Współpraca międzynarodowa w oparciu o godną zaufania Sztuczną Inteligencję

Newsletter sekcji – wydanie 3.

kwiecień 2025

Ministerstwo Cyfryzacji | Tu tworzymy przyszłość

GRAI GRUPA ROBOCZA DS. SZTUCZNEJ INTELIGENCJI

Od redakcji

Szanowni Państwo,
Drodzy Przyjaciele i Przyjaciółki Sekcji,

To już trzecie wydanie newslettera naszej Sekcji ds. współpracy międzynarodowej w oparciu o godną zaufania sztuczną inteligencję. Dzięki wsparciu Ministerstwa Cyfryzacji oraz zaangażowaniu Grupy Roboczej ds. Sztucznej Inteligencji (GRAI) tworzymy publikację, która ma na celu wspieranie wymiany wiedzy i budowanie międzynarodowego dialogu w tej dynamicznie rozwijającej się dziedzinie.

To wyjątkowe wydanie – w związku z zainteresowaniem innych sekcji naszej Grupy Roboczej, zaprosiliśmy do udziału w naszych pracach nowych autorów i autorki. Witamy na pokładzie i dziękujemy za wartościowy wkład!

W niniejszym numerze znajdziecie 7 tekstów autorów i autorek. Mamy nadzieję, że prezentowane tu teksty przybliżą naszym czytelnikom i czytelniczkom nowe zagadnienia, dzięki którym jeszcze lepiej będziemy w stanie poruszać się w świecie sztucznej inteligencji i innych pokrewnych zaawansowanych technologii.

Ponownie zachęcamy do włączania się w nasze prace wszystkich członków i członkinie GRAI – wierzymy, że właśnie tego typu zaangażowanie może przynosić satysfakcję i prowadzić do prawdziwej synergii działań[1].

Michał Jackowski
Zuzanna Karcz

[1] Newsletter przygotowali eksperci działający na zasadach pro publico bono w Grupie Roboczej ds. Sztucznej Inteligencji (GRAI) przy Ministerstwie Cyfryzacj w ramach sekcji Współpraca międzynarodowa w oparciu o godną zaufania Sztuczną Inteligencję.

Projekt ustawy o ochronie małoletnich przed dostępem do treści szkodliwych w internecie. Przykłady określenia treści szkodliwych i metod weryfikacji wieku w innych krajach.

Mariusz Krzysztofek

Projekt ustawy o ochronie małoletnich przed dostępem do treści szkodliwych w internecie opracowany przez Ministerstwo Cyfryzacji i opublikowany 24.02.2025 na stronie RCL, wpisuje się w analogiczną międzynarodową tendencję. Podobne ustawy są m.in. we Francji, Niemczech, Wielkiej Brytanii, Australii, Chinach i Stanach Zjednoczonych. Choć różnie w różnych ustawodawstwach postrzegane są „treści szkodliwe w internecie", przed którymi małoletni mają być chronieni. Różne też mogą być metody weryfikacji wieku użytkownika.

1. Zakres pojęcia „treści szkodliwych" w projekcie MC i w innych krajach

Projekt Ministerstwa Cyfryzacji, jak wskazuje jego uzasadnienie, koncentruje się na treściach pornograficznych. Zostało to poparte badaniami, w tym raportem NASK pt. „Nastolatki wobec pornografii cyfrowej. Ale ani tytuł ustawy ani taka definicja treści szkodliwych nie wykluczają objęcia zakresem ustawy innych treści, jak patostreaming.

Warto dla porównania spojrzeć na dwa inne podejścia – australijskie i chińskie – do ochrony dzieci w internecie, których wspólnym mianownikiem z pozostałymi jest wprawdzie ochrona dzieci przed pornografią (np. amerykańska federalna COPPA przez treści szkodliwe rozumie „prurient interest", do których należą treści pornograficzne), ale które rozszerzają tę ochronę o dwa interesujące czynniki: Australia – o media społecznościowe, Chiny - o gry internetowe. Australia uchwaliła (w listopadzie 2024 r.) ustawę zakazującą osobom do 16. roku życia korzystania z mediów społecznościowych, co więcej – przy aprobacie społecznej, bo zgodnie z sondażami zakaz popierało aż 77% obywateli.

Chiński ustawodawca położył szczególny nacisk na gry online. Ponadto w reakcji na plagę uzależnienia od gier online wśród dzieci znowelizowane przepisy ustawy z 1.10.2019 r. o ochronie danych osobowych dzieci w Internecie oraz przepisy wykonawcze Krajowego Urzędu Prasy i Publikacji wprowadzają limit czasowy grania w Internecie przez dzieci poniżej 18 roku życia dla każdego dnia powszedniego i w weekendy. Chińska Administracja Cyberprzestrzeni (Cyberspace Administration of China, CAC) opublikowała 15.11.2024 r. wytyczne dotyczące mobilnego Internetu dla nieletnich („Guidelines for the Construction of Mobile Internet Mode for Minors"), które określają funkcjonalność urządzeń mobilnych i aplikacji, aby umożliwić egzekwowanie tego dziennego limitu.

