rodzić poważne dylematy moralne i etyczne. Należy pamiętać, że technologie sztucznej inteligencji uczą się na podstawie danych, na których są szkolone. Jeśli dane te zawierają uprzedzenia, także one mogą zostać przejęte i utrwalone przez system sztucznej inteligencji.

Co więcej, technologie sztucznej inteligencji nie posiadają zdolności rozumienia ludzkiego pojmowania rzeczywistości. Nie wiedzą, z czym wiążą się indywidualne trudności ucznia, w tym na przykład walka ze wstydem czy z lękiem przed porażką. Brak zrozumienia tych okoliczności może prowadzić do nieadekwatnej oceny sytuacji i braku odpowiedniego wsparcia w obszarach, które są kluczowe dla dalszego rozwoju ucznia.

Wprowadzenie sztucznej inteligencji do systemu edukacji niesie za sobą także ryzyko nadmiernej standaryzacji. Kiedy algorytmy decydują o tym, które treści są najbardziej

odpowiednie dla uczniów, istnieje realne niebezpieczeństwo, że indywidualność w procesie uczenia zostanie zepchnięta na drugi plan. Z ostrożnością należy więc podchodzić do narzędzi, które umożliwiają automatyczną ocenę pracy uczniów. Ich bezrefleksyjne używanie skutkować może bowiem nadmiernym skupieniem uwagi na konkretnych wytycznych, kryteriach i nie uwzględniać głębszego, odmiennego rozumienia tematu przez ucznia. Automatyczne spojrzenie może oceniać jedynie powierzchowne aspekty pracy, nie zaś niuanse opracowania czy przejawy kreatywnego podejścia ucznia do zadania. Niepożądana byłaby sytuacja, w której uczniowie, zamiast w pełni zrozumieć i przyswoić materiał, zaczęliby dostosowywać się do algorytmu, żeby tylko osiągnąć jak najwyższy wynik.

Obawy o przyszłość edukacji budzi także ewentualna utrata autonomii decyzyjnej nauczycieli, która może wynikać z rosnącej roli narzędzi sztucznej inteligencji. Wspomaganie nauczania przez sztuczną inteligencję może powodować, że decyzje dotyczące oceny, personalizacji materiałów czy nawet planowania lekcji będą coraz bardziej narzucane przez algorytm. Tymczasem **sztuczna inteligencja powinna przede wszystkim wspierać nauczycieli, umożliwiając im dostosowanie narzędzi do indywidualnych potrzeb uczniów, a nie pozbawiać ich możliwości podejmowania kluczowych decyzji pedagogicznych.**

W kontekście sztucznej inteligencji pojawia się również pytanie o autorytet nauczyciela i możliwość jego zakwestionowania. Niewątpliwie istotne jest, aby w miarę wspomaganie procesów edukacyjnych przez technologie AI podkreślać, że to nauczyciele posiadają unikalne umiejętności, które trudno jest zastąpić algorytmami. Ich zdolność do odkrywania potencjału ucznia, do inspirowania do nauki, dostosowywania się do indywidualnych potrzeb uczniów oraz kreatywnego i abstrakcyjnego myślenia jest niezastąpiona. Nie wolno zapominać także, że to z nauczycielem, a nie z narzędziem sztucznej inteligencji, uczeń może budować relację, która będzie wspomagała uczenie się. Zbyt częste wykorzystywanie technologii w edukacji może skutkować rozluźnieniem więzi społecznych i spadkiem interakcji uczniów zarówno z nauczycielami, jak i rówieśnikami. W rezultacie **nieumiejętne i nadmierne korzystanie z technologii AI powoduje wzrost poczucia samotności uczniów, ich izolację czy właśnie pogorszenie się relacji międzyludzkich, które są filarem zdrowia psychicznego.**

Ważnym aspektem jest także kwestia prywatności i ochrony danych osobowych. Gdy uczniowie lub nauczyciele korzystają z technologii AI, ich rozmowy i dane osobowe mogą być przechowywane i analizowane przez dostawcę narzędzia, co stanowi zagrożenie dla ich prywatności. W przypadku publicznych narzędzi sztucznej inteligencji to nauczyciele powinni zachować szczególną ostrożność i powstrzymać się od wprowadzania lub ujawniania poufnych informacji o sobie lub uczniach, w tym między innymi prywatnej korespondencji, danych osobowych, wyników w nauce, informacji finansowej czy też danych objętych tajemnicą medyczną.

Uwagę należy poświęcić także problemowi niedostatecznego przeszkolenia nauczycieli w zakresie efektywnej pracy z narzędziami sztucznej inteligencji. W kontekście dynamicznie zmieniającego się świata technologii istnieje pilna potrzeba zorganizowania szkoleń, w czasie których edukatorzy uzyskają nie tylko podstawową wiedzę na temat AI, ale i poznają praktyczne aspekty łączenia tych narzędzi z procesem nauczania. Jest to konieczne, aby nauczyciele mogli być przewodnikami dla uczniów, realnym wsparciem w scyfryzowanym świecie, w którym coraz większą rolę odgrywać będą technologie sztucznej inteligencji.

