

Sztuczna Inteligencja: Kluczowe Zagadnienia i Zalecenia

Niniejszy dokument stanowi zwięzłe podsumowanie kluczowych zagadnień i zaleceń Szkolenia dotyczącego wykorzystania **sztucznej inteligencji (AI - Artificial Intelligence)** w obszarze kultury. Szczególną uwagę zwrócono na weryfikację wytworów generowanych przez AI oraz na świadome i odpowiedzialne korzystanie z tych technologii.

Bezpieczna Lista Kontrolna (Checklist Level Prudent)

Świadome i Odpowiedzialne Korzystanie z AI:

- **Unikanie generowania szkodliwych treści** za pomocą AI, takich jak **dezinformacja** czy **mowa nienawiści**.
- **Świadomość wpływu AI na środowisko naturalne.** Każde zapytanie do modelu AI wiąże się ze zużyciem energii i emisją CO₂. Przykładowo:
Pojedyncze zapytanie do ChatGPT generuje emisję ok. 4.32g CO₂.
16 zapytań odpowiada emisji związanej z zagotowaniem czajnika wody (ok. 70g CO₂).
Wykonanie 139 zapytań generuje emisję porównywalną ze standardowym praniem w temperaturze 30 stopni Celsjusza (ok. 600g CO₂).
Około 92000+ zapytań wiąże się z emisją CO₂ na poziomie przelotu samolotem na trasie Kraków-Madryt i z powrotem (ok. 2160 km w jedną stronę), wynoszącą ok. 400 kg CO₂.
- **Przeciwdziałanie zjawisku erozji myślenia krytycznego.** Nadmierne poleganie na AI może prowadzić do spadku umiejętności samodzielnego weryfikowania informacji, rozwiązywania problemów i krytycznej analizy.

Weryfikacja i Krytyczne Podejście: Należy zachować krytyczne podejście do treści generowanych przez AI. Nie należy traktować ich jako niepodważalnego źródła wiedzy.

- **Wytwory AI jako punkt wyjścia** do dalszych analiz, prac i badań, a nie ostateczny produkt (kwestie praw autorskich oraz podejścia bezpiecznego)
- **Analiza** generowanych treści pod kątem **błędów logicznych i niespójności**.
- **Jednolitość stylu.** Nagłe zmiany mogą wskazywać na tzw. **halucynacje (majaczenie)**
- **Weryfikowanie informacji** generowanych przez AI poprzez porównanie ich z wiarygodnymi źródłami (np. Wikipedia, publikacje naukowe, uznane opracowania). W tym **aktualności informacji**. Wiedza modeli AI jest ograniczona do określonego momentu w czasie (np. dla ChatGPT jest to **grudzień 2023**).
- **Również obrazy, filmy (i dźwięki) wygenerowane przez AI wymagają weryfikacji autentyczności.** Należy zachować ostrożność wobec zjawiska **deepfake** i innych technik manipulacji obrazem.
- **Podwójne upewnienie się, że przez użycie AI nie przypisujemy sobie autorstwa istniejących dzieł.**

Fakty i Mity AI

MIT: Interakcje z modelami AI (np. ChatGPT) prowadzą do ich bezpośredniego uczenia się.

FAKT: Uczenie dużych modeli językowych w czasie rzeczywistym, na podstawie każdej interakcji, jest obecnie niepraktyczne i zbyt kosztowne obliczeniowo. Modele AI są trenowane na specjalnie wyselekcjonowanych i przygotowanych zestawach danych.

MIT: AI jest nieomylna i zawsze generuje prawdziwe informacje.

FAKT: Modele AI mogą generować informacje fałszywe, nieprecyzyjne lub zniekształcone (tzw. **halucynacje**). Weryfikacja jest niezbędna.

MIT: AI posiada zdolność rozumienia treści w sposób analogiczny do ludzkiego. Może oszukiwać, może się denerwować, może się zniechęcać.

FAKT: Działanie AI opiera się na analizie wzorców statystycznych, bez głębokiego rozumienia kontekstu i znaczenia generowanych treści.

MIT: AI całkowicie zastąpi pracę ludzką.

FAKT: AI będzie przede wszystkim narzędziem wspomagającym pracę człowieka. W wielu dziedzinach **kreatywność (creativity)**, **empatia (empathy)** i **kompetencje społeczne (social competencies)** pozostaną kluczowe i niezastąpione.

MIT: Dzieła generowane przez AI są chronione prawem autorskim.

FAKT: Według stanu na dzień dzisiejszy, wytwory AI powinny być uznawane za element **domeny publicznej (public domain)**, nieobjęte ochroną

Wtórne Zagrożenia

Upředzenia (Bias): Model AI może odzwierciedlać, a nawet wzmacniać **ukryte upředzenia (bias)** obecne w danych treningowych.