2. Metody weryfikacji wieku w projekcie MC i w innych krajach

Projekt ustawy Ministerstwa Cyfryzacji zakłada obowiązek weryfikacja wieku, aby wykluczać dostęp małoletnich użytkowników do szkodliwych treści w internecie. Sama ustawa nie będzie określać mechanizmów weryfikacji, lecz zobowiązuje do ich określenia Ministra Cyfryzacji, w porozumieniu z Prezesem Urzędu Komunikacji Elektronicznej. Mają one zapewniać skuteczności weryfikacji wieku, a także uwzględniać ochronę danych osobowych i możliwości techniczne usługodawców (ale też brać pod uwagę ryzyko obejścia zakazu przez użycie VPN).

Zgodnie z wytycznymi Europejskiej Rady Ochrony Danych rozwiązania służące tej weryfikacji „powinny być proporcjonalne do charakteru czynności przetwarzania i związanego z nimi ryzyka”. „Weryfikacja wieku nie powinna prowadzić do nadmiernego przetwarzania danych”, więc „w niektórych sytuacjach niskiego ryzyka właściwe może być zobowiązanie nowego abonenta usługi do ujawnienia roku urodzenia lub wypełnienia formularza, w którym oświadczy, że (nie) jest osobą małoletnią”. Ale ze względu na szkodliwy charakter treści, których dotyczy ustawa taka weryfikacja nie będzie tu wystarczająca.

Ministerstwo rozważa, bo takie metody są powszechne w wielu krajach, m.in. weryfikację przy pomocy logowania do bankowości elektronicznej, karty kredytowej, portfeli tożsamości cyfrowej, ale wybór metody będzie należał do operatora.

Metody weryfikacji wieku w różnych krajach są zróżnicowane i mogą służyć jako inspiracja dla lokalnych wytycznych w powiązaniu z innymi metodami i z uwzględnieniem ryzyka.

Użytecznym przykładem jest opracowany 11.10.2024 r. przez francuski Urząd Regulacji Audiowizualnych i Komunikacji Cyfrowej (Autorité de régulation de la communication audiovisuelle et numérique) standard weryfikacji wieku użytkowników stron pornograficznych, który został opracowany dla treści wymagających szczególnie starannej kontroli dostępu. W dokumencie zwrócono uwagę m.in. na odporność weryfikacji wieku na manipulacje, np. deepfake’i i spoofing, a także na mechanizm „podwójnej anonimowości”, przedstawiony wcześniej w rekomendacji CNIL, który zapewnia, że witryna nie może gromadzić danych użytkownika podczas korzystania z sytemu do weryfikacji wieku w przyszłości.

Ale intersujące jest też spojrzenie na inne przykłady.

W niektórych serwisach wykorzystywana jest sztuczna inteligencja (np. na podstawie selfie” przesłanego telefonem) lub sprawdzanie wieku na innych powiązanych kontach. To metody,

które mają być alternatywami wobec wymagania przesyłania skanu dokumentu tożsamości, który zgodnie z wytycznymi wielu regulatorów jest naruszeniem zasady minimalizacji oraz jest postrzegany przez wielu użytkowników jako ingerencja w prywatność, zresztą nie wszystkie dzieci mają dokumenty tożsamości. Trzeba tu również wyciągnąć wnioski z wyroku, który zablokował wejście w życie kalifornijskiej ustawy CAADCA (United States Court of Appealsfor The Ninth Circuit, 16.08.2024, Netchoice, Llc V. Bonta, No. 23-2969 (9th Cir. 2024). Potwierdzenie tożsamości rodzica lub opiekuna przez wykorzystanie jego adresu e-mail jest metodą wskazaną w amerykańskiej federalnej COPPA. COPPA wskazuje także płatności kartami kredytowymi, które wymagają powiadomienia posiadacza karty o transakcji, w 3D Secure security protocol, lub komunikację z personelem przez bezpłatną infolinię lub wideokonferencję.

Na podstawie 50 najpopularniejszych aplikacji dla zróżnicowanych odbiorców i 10 najpopularniejszych aplikacji dla dorosłych z Tencent App Store (stan na luty 2023 r.) najpowszechniej stosowaną w Chinach metodą weryfikacji wieku okazuje się deklaracja użytkownika, a więc np. wpisanie daty urodzenia i odmowa rejestracji konta użytkownika, jeżeli ten oświadczy, że nie ukończył 14 lat (art. 28 chińskiej PIPL zalicza dane osobowych małoletnich w wieku poniżej 14 lat do katalogu danych wrażliwych). Inną popularną metodą jest wymaganie numeru telefonu komórkowego przy rejestracji konta użytkownika aplikacji.

Sztuczna inteligencja w zarządzaniu miastem: co sądzą mieszkańcy Warszawy?

