Stanowisko
Izby Wydawców Prasy oraz Stowarzyszenia Dziennikarzy i Wydawców Repropol
do projektu „Polityki rozwoju sztucznej inteligencji w Polsce do 2030 roku"

I. Wstęp

Projekt „Polityki rozwoju sztucznej inteligencji w Polsce do 2030 r." stanowi ważny krok w kierunku kształtowania przyszłości cyfrowej Polski. Dokument w sposób kompleksowy omawia szanse i wyzwania związane z rozwojem AI, odwołując się do standardów międzynarodowych i europejskich. Pozwalamy sobie jednak na uwagę, że przedstawiona do konsultacji polityka nie dostrzega w wystarczającym stopniu jednego z kluczowych wyzwań współczesnej polityki innowacyjności: potrzeby zapewnienia **symetrycznego, sprawiedliwego i zrównoważonego rozwoju AI**, który nie prowadzi do koncentracji korzyści w rękach technologicznie uprzywilejowanych podmiotów kosztem twórców czy producentów dóbr kultury.

Poniżej prezentujemy zasadnicze zagadnienia, które stanowią podstawę dla sformułowania przez nas konkretnych postulatów uzupełniających projekt przedłożonej do konsultacji polityki.

Fałszywe przeciwstawienie rozwoju technologicznego oraz ochrony kultury

W szczególności brakuje nam jednoznacznego zakwestionowania **fałszywej narracji opozycji między rozwojem technologicznym a ochroną „ludzkiej" kultury**. Takie przeciwstawienie ignoruje fakt, że to właśnie **bogactwo kultury jest fundamentami odpowiedzialnego rozwoju technologii**, w tym systemów opartych na sztucznej inteligencji. Jest tak zarówno w wymiarze metaforycznym, jak i dosłownym: wieloletnie, bezumowne korzystanie z utworów pozyskanych z sieci internetowej jest źródłem wykładniczego postępu narzędzie generatywnej AI. Polityka powinna tymczasem wyraźnie uznać, że **rozwój AI musi iść w parze z ochroną i wsparciem dla ludzkiej kreatywności oraz poszanowaniem podmiotowości twórców**.

Z zadowoleniem należy natomiast odnotować, że projekt odnosi się do międzynarodowych standardów, takich jak wytyczne OECD z 2019 r., które promują rozwój AI z poszanowaniem praw człowieka i wartości demokratycznych, w tym zasad sprawiedliwości, przejrzystości, prywatności i wyjaśnialności działania AI. Również rekomendacje UNESCO w sprawie etyki AI z 2021 r. jednoznacznie wskazują, że ochrona godności ludzkiej i praw człowieka powinna być fundamentem wszelkich działań w zakresie rozwoju AI, co znajduje wyraz w postulatach wzmocnionego nadzoru ludzkiego (*human oversight*) oraz transparentności ich działania. Wartości te leżą również u podstaw europejskiego Aktu w sprawie sztucznej inteligencji (AI Act).

W świetle powyższych standardów polityka koniecznie wymaga **wzmocnienia o postulaty,** które uwzględniałyby interesy podmiotów kultury, twórców jako równoprawnych uczestników procesu transformacji cyfrowej. Ci ostatni są przecież nie tylko beneficjentami technologii, lecz także jej źródłem.

AI jako narzędzie wspierające człowieka

Uważamy, że sztuczna inteligencja powinna być narzędziem wspierającym człowieka, nigdy zaś jego substytutem. Systemy AI muszą pozostawać pod nadzorem i kontrolą człowieka, zwłaszcza w obszarach wymagających odpowiedzialności oraz zgodności z podstawowymi zasadami etycznymi. Z uznaniem odnotowujemy, że projekt Polityki AI odwołuje się do idei *ethics by design* oraz zapowiada ramy regulacyjne sprzyjające innowacjom, a zarazem zgodne z podstawowymi wartościami demokratycznymi, takimi jak godność osoby ludzkiej, prawa człowieka czy uczciwa konkurencja. W naszej ocenie wszelkie systemy AI wdrażane w Polsce powinny funkcjonować wyłącznie dla dobra człowieka i podlegać jego skutecznej kontroli. Takie podejście jest spójne z międzynarodowymi standardami oraz opiniami środowisk eksperckich.

Zagadnieniem pilnie wymagającym doprecyzowania pozostaje natomiast ochrona twórczości ludzkiej w kontekście generatywnej AI. Rozdział 6.5 projektu trafnie wskazuje, że treści generowane przez AI nie podlegają ochronie prawa autorskiego z uwagi na brak wkładu człowieka. Podzielamy tę ocenę i opowiadamy się za jej utrzymaniem. Zwracamy jednocześnie uwagę, że masowa produkcja syntetycznych treści rodzi szereg ryzyk zarówno dla pozycji ekonomicznej twórców, jak i jakości informacji oraz zaufania do przekazów funkcjonujących w przestrzeni publicznej.