Etyka i transparentność w korzystaniu z AI

Wszechobecny dostęp do narzędzi sztucznej inteligencji może przyczynić się do nieograniczonego korzystania z ich możliwości, co oznacza zarówno szanse, jak i zagrożenia. Istnieje duże zagrożenie, że treści wygenerowane przez AI zostaną potraktowane jako w pełni wytworzone przez człowieka. Bardzo ważne jest więc, aby oznaczać fragmenty przygotowane przez sztuczną inteligencję oraz opisywać jej udział w tworzeniu materiałów autorskich.

W związku z powyższym **każdy użytkownik narzędzi sztucznej inteligencji musi mieć świadomość konieczności przestrzegania standardów etycznych i wpajania ich dzieciom od najmłodszych lat**. Pytanie jednak brzmi: czy taki proces jest w ogóle możliwy? Tak. I warto go rozpocząć od redefinicji prac domowych. Wraz z rozwojem narzędzi AI rośnie ryzyko, że uczeń obciążony zbyt dużą liczbą wypracowań oraz zadań matematycznych wykorzysta sztuczną inteligencję do wykonania pracy domowej. To zaś wymusi zmianę prowadzenia lekcji w szkołach, a z czasem również rewizję podstawy programowej. Nie da się bowiem w sposób efektywny i transparentny kontrolować

potencjalnych nadużyć ani związanej z nimi pokusy. Praca domowa ulegnie radykalnej zmianie i będzie musiała mieć charakter bardziej praktyczny: skupiać się na realizacji projektów i interakcji z rówieśnikami.

Dostępność technologii

Koszty i dostępność

Różnice w dostępie do odpłatnych technologii wykorzystujących AI mogą prowadzić do powstania nierówności edukacyjnych. Taka dysproporcja stwarza zagrożenie szczególnie dla uczniów mniej zamożnych, których nie stać na korzystanie z odpłatnej wersji AI, zawierającej najaktualniejsze dane. Może to skutkować zmniejszeniem szans mniej zamożnych uczniów na szybszy rozwój.

Zależność od dostawców technologii

Uzależnienie od jednego dostawcy AI może prowadzić do ujednolicenia wybranych technologii oraz ograniczonego wprowadzania innowacyjnych rozwiązań w przyszłości. Może też skutkować brakiem elastyczności w dostosowywaniu narzędzia do zmieniających się potrzeb edukacyjnych. W przypadku wystąpienia problemów technicznych lub nagłej zmiany warunków użytkowania szkoły mogą mieć trudności w utrzymaniu ciągłości nauczania. W obliczu dynamicznie rozwijającej się technologii kluczem do efektywnego wykorzystania sztucznej inteligencji w procesie edukacji jest zatem dywersyfikacja rozwiązań oraz wykorzystywanych technologii i sprzętu.

Należy także mieć świadomość, że technologia sama w sobie nie jest ani dobra, ani zła, jej wartość uzależniona jest bowiem od sposobu jej używania, dlatego istotne jest, aby proces wdrażania AI w środowisku edukacyjnym był przemyślany i poparty solidnymi badaniami, uwzględniającymi zarówno potencjalne korzyści, jak i zagrożenia wynikające z jej używania. Wymaga to jednak współpracy między Ministerstwem Edukacji i Nauki, firmami z branży EdTech, nauczycielami, rodzicami i uczniami. Musimy mieć bowiem pewność, że technologia służy prawdziwym potrzebom uczniów. Im bardziej świadomie będziemy kształtować polskie środowisko edukacyjne w erze cyfrowej, tym większą zyskamy pewność, że w przyszłości sztuczna inteligencja będzie pracować ramię w ramię z człowiekiem, a co za tym idzie – także wesprze i urozmaici proces nauczania.

Podsumowanie

Przygotowując podsumowanie, posłużyliśmy się modelem językowym ChatGPT-4, aby pokazać praktyczne zastosowanie narzędzia. Sądzymy, że jest to ciekawe przedstawienie jednego z wielu sposobów wykorzystywania modeli językowych opisywanych w tej broszurze.

Dostarczyliśmy botowi jako prompt broszurę w formacie .pdf i wydaliśmy mu polecenie:

„Napisz podsumowanie załączonej broszury edukacyjnej dla nauczycieli w Polsce, koncentrującej się na wykorzystaniu sztucznej inteligencji w edukacji. Tekst powinien mieć długość od pół do jednej strony A4 i stanowić ostatni podrozdział broszury, skutecznie podsumowując wszystkie zawarte w niej tematy. Należy uwzględnić krótką refleksję na temat poruszanych zagadnień. Tekst powinien być napisany po polsku, z zachowaniem właściwej struktury i stylu, odpowiedniego dla edukacyjnego materiału informacyjnego, mając na uwadze jego publikację na końcu broszury”.

Oto tekst, który został wygenerowany przez LLM:

„Niniejsza broszura edukacyjna, zatytułowana *Do czego AI nie służy*, jest skierowana do nauczycieli i nauczycielek polskich szkół. Jest to kontynuacja publikacji *Chat GPT w szkole. Szanse i zagrożenia* i skupia się na kreatywności generatywnej sztucznej inteligencji (Gen AI), etyce i moralności w kontekście AI, wykorzystaniu modeli językowych, a także wskazuje sytuacje, w których korzystanie AI w szkole może być niepożądane.