Anna Domaradzka, Mateusz Trochymiak
Zespół Civil City Lab, Instytut Studiów Społecznych, Uniwersytet Warszawski

Sfinansowano w ramach projektu: "Prawo do inteligentnego miasta: wpływ nowych technologii na jakość życia, relacje społeczne i politykę miejską" ze środków Narodowego Centrum Nauki (2018/30/E/HS6/00379)

Projekt "Prawo do inteligentnego miasta" to seria badań dotyczących społecznych i psychologicznych konsekwencji wdrażania nowoczesnych technologii w kontekście "prawa do miasta". Celem projektu jest zbadanie konsekwencji rozwoju technologicznego w miastach z perspektywy mieszkańców i mieszkanek. W ramach projektu powstał raport pt. „Stosunek do nowych technologii w zarządzaniu miastem".

Henri Lefebvre, autor idei prawa do miasta podkreślał, że wszyscy mieszkańcy miast powinni mieć prawo uczestniczyć w decyzjach kształtujących ich środowisko życia. W obliczu przemian technologicznych hasło prawa do miasta przypomina o tym, że to obywatele

i obywatelki są głównymi użytkownikami miasta, usług dla mieszkańców i adresatami miejskich polityk.

Nasze badanie miało na celu weryfikację stosunku mieszkańców Warszawy do stosowania sztucznej inteligencji w procesie zarządzania miastem. Zapytaliśmy warszawiaków o postawy dotyczące sztucznej inteligencji, jej roli w procesach decyzyjnych i zagrożeniach z tym związanych.

Wyniki pokazują, że chociaż mieszkańcy Warszawy wykazują się optymizmem, jeżeli chodzi o AI i nowe technologie, chętniej akceptują rozwiązania ukierunkowane na rozwiązywanie konkretnych problemów i wybranych grup, niż takie, które są dedykowane wszystkim mieszkańcom. Pewna niechęć do rozwiązań takich jak chatboty czy autonomiczne autobusy wynika prawdopodobnie z obawy przed osłabieniem lub zastąpieniem relacji z człowiekiem. Sztuczna inteligencja jest postrzegana jako rozwiązanie poprawiające jakość decyzji miejskich, nie powinna jednak prowadzić do zastąpienia pracowników miejskich.

Opracowaliśmy wskaźnik akceptacji AI w usługach miejskich. Wypracowany wskaźnik jest składową zmiennych z pytań dotyczących tego, czy mieszkańcy dopuszczają, aby AI decydowała o wsparciu socjalnym, inwestycjach w infrastrukturę, siatce połączeń komunikacji publicznej czy interwencjach służb porządkowych. Wynik 47.8 na skali 1-100 wskazuje na bardzo ostrożną postawę warszawiaków w stosunku do zastosowania AI w zarządzaniu miastem. Podobną, chociaż jeszcze niższą wartość nasz wskaźnik osiągnął w Tallinnie (40.6), z kolei dla Singapuru przyjął wartość aż 62.5.

Poziom akceptacji rozwiązań opartych o AI był stosunkowo podobny między kobietami i mężczyznami, w różnych grupach wiekowych czy mających różny poziom kompetencji cyfrowych. Wyjątkiem była samoocena sytuacji materialnej badanych. Grupa osób deklarujących złą lub bardzo złą sytuację ekonomiczną jest wyraźnie mniej chętna zaakceptować rozwiązania oparte o AI. Z kolei osoby w dobrej sytuacji materialnej wykazywały większą akceptacją dla technologii niż pozostali.

Wyniki wskazują na akceptację roli AI jako wsparcia w podejmowaniu przez urząd miasta decyzji technicznych i równocześnie niechęć warszawiaków do stosowania AI w sprawach, które wymagają czynnika ludzkiego. Poziom akceptacji jest wyższy w przypadku spraw, w których oczekuje się przede wszystkim decyzji obiektywnych, opartych o racjonalne przesłanki, tj. planowania siatki połączeń komunikacji, kierowania patrolami policji czy zamykania obszarów o wysokim poziomie zanieczyszczenia powietrza. Akceptacja dla AI jest mniejsza w odniesieniu do kwestii postrzeganych jako złożone, wymagającymi osiągnięcia społecznego kompromisu czy emocjonalnie angażującymi, tj. inwestycji

w infrastrukturę, prawa do demonstracji, prawa do zasiłków, czy zastąpienia urzędników miejskich przez algorytm sztucznej inteligencji.

Badanie zrealizowano w 2022 roku na reprezentatywnej próbie 700 dorosłych warszawiaków.

Propozycje polskich ekspertów dotyczące Kodeksu Praktyk dla dostawców modeli GPAI

Piotr Brzyski

W kontekście trwających prac nad wdrażaniem unijnego Aktu o Sztucznej Inteligencji (AI Act), w marcu 2025 roku grupa polskich ekspertów AI opublikowała dokument prezentujący stanowisko w sprawie Kodeksu Praktyk dla modeli AI ogólnego przeznaczenia (GPAI). Dokument pojawił się krótko po opublikowaniu trzeciego projektu Kodeksu, którego ostateczna wersja ma zostać przyjęta do 2 maja 2025 roku. Dokument z treścią kodeksu został opublikowany online.