Dezinformacja (Disinformation): Istnieje ryzyko wykorzystania AI do generowania i rozpowszechniania **fałszywych informacji (false information)** na szeroką skalę.

Bezpieczeństwo (Security): Należy brać pod uwagę możliwość użycia AI w **cyberatakach (cyberattacks)** lub innych działaniach o charakterze przestępczym (criminal activities).

Appendix

A. Definicja "majaczenia" AI:

Termin *hallucinations* - "majaczenie" (**halucynacje, zmyślanie**) w kontekście AI odnosi się do generowania przez model informacji niezgodnych z prawdą, nielogicznych lub zmyślonych. Wynika to z faktu, iż AI uczy się poprzez identyfikację wzorców statystycznych w danych, nie zaś poprzez rozumienie treści w sposób właściwy człowiekowi. - [https://en.wikipedia.org/wiki/Hallucination_\(artificial_intelligence\)](https://en.wikipedia.org/wiki/Hallucination_(artificial_intelligence))

B. Ograniczenie Czasowe Wiedzy AI:

Modele językowe AI, takie jak ChatGPT, są trenowane na obszernych zbiorach danych, aktualnych do określonego momentu (**knowledge cutoff**). Na przykład dla modelu ChatGPT moment ten przypada na **grudzień 2023**. Po tej dacie model nie ma dostępu do nowszych informacji - <https://chatgpt.com/share/676dc9d1-4e3c-8013-b112-0ae37b11cb87>

C. Wpływ na Środowisko:

Korzystanie z AI wiąże się ze znacznym zużyciem energii i emisją CO₂. Należy dążyć do optymalizacji wykorzystania zasobów obliczeniowych. **Źródło:** <https://smartly.ai/blog/the-carbon-footprint-of-chatgpt-how-much-co2-does-a-query-generate> - *The carbon footprint of ChatGPT – how much CO₂ does a query generate?*

D. Prawa Autorskie:

Aktualnie, wytwory AI powinny być uznawane za **domenę publiczną (public domain)**, czyli być nieobjęte prawami autorskimi (without copyright). **Źródło:** <https://akademia.zaiks.org.pl/wiedza/czy-ai-moze-byc-autorem-copy> - *Czy AI może być autorem? | Akademia ZAIKS*

E. Zasoby Dodatkowe (Linki):

Analiza problemu majaczenia AI: <https://arxiv.org/pdf/2401.11817> - "Hallucination is Inevitable: An Innate Limitation of Large Language Models"

Artykuł dot. wpływu AI na edukację: <https://www.nature.com/articles/s41599-023-01787-8> - *Nature*, 9 Cze 2023

Osiągnięcia AI w dziedzinie matematyki: <https://deepmind.google/discover/blog/ai-solves-imo-problems-at-silver-medal-level/> - *DeepMind*, 25 Lip 2024

Przykład zastosowania AI do koloryzacji wideo: <https://neural.love/blog/sota-video-colorization> - *Neural Love*

Wywiad z Demisem Hassabisem (Google DeepMind):

<https://www.cbsnews.com/news/google-artificial-intelligence-future-60-minutes-transcript-2023-06-11/> - CBS News

Raport AI Index 2024 (Stanford University): <https://aiindex.stanford.edu/report/>

Mapa MAD (Machine Learning, Artificial Intelligence & Data) 2024: Źródło:
mad.firstmark.com

Zagadnienie Fałszywych Informacji (Fake News) i Wiarygodności:

<https://www.yahoo.com/news/fact-check-real-pic-baby-184500737.html>;

https://www.reddit.com/r/interestingasfuck/comments/1fz72q8/the_human_internet_is_dying_ai_images_taking_over/

Projekt Deep Dream: <https://www.artnome.com/news/2018/12/30/deepdream-creator-unveils-very-first-images-after-three-years>

!!! (Ważne) Narzędzia

 <https://aistudio.google.com/>

 <https://claude.ai/>

 <https://chatgpt.com/>

 <https://runwayml.com/>

 <https://suno.com/>

 <https://www.udio.com/>

 <https://elevenlabs.io/>

 <https://www.midjourney.com/>

 <https://labs.google/>

 <https://notebooklm.google.com/>

Recap

Wykorzystanie AI niesie ze sobą zarówno szanse, jak i zagrożenia. Kluczowe jest świadome i odpowiedzialne korzystanie z tych narzędzi, z uwzględnieniem ich ograniczeń, zagrożeń prawnych oraz potencjalnego wpływu na społeczeństwo i środowisko.

Glosariusz

AI (Artificial Intelligence - Sztuczna Inteligencja / Inteligencja Maszynowa): Dziedzina informatyki zajmująca się tworzeniem systemów komputerowych zdolnych do wykonywania zadań, które typowo wymagają ludzkiej inteligencji, takich jak uczenie się, rozumienie języka, rozpoznawanie obrazów czy rozwiązywanie problemów.