Publikacja rozważa kreatywność modeli językowych, ich zdolności do generowania nowych treści i potencjalnych zastosowań w edukacji. Porusza również kwestię etyki i moralności w pracy z dużymi modelami językowymi, prezentując wskazówki i przestrogi dla nauczycieli. Wskazane są również przykłady działań niepożądanych w pracy z modelami językowymi w szkole.

Niektóre z zagadnień, które nie powinny być podejmowane z użyciem modeli językowych, to leczenie i diagnoza, naukowe lub akademickie badania, porady prawne, tłumaczenia wymagające dużej dokładności, przetwarzanie poufnych informacji, krytyczne decyzje życiowe, sytuacje wymagające zrozumienia emocji, kwestie etyczne i moralne, sytuacje awaryjne, oraz dane personalne.

Znaczenie AI w edukacji wzrasta, a nauczyciele muszą być przygotowani do korzystania z narzędzi AI w etyczny i odpowiedzialny sposób. Etyka odgrywa kluczową rolę w podejmowaniu decyzji moralnych i rozwijaniu zasad współżycia społecznego, co jest istotne przy projektowaniu i wykorzystywaniu technologii AI w edukacji.

Podsumowując, ta broszura to przewodnik dla nauczycieli, który pomaga zrozumieć, jak korzystać z AI w sposób odpowiedzialny i etyczny, zwracając uwagę na ograniczenia i potencjalne zagrożenia związane z jej wykorzystaniem w edukacji. Cel to wspieranie nauczycieli w przygotowaniu uczniów do życia w świecie, gdzie AI odgrywa coraz większą rolę, jednocześnie zachowując świadomość odpowiedzialności i etycznych aspektów korzystania z tych narzędzi”.

Do kogo należy, będący odpowiedzią na prompt, powyższy fragment, który został wygenerowany przez model językowy? O tym już w następnej publikacji.

O autorach

Tomasz Łukawski (t.lukawski@ibe.edu.pl)

Doktor nauk społecznych w dyscyplinie pedagogika. Doktorat obronił na Uniwersytecie im. Adama Mickiewicza w Poznaniu. Tytuł dysertacji: *Wzajemne uczenie się uczniów dwóch społeczności szkolnych (w kontekście tworzenia cyfrowych narzędzi edukacyjnych)*. Absolwent studiów podyplomowych na Wydziale Zarządzania Uniwersytetu Warszawskiego oraz Instytutu Informatyki Uniwersytetu Wrocławskiego. Studia magisterskie ukończył na Uniwersytecie Warszawskim na kierunku pedagogika. Ekspert w Centrum Transformacji Cyfrowej, autor *Polityki cyfrowej transformacji edukacji w Polsce*, koordynujący prace nad tym dokumentem w Instytucie Badań Edukacyjnych, członek Rady Programowej Szerokiego Porozumienia na rzecz Umiejętności Cyfrowych i Zatrudnienia.

Adam Łukawski (adam.lukawski@orpheusinstituut.be)

Kompozytor i programista. Wykładowca przedmiotu *muzyka ze sztuczną inteligencją* w konserwatorium w Amsterdamie. Artysta-naukowiec (artist researcher) w grupie naukowej MetamusicX w Orpheus Institute w Gandawie. Doktorant na Uniwersytecie w Lejdzie w Holandii – praca doktorska na temat komponowania muzyki z wykorzystaniem technologii AI i blockchain. Ukończył studia magisterskie z kompozycji muzyki w Amsterdamie (Conservatorium van Amsterdam) i w Londynie (Guildhall School of Music & Drama). Jego kompozycje są wydawane przez Donemus Publishing w Hadze, a ich premiery były grywane w takich instytucjach kultury jak Opera Narodowa i Balet w Amsterdamie, Gaudeamus Festival w Utrechcie, Studio Koncertowe Polskiego Radia i m. Witolda Lutosławskiego w Warszawie, Barbican Milton Court w Londynie czy Erhbar Hall w Wiedniu. Jako programista zdobył doświadczenie, pracując w start-upach, w Polskim Towarzystwie Informatycznym oraz jako technologiczny keynote speaker w Cortland/Apple Inc. Jako wykładowca bierze udział w licznych międzynarodowych wydarzeniach naukowych i artystycznych.

Mateusz Rafał (mateusz.rafał@mein.gov.pl)

Absolwent WPIA UW, doktorant w Kolegium Zarządzania i Finansów Szkoły Głównej Handlowej, ekspert Polskiego Instytutu Ekonomicznego, autor i współautor licznych publikacji, wykładowca akademicki, obecnie dyrektor Centrum Transformacji Cyfrowej

w Ministerstwie Edukacji i Nauki. W przeszłości pełnił funkcję członka zespołu eksperckiego Prezesa Rady Ministrów i doradcy ministrów konstytucyjnych. Pracował w sektorze bankowym oraz akademickim.