Dokument polskich ekspertów proponuje dwa kluczowe podejścia do Kodeksu Praktyk:

Adaptacyjna taksonomia zarządzania ryzykiem AI

Eksperci postulują, by wytyczne regulacyjne obejmowały tylko te rodzaje ryzyka, dla których istnieją już sprawdzone środki techniczne lub zarządcze. Adaptacyjna taksonomia oznacza elastyczny system klasyfikacji ryzyka, który powinien być regularnie aktualizowany, odzwierciedlając postęp w badaniach nad AI i praktykach branżowych. Nowe kategorie ryzyka miałyby zastosowanie tylko do modeli GPAI wprowadzonych lub istotnie zaktualizowanych po dodaniu tych kategorii do taksonomii. Podejście to ma na celu utrzymanie wymogów regulacyjnych w granicach możliwych do praktycznego wdrożenia przez dostawców.

Projektowanie z myślą o godnej zaufania i odpowiedzialnej AI (TRAI)

Druga propozycja dotyczy promowania wbudowania mechanizmów odpowiedzialności i wiarygodności w systemy AI od początku ich rozwoju. TRAI (Trustworthy, Responsible AI by Design) to podejście, zgodnie z którym zamiast koncentrować się na ocenach i poprawkach po wdrożeniu, należy uwzględniać kwestie etyczne i środki łagodzące ryzyko już w procesie projektowania. Koncepcja TRAI obejmuje zarządzanie ryzykiem w całym cyklu życia systemu AI, wzorce architektoniczne zapewniające przejrzystość i wyjaśnialność, metodologie zarządzania kompromisami w rozwoju AI (np. między dokładnością a uczciwością) oraz wytyczne uwzględniające wpływ systemów AI na ludzi.

Według autorów dokumentu, takie proaktywne podejście może skutecznie rozwiązać problemy takie jak pułapki związane z wyjaśnialnością czy tzw. ciemne wzorce (manipulacyjne techniki projektowe znane jako "dark patterns"), koncentrując się na rozwiązaniach projektowych zachęcających do krytycznego myślenia.

Potencjalne skutki propozycji

Przyjęcie proponowanych zasad mogłoby przekształcić Kodeks Praktyk w bardziej elastyczne ramy regulacyjne, które ewoluują wraz z rozwojem technologii AI. Stanowisko przewiduje Kodeks wspierający dynamiczne zarządzanie ryzykiem, umożliwiający dostawcom AI wykazanie zgodności poprzez różne podejścia branżowe i kładący nacisk na środki zapobiegawcze zamiast reaktywnego zarządzania ryzykiem. Podejście to mogłoby również ułatwić zgodność z międzynarodowymi ramami bezpieczeństwa i zarządzania AI, zapewniając odzwierciedlenie globalnych najlepszych praktyk w europejskich regulacjach.

Dokument został poparty przez Polską Izbę Informatyki i Telekomunikacji oraz podpisany przez grupę polskich ekspertów AI, w tym naukowców, inżynierów i prawników. Odnosi się do istniejących ram i narzędzi, takich jak ramy Wdrażania Godnej Zaufania AI OECD, ramy Samooceny Godnej Zaufania AI Komisji Europejskiej oraz zasoby z Globalnego Partnerstwa na rzecz AI.

Stanowisko polskich ekspertów koncentruje się na zapewnieniu, by regulacje AI były zarówno skuteczne w minimalizowaniu ryzyka, jak i praktyczne we wdrażaniu. Podkreśla potrzebę równowagi między ochroną przed potencjalnymi szkodami a wspieraniem innowacji w dziedzinie AI. Opublikowany dokument stanowi wkład w trwający europejski dyskurs na temat kształtu przyszłych regulacji AI i może wpłynąć na ostateczną formę Kodeksu Praktyk GPAI.

„Ghiblifikacja” internetu: czy AI zagraża twórcom?

Natalia Lener-Bobek

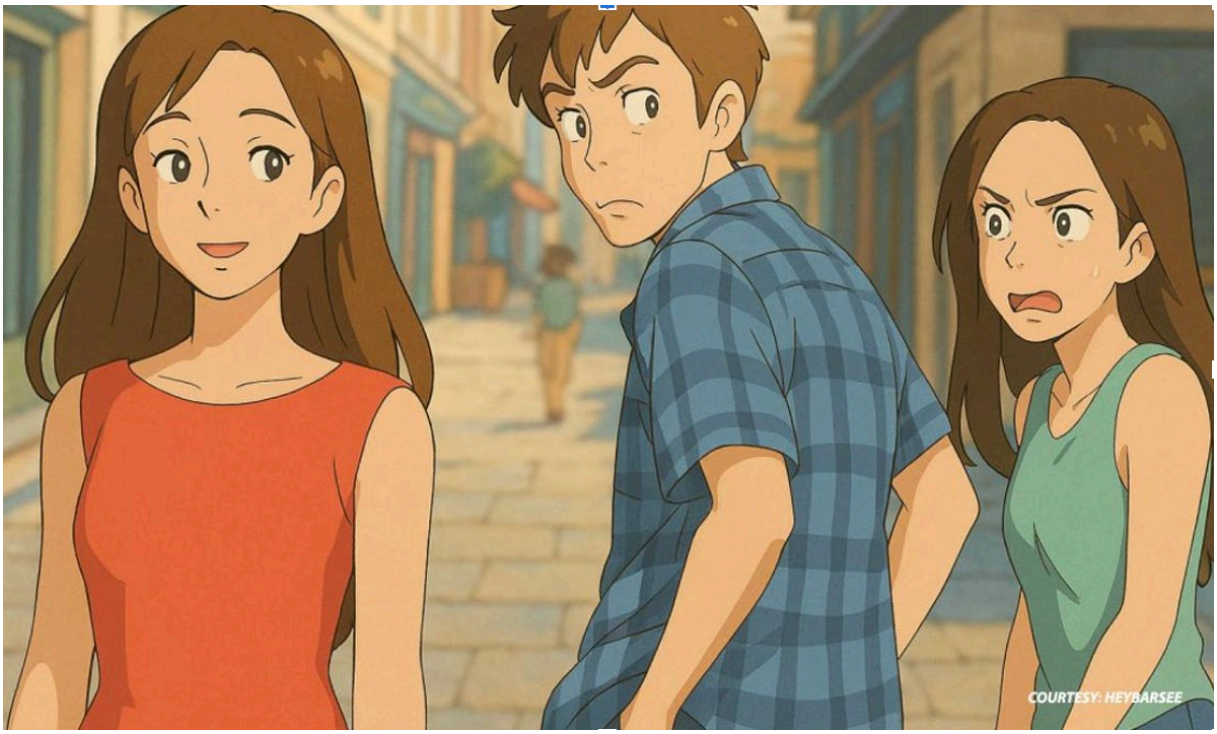
W ostatnich dniach internet zalała fala grafik inspirowanych kultowym stylem Studia Ghibli.

Jak to możliwe? Wszystko dzięki nowemu generatorowi obrazów od OpenAI, który pojawił się w ramach aktualizacji GPT-4o. To narzędzie bez trudu tworzy wizualizacje przypominające te z filmów Ghibli.

Studio Ghibli to japońska wytwórnia znana z ręcznie rysowanych animacji, pełnych detali i emocjonalnie poruszających historii. Ich filmy powstają często latami i mają niepowtarzalny charakter.

Gdy sam Altman przerobił swoje zdjęcie na styl Ghibli, trend błyskawicznie stał się viralem.

Można więc z dużym prawdopodobieństwem założyć, że model GPT-4o był trenowany na grafikach Studia Ghibli.



Należy jednak zadać sobie pytanie, czy takie działanie jest legalne i etyczne? Czy OpenAI zawarł jakąś umowę ze studiem Ghibli? A może to kolejna historia o tym, jak OpenAI „luźno” traktuje kwestie praw autorskich, wychodząc z podstawowego założenia bigtechów, że lepiej przepraszać niż prosić o pozwolenie?

Co to wszystko oznacza?

Temat praw autorskich i danych, na których trenowane są narzędzia AI, będzie wracał jak bumerang. Debata na temat etycznego rozwoju AI jest dziś pilniejsza niż kiedykolwiek. Czy fakt, że generator obrazów potrafi naśladować styl konkretnego artysty, jest naruszeniem prawa autorskiego? W Unii Europejskiej sam styl nie jest chroniony, ale wszyscy czujemy, że nie tędy droga do etycznej AI. Gdzie więc przebiega granica?

Obecne przepisy nie nadążają za technologią i nie są wystarczające do ochrony twórczości artystów która jest zasysana przez sztuczną inteligencję.

Naśladowanie stylu Ghibli mogłoby podpadać pod naruszenie dóbr osobistych twórców czy czyn nieuczciwej konkurencji.

Wrzucając w generator ChataGPT zdjęcia nasze i naszych bliskich dostarczamy je do OpenAI udzielając tym samym rodowskiej zgody na ich dalsze wykorzystanie. Efekt Ghibli to tylko wierzchołek góry lodowej, jeśli chodzi o kwestie prawne związane z wykorzystaniem sztucznej inteligencji i jej wpływu na prawa autorskie i RODO. Z całą pewnością ten temat nie pozwoli nam się nudzić!

Regulacje generatywnego AI w grach video

Mateusz Hyla

Sztuczna inteligencja generująca treści znacząco przyspiesza proces produkcji gier wideo. Pozwala tworzyć większe i bardziej rozbudowane produkcje przy znacznie niższych kosztach. Tego rodzaju automatyzacja bywa jednak krytykowana przez graczy, dlatego niektórzy producenci starają się ukrywać informacje o wykorzystaniu AI.