GenAI (Generative AI - Generatywna Sztuczna Inteligencja): Rodzaj AI, który koncentruje się na tworzeniu nowych treści, takich jak tekst, obrazy, muzyka, kod, itp., na podstawie wzorców i danych, na których został wytrenowany.

LLM (Large Language Model - Duży Model Językowy): Zaawansowany model AI, zazwyczaj oparty na architekturze Transformer, wytrenowany na ogromnych zbiorach danych tekstowych, zdolny do rozumienia i generowania języka naturalnego na wysokim poziomie. Przykłady: GPT, LLaMA, PaLM.

RAG (Retrieval-Augmented Generation): Technika, która łączy możliwości generatywnych modeli językowych (np. LLM) z zewnętrznymi bazami wiedzy. Pozwala to na generowanie bardziej precyzyjnych i wiarygodnych odpowiedzi, opartych na aktualnych i sprawdzonych informacjach.

Diffusion (Dyfuzja): W kontekście generatywnego AI, proces stopniowego dodawania szumu do danych (np. obrazu), a następnie odwrócenie tego procesu w celu wygenerowania nowych, podobnych danych. Modele dyfuzyjne są obecnie szeroko stosowane do tworzenia obrazów o wysokiej jakości.

Prompt: Tekstowe polecenie, instrukcja lub pytanie wprowadzane do modelu AI (np. LLM) w celu uzyskania określonej odpowiedzi lub wygenerowania konkretnego rodzaju treści.

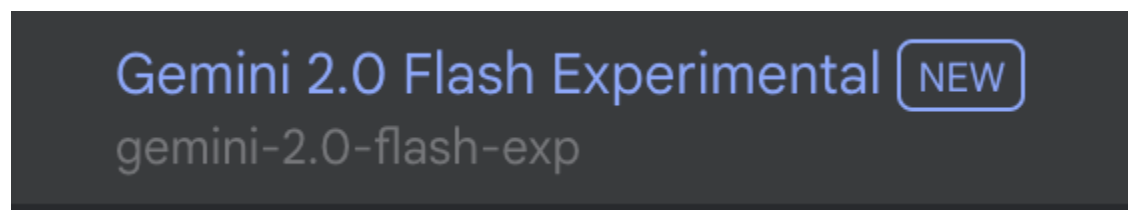
Token: Podstawowa jednostka informacji przetwarzana przez modele językowe. W zależności od modelu, token może być pojedynczym znakiem, sylabą, całym słowem lub fragmentem słowa.

Fine-tuning (Dostrajanie): Proces dalszego uczenia wytrenowanego modelu AI na mniejszym, bardziej specyficznym zbiorze danych w celu dostosowania go do konkretnego zadania lub dziedziny.

Agent (w kontekście systemów AI): Autonomiczny podmiot programowy, wykorzystujący algorytmy sztucznej inteligencji, zdolny do wykonywania złożonych zadań w określonym środowisku. Agent charakteryzuje się percepcją, reaktywnością, proaktywnością oraz zdolnością do uczenia się i adaptacji. Działa niezależnie, podejmując decyzje na podstawie własnych celów i analizy otoczenia, może komunikować się z innymi systemami lub ludźmi, dążąc do osiągnięcia wyznaczonych zadań. Agent AI znajduje zastosowanie w różnorodnych dziedzinach, od wirtualnych asystentów po systemy automatyzacji, a w kontekście kultury, może być narzędziem do tworzenia nowych form ekspresji oraz ułatwiania dostępu do wiedzy.

Ray Kurzweil – Osobliwość Coraz Bliżej – warto przeczytania wprowadzenie do technologii i futurologia - <https://lubimyczytac.pl/ksiazka/5168014/nadchodzi-osobliwosc-kiedy-czlowiek-przekroczy-granice-biologii>

Podsumowanie Pytań i Odpowiedzi Poszkoleniowe



Temat Spotkania: Sztuczna Inteligencja w Kulturze: Szanse i Zagrożenia

Cel Spotkania: Przedstawienie uczestnikom kompleksowego przeglądu aktualnych zagadnień związanych z rozwojem i implementacją sztucznej inteligencji (AI) w kontekście kulturalnym, w tym:

- Omówienie technicznych aspektów działania AI (sieci neuronowe, uczenie maszynowe).
- Analiza potencjalnych korzyści i zagrożeń wynikających z zastosowania AI w kulturze.
- Prezentacja praktycznych narzędzi i przykładów wykorzystania AI.
- Dyskusja nad etycznymi i prawnymi aspektami generowania treści przez AI.
- Podkreślenie roli instytucji kultury w kontekście dynamicznego rozwoju AI.