Warto, by projekt polityki wyraźniej uwzględniał te wyzwania i wzmacniał ochronę utworów ludzkich oraz pozycję.

Poszanowanie zjawiska twórczości

Zagadnieniem najistotniejszym dla środowiska, które reprezentujemy, jest kwestia wykorzystania utworów do trenowania modeli AI. Pozostaje poza sporem, że dziejąca się na naszych oczach rewolucja **nie byłaby możliwa bez wykorzystania ogromnych zasobów cudzych utworów** (dziennikarskich, literackich, artystycznych, audiowizualnych, etc.) które są masowo kopiowane bez zgody czy wiedzy twórców. Taka sytuacja jest nadzwyczaj alarmująca: twórcy obawiają się, że ich praca stała się zwykłym paliwem dla modeli AI – bez ich kontroli, bez jakiegokolwiek wynagrodzenia.

Bodaj najlepszym komentarzem do całego zestawu problemów powstających na styku modeli AI oraz prawa autorskiego jest najnowszy raport pt. „*Generative AI & Copyright. Balancing creative rights, legal integrity, and accountability in the AI age*" opracowany na zlecenie Komisji Prawnej Parlamentu Europejskiego (JURI). Raport powinien być oficjalnie udostępniony w czerwcu bądź lipcu br. To opracowanie stanowi kompleksową analizę skutków rozwoju generatywnej sztucznej inteligencji dla prawa autorskiego Unii Europejskiej. Raport poddaje krytycznej ocenie obowiązujące ramy regulacyjne – w szczególności dyrektywę 2019/790 (dyrektywa o prawie autorskim a jednolitym rynku cyfrowym, CSDM) – wskazując na jej nieprzystawalność do wyzwań wynikających z masowego i komercyjnego wykorzystania utworów chronionych w procesie trenowania modeli generatywnej AI. Głównymi wnioskami wynikającymi z raportu są:

a) dyrektywa CDSM nie była projektowana z myślą o specyfice i skali działania generatywnej AI, co czyni wyjątek TDM nieadekwatnym wobec współczesnych modeli generatywnej AI,

b) modele generatywnej AI są trenowane na ogromnych zbiorach zawierających chronione utwory bez zgody ani wynagrodzenia dla uprawnionych,

c) AI generuje treści wypierające twórczość pochodzącą od człowieka z rynku, co zagraża istnieniu całych zawodów kreatywnych,

d) trenowanie AI wiąże się z techniczną reprodukcją utworów po etapie trenowania, co budzi oczywiste wątpliwości w świetle prawa autorskiego z uwagi na ograniczenie wyjątku eksploracji tekstów i danych do etapu trenowania modeli AI; modele AI nie ograniczają się do pozyskania wiedzy faktograficznej, lecz kodują w sobie samą zawartość utworów, co narusza sam rdzeń ochrony prawa autorskiego,

e) przewidziany w art. 4 CDSM wyjątek dla komercyjnego TDM nie może usprawiedliwiać trenowania generatywnych modeli AI, a rozszerzającą wykładnia wyjątku TDM jest błędna, a ponadto sprzeczna z tzw. testem trzystopniowym (*three-step-test*) wynikającym konwencji berneńskiej,

f) mechanizm opt-out jest nieskuteczny wobec masowego pozyskiwania utworów z sieci internetowej (web-scraping) i braku wypracowania wspólnego standardu technicznego dla takiego zastrzeżenia,

g) brak systemu wynagradzania twórców za wykorzystanie ich utworów podczas trenowania AI pogłębia tzw. lukę wartości (*value gap*, czyli sytuacja, w której olbrzymia wartość ekonomiczna generowana przez modele AI na podstawie cudzej twórczości paradoksalnie nie staje się w żaden sposób udziałem twórców),

wreszcie

h) zaniechanie reform w zakresie prawa autorskiego oraz reżimów pokrewnych grozi nie tylko dalszą koncentracją rynku AI, ale również marginalizacją twórców.

Nieakceptowalny free-riding: komercyjny TDM

Obecny porządek prawny niestety nie gwarantuje ochrony twórczości.

Dyrektywa CDSM wprowadziła wyjątki od ochrony prawnoautorskiej w postaci eksploracji tekstów i danych (*text and data mining, TDM***)**, które zezwalają na zwielokrotnianie utworów do trenowania algorytmów AI, wszakże jedynie w pewnych sytuacjach. Art. 4 dyrektywy CDSM zezwala na eksplorację tekstów i danych dla celów komercyjnych na legalnie dostępnych treściach, o ile twórca nie zastrzegł wyraźnie swoich praw (mechanizm *opt-out*). Do powszechnej wiedzy należy jednak, że w praktyce autorzy nie mają często nawet świadomości, że jego dzieło zostało wykorzystane w treningu AI, gdyż korporacje tworzące modele AI po prostu nie ujawniają źródeł danych