Na ten moment nie istnieją przepisy regulujące tę praktykę. Spośród trzech najistotniejszych międzynarodowych platform sprzedaży gier, jedynie Steam przyjął restrykcyjną politykę wobec treści generowanych przez sztuczną inteligencję. Początkowo, jeśli podczas procesu zatwierdzania gry wykryto elementy stworzone przez AI, tytuł był odrzucany. Twórcy musieli wtedy usunąć wszystkie tego typu materiały.

Na początku 2024 roku Steam złagodził te zasady. Wprowadzono ankietę, w której deweloperzy zobowiązani są poinformować o wykorzystaniu generatywnego AI. Twórcy muszą określić, czy gra zawiera wcześniej wygenerowane treści — takie jak grafiki, kod, dźwięki, obrazy, itp. — oraz potwierdzić, że zostały one stworzone zgodnie z obowiązującym prawem i nie naruszają praw innych podmiotów. Dodatkowo, jeśli gra generuje treści na żywo, deweloperzy muszą wskazać zastosowane zabezpieczenia, które zapobiegają tworzeniu treści nielegalnych lub nieetycznych.

Podsumowanie tych informacji jest wyświetlane w widocznym miejscu na stronie sprzedaży gry — pod materiałami marketingowymi, nad sekcją recenzji i wymaganiami systemowymi. Dzięki temu użytkownicy mogą świadomie podjąć decyzję o zakupie.

Steam udostępnił również narzędzie dla graczy, umożliwiające zgłaszanie tytułów, które ukrywają obecność treści generowanych przez AI lub wykorzystują je niezgodnie z prawem. Po weryfikacji zgłoszenia platforma może:

- nakazać deweloperowi uzupełnienie ankiety,
- nakazać usunięcie nielegalnych materiałów,
- zawiesić sprzedaż gry,
- usunąć grę na stałe z platformy,
- zakończyć współpracę z twórcą (w przypadku recydywy).

Epic oraz GOG

Największa konkurencja Steam – chińska platforma Epic Games Store i polska GOG jak dotąd nie ogłosiły oficjalnych regulacji dotyczących treści generowanych przez AI. Należy jednak zwrócić uwagę, że proces weryfikacji gier bywa tam bardziej rozbudowany.
Na oficjalnych forach pojawiają się głosy graczy domagających się wprowadzenia podobnych zasad jak na Steam, ale jak dotąd nie spotkało się to z reakcją ze strony tych platform.

Rozwiązania wprowadzone przez Steam są obecnie najpełniejsze. Dają równowagę między ochroną interesów graczy a swobodą twórczą deweloperów. Wciąż jednak brakuje precyzyjnych wytycznych dla innych form wykorzystania AI, na przykład wyraźnego oznaczania przeciwników sterowanych przez sztuczną inteligencję w grach wieloosobowych.

Analogiczna procedura informowania o wykorzystywaniu sztucznej inteligencji oraz weryfikacji praktyk, mogłaby być stosowana w przypadku wszystkich programów komputerowych.

Oficjalne oświadczenie Steam o wprowadzeniu systemu informowania o wykorzystywaniu AI

Przykładowy apel do GOG o oznaczanie gier korzystających z AI

Zjednoczone Emiraty Arabskie liderem w dziedzinie sztucznej inteligencji do 2031 roku

Lidia Sobańska

Zjednoczone Emiraty Arabskie (ZEA) mają ambitną wizję, aby stać się jednym z wiodących krajów w dziedzinie sztucznej inteligencji (AI) do 2031 roku. Kraj ten aktywnie realizuje swoją strategię, inwestując w ludzi, technologie i sektory kluczowe dla osiągnięcia tego celu.

Strategiczne cele i priorytety

Strategia ZEA w dziedzinie AI opiera się na ośmiu strategicznych celach, które stanowią fundament dla rozwoju i wdrażania AI w kraju. Cele te obejmują:

- Budowanie reputacji ZEA jako globalnego centrum AI, przyciągającego talenty i inwestycje z całego świata.
- Zwiększenie konkurencyjności ZEA w kluczowych sektorach, takich jak energia, logistyka, turystyka, opieka zdrowotna i cyberbezpieczeństwo, poprzez wdrażanie innowacyjnych rozwiązań AI.
- Rozwój ekosystemu wspierającego innowacje i rozwój AI, obejmującego finansowanie, mentoring i wymianę wiedzy.
- Wykorzystanie AI do ulepszania usług publicznych i jakości życia obywateli, poprzez rozwiązywanie kluczowych wyzwań społecznych i ekonomicznych.
- Przyciąganie i szkolenie talentów w dziedzinie AI, zarówno lokalnych, jak i międzynarodowych, aby zapewnić odpowiednią kadrę dla rozwijającego się sektora.
- Wspieranie badań naukowych i współpracy między sektorem akademickim a przemysłem, aby stymulować innowacje i transfer technologii.
- Zapewnienie dostępu do danych i infrastruktury niezbędnej do testowania i wdrażania AI, przy jednoczesnym zapewnieniu bezpieczeństwa i prywatności danych.
- Wdrożenie skutecznych ram zarządzania i regulacji AI, które będą promować etyczne i odpowiedzialne wykorzystanie tej technologii.

Współpraca międzynarodowa i wymiana wiedzy

ZEA zdają sobie sprawę z kluczowej roli współpracy międzynarodowej w rozwoju AI. Kraj ten aktywnie współpracuje z międzynarodowymi ekspertami, firmami i instytucjami w celu wymiany wiedzy i doświadczeń w dziedzinie AI. ZEA organizują również międzynarodowe konferencje i fora, aby stać się globalnym centrum dyskusji i innowacji w obszarze AI.

Inwestycje w kluczowe sektory

ZEA koncentrują się na wdrażaniu AI w pięciu kluczowych sektorach, w których technologia ta ma potencjał przynieść znaczące korzyści gospodarcze i społeczne. Sektory te to:

- **Zasoby i energia:** ZEA, jako znaczący eksporter ropy naftowej, wykorzystują AI do optymalizacji procesów wydobycia i zarządzania zasobami. Kraj ten inwestuje również w AI w celu rozwoju energii odnawialnej, inteligentnych sieci energetycznych i efektywnego gospodarowania wodą.
- **Logistyka i transport:** ZEA, będąc ważnym węzłem komunikacyjnym, wdrażają AI w celu automatyzacji zarządzania ruchem lotniczym i morskim, obsługi bagażu, optymalizacji łańcuchów dostaw i rozwoju autonomicznych pojazdów.
- **Turystyka i gościnność:** ZEA wykorzystują AI do personalizacji usług turystycznych, przewidywania potrzeb klientów, automatyzacji obsługi klienta i tworzenia inteligentnych asystentów dla turystów.
- **Opieka zdrowotna:** ZEA inwestują w AI w celu poprawy diagnostyki chorób, personalizacji leczenia, rozwoju badań genetycznych i farmaceutycznych oraz optymalizacji zarządzania placówkami medycznymi.
- **Cyberbezpieczeństwo:** ZEA, w obliczu rosnących zagrożeń cybernetycznych, rozwijają AI w celu ochrony infrastruktury krytycznej, wykrywania i zapobiegania atakom, a także budowania odporności na cyberprzestępczość.

Rozwój talentów i edukacja

ZEA zdają sobie sprawę z kluczowej roli edukacji i rozwoju talentów w zapewnieniu zrównoważonego rozwoju AI. Kraj ten inwestuje w programy szkoleniowe, kursy i inicjatywy edukacyjne, mające na celu podnoszenie świadomości na temat AI, rozwijanie umiejętności cyfrowych i kształcenie specjalistów w dziedzinie AI.

Wizja na przyszłość

ZEA mają ambitną wizję, aby stać się światowym liderem w dziedzinie AI. Kraj ten dąży do stworzenia ekosystemu, który przyciągnie talenty, inwestycje i innowacje, a także do wykorzystania AI do rozwiązywania kluczowych wyzwań społecznych i gospodarczych, poprawy jakości życia obywateli i budowania zrównoważonej przyszłości.

Więcej w raporcie “UAE NATIONAL STRATEGY FOR ARTIFICIAL INTELLIGENCE 2031”

Cyfrowe bliźniaki w ochronie zdrowia: wirtualne modele, realne rezultaty

Jagoda Miszewska

Rozwój cyfrowych bliźniaków nabiera coraz większego tempa, przekształcając je w kluczowe narzędzie modernizacji systemów opieki zdrowotnej. Technologia ta, wykorzystująca sztuczną inteligencję oraz integrację danych w czasie rzeczywistym, umożliwia zaawansowane symulacje procesów szpitalnych lub personalizację terapii na niespotykaną dotąd skalę.

W sektorze ochrony zdrowia funkcjonują dwa zasadnicze typy cyfrowych bliźniaków. **Cyfrowy bliźniak szpitala** stanowi wirtualną replikę procesów operacyjnych placówki, obejmującą przepływ pacjentów, alokację łóżek oraz harmonogramy zabiegów. System ten integruje dane z elektronicznej dokumentacji medycznej (EHR), czujników IoT oraz rejestrów szpitalnych. Jego podstawowym celem jest optymalizacja zarządzania zasobami, redukcja kolejek oraz poprawa efektywności pracy personelu. Przykładem skutecznej implementacji jest Szpital Mater w Dublinie, gdzie cyfrowy bliźniak oddziału radiologii, opracowany we współpracy z Siemens Healthineers, doprowadził do skrócenia czasu oczekiwania na badania MRI o 33% oraz zwiększenia wykorzystania tomografów z 68% do 89% poprzez optymalizację harmonogramów.

Równolegle rozwija się koncepcja **cyfrowego bliźniaka pacjenta**, który stanowi dynamiczny model zdrowia pojedynczej osoby, uwzględniający dane genetyczne, wyniki badań diagnostycznych oraz informacje z urządzeń noszonych na ciele. Służy on do symulacji reakcji organizmu na leki, prognozowania progresji chorób oraz personalizacji schematów terapeutycznych. Francuski projekt Meditwin, rozwijany przez Dassault Systèmes we współpracy z FDA, umożliwia tworzenie wirtualnych modeli narządów pacjentów. Technologia ta przyspieszyła procesy testów klinicznych dla leków kardiologicznych o 40% oraz zredukowała częstotliwość występowania błędów terapeutycznych dzięki możliwości weryfikacji skuteczności terapii w środowisku wirtualnym (in silico). Innym z potwierdzonych przykładów implementacji tej technologii jest projekt realizowany przez firmę Phesi, koncentrujący się na przewlekłej chorobie przeszczep przeciwko gospodarzowi (cGvHD). Zespół badawczy Phesi skonstruował precyzyjny cyfrowy odpowiednik grupy pacjentów poddawanych standardowej terapii prednizonem, stanowiącej konwencjonalną metodę leczenia w grupach kontrolnych u pacjentów po transplantacji. Opracowanie to umożliwiło zastąpienie tradycyjnych kohort kontrolnych symulacjami cyfrowymi, co znacząco przyspieszyło proces ewaluacji nowych terapii. W obszarze badań nad chorobą Alzheimera

zastosowanie metodologii cyfrowych bliźniaków przyniosło równie imponujące rezultaty. Implementacja tej technologii umożliwiła redukcję czasu trwania badań klinicznych z typowego okresu 18-24 miesięcy do zaledwie 12 miesięcy. Wykorzystanie cyfrowych bliźniaków pozwoliło na precyzyjne modelowanie progresji procesów neurodegeneracyjnych, z obserwacją spadku funkcji poznawczych na poziomie 20-30% w krótszym horyzoncie czasowym, co przełożyło się na istotną optymalizację efektywności oraz redukcję kosztów badań.

Cyfrowe bliźniaki to już nie futurystyczna wizja, lecz realne narzędzia transformujące ochronę zdrowia. Wiele placówek medycznych na świecie już aktywnie eksploruje możliwości wykorzystania tej metodologii w zarządzaniu chorobami przewlekłymi, jak i procedurami szpitalnymi. Rynek cyfrowych bliźniaków w ochronie zdrowia ma osiągnąć wartość 3,5 miliarda dolarów do końca 2025 roku, choć ponad 50% organizacji wskazuje integrację danych jako główną barierę we wdrażaniu tej technologii. Polska, dysponująca znaczącym potencjałem innowacyjnym w sektorze medtech, ma realną szansę na aktywny udział w tym dynamicznie rozwijającym się rynku, wykorzystując swoje zaplecze naukowe oraz technologiczne do budowania konkurencyjnych rozwiązań na skalę międzynarodową.

Sekcja ds. współpracy międzynarodowej w oparciu o godną zaufania sztuczną inteligencję

Grupa Robocza ds. Sztucznej Inteligencji

Sekcja działa w ramach Grupy Roboczej ds. Sztucznej Inteligencji (GRAI) przy Ministerstwie Cyfryzacji. Jej głównym celem jest wspieranie międzynarodowej współpracy na rzecz odpowiedzialnego rozwoju AI. Dążymy do budowania pomostów pomiędzy administracją publiczną, biznesem i środowiskiem naukowym, aby tworzyć warunki dla rozwoju sztucznej inteligencji zgodnej z zasadami etyki, cyberbezpieczeństwa oraz praw człowieka.

Pośród naszych kluczowych działań znalazły się:

- **Konsultacje prezentowanych przez Ministerstwo Cyfryzacji polityk** – bierzemy aktywny udział w opiniowaniu dokumentów strategicznych oraz tworzenie spójnych stanowisk w imieniu grupy.
- **Katalog ekspercki** to baza specjalistów gotowych wspierać rozwój AI w Polsce, w tym w obszarze współpracy z MŚP i administracją.
- **Newsletter specjalistyczny** to regularne publikacje prezentujące analizy, komentarze i wydarzenia związane z AI na poziomie lokalnym

Sekcja organizuje także cykliczne spotkania, które umożliwiają wymianę doświadczeń i integrację środowiska. Dzięki wspólnemu zaangażowaniu tworzymy przestrzeń dla dyskusji, które kształtują przyszłość AI jako technologii transparentnej, godnej zaufania i dostępnej dla wszystkich.

Członkowie i członkinie naszej sekcji działają na zasadach *pro publico bono,* przyczyniając się do dynamicznego rozwoju tej kluczowej dziedziny technologicznej na arenie międzynarodowej. Dołącz do nas i współtwórz przyszłość sztucznej inteligencji!

Osoby zainteresowane dołączeniem do GRAI i naszej sekcji, współtworzenia opracowań eksperckich na zasadach pro bono prosimy o przesłanie zgłoszenia na adres mailowy: grai@cyfra.gov.pl.

Dowiedz się więcej na stronie GRAI.