Kluczowe Punkty Spotkania:

1. Wprowadzenie do zagadnienia:

- Definicja i omówienie pojęć "inteligencja maszynowa" i "uczenie maszynowe".
- Przykłady wizualizacji sieci neuronowych i ich zastosowania w kulturze.
- Wprowadzenie do prezentacji i jej struktury: od zagrożeń, przez szanse, do warsztatu.



2. Techniczne aspekty działania AI:

- Wyjaśnienie działania sieci neuronowych i ich roli w generowaniu treści.
- Omówienie procesu uczenia maszynowego i jego wpływu na jakość generowanych treści.
- Przykłady narzędzi opartych na AI (np. Google DeepMind, ChatGPT, Midjourney).

3. Zagrożenia związane z AI:

- **Hacking i manipulacja:** Możliwość wykorzystania AI do celów nieetycznych, w tym manipulacja informacjami i oszukiwanie systemów.
- **Halucynacje:** Skłonność AI do tworzenia nieprawdziwych i niespójnych treści (tzw. "zmyślanie"), zacieranie granic między prawdą a fikcją.
- **Erozja wiarygodności:** Wpływ generowanych przez AI treści na obniżenie zaufania do źródeł informacji, w tym wizualnych.
- **Utrata krytycznego myślenia:** Potencjalne osłabienie zdolności analitycznych i decyzyjnych użytkowników AI.
- **Upředzenia i nierówności:** Problem upředzeń ukrytych w algorytmach i systemach AI, potencjalnie prowadzących do nieuczciwych wyników.

4. Szanse związane z AI:

- **Nowe możliwości artystyczne:** AI jako narzędzie do tworzenia nowych form ekspresji i interakcji z kulturą.
- **Ułatwienie dostępu do treści kultury:** AI jako pomoc w tworzeniu audiodeskrypcji, tłumaczeń i innych form adaptacji.
- **Edukacja przez praktykę:** Wykorzystanie AI do wspomagania procesu nauczania i zdobywania wiedzy.
- **Automatyzacja żmudnych procesów:** Odciążenie twórców i edukatorów z powtarzalnych zadań.
- **Prace badawcze:** Wykorzystanie AI jako wsparcia do badań.

5. Kwestie etyczne i prawne:

- **Autentyczność i wiarygodność:** Konieczność weryfikacji treści generowanych przez AI.
- **Prawa autorskie:** Problematyka własności intelektualnej dzieł stworzonych przy wsparciu AI.
- **Ochrona danych:** Zagadnienie prywatności i bezpieczeństwa danych w kontekście systemów AI.
- **Wpływ na środowisko:** Emisja CO2 związana z obliczeniami AI.

6. Rola instytucji kultury:

- Potencjalna rola jako gwaranta wiarygodności informacji w kontekście generowanych treści.
- Tworzenie standardów etycznych i zasad korzystania z AI.

- Wspieranie rozwoju kompetencji w zakresie wykorzystania AI w kulturze.

Zadania i wnioski:

1. **Dalsze badania i eksploracja:** Kontynuacja analizy wpływu AI na kulturę i sztukę.
2. **Weryfikacja źródeł informacji:** Konieczność krytycznej oceny treści, w tym obrazów, tekstu, dźwięku generowanych przez AI.
3. **Rozwój kompetencji cyfrowych:** Udoskonalenie umiejętności wykorzystania narzędzi AI w kontekście kulturalnym.
4. **Opracowanie etycznych wytycznych:** Wdrożenie zasad uczciwego i odpowiedzialnego korzystania z AI w kulturze.
5. **Współpraca interdyscyplinarna:** Nawiązanie dialogu między ekspertami technicznymi, artystami, instytucjami kultury i prawnikami.

Materiały Dodatkowe:

- Prezentacja udostępniona uczestnikom spotkania.
- Linki do stron i materiałów referencyjnych (zostały przekazane uczestnikom).
- Dostarczenie uczestnikom dokumentu do wykorzystania po spotkaniu.

Podsumowanie:

Spotkanie miało na celu dostarczenie uczestnikom wiedzy i narzędzi niezbędnych do zrozumienia i efektywnego wykorzystania sztucznej inteligencji w kontekście kultury. Podkreślono zarówno szanse, jak i potencjalne zagrożenia związane z tą technologią, wskazując na konieczność etycznego i odpowiedzialnego podejścia do jej implementacji. Uczestnicy zostali zachęcani do dalszej eksploracji i aktywnego kształtowania przyszłości AI w obszarze kultury.

<https://suno.com/song/cded40ea-53f6-4c64-9850-5f3f21e797a8>

<https://suno.com/song/1e1aae68-4721-40c8-bce4-0b1fbe0ca5f9>

<https://suno.com/song/201a8ce9-d170-4add-958a-9e57a77cfdbf>

<https://suno.com/song/b1c08b26-5dce-4ff1-b4fc-022d6fb9f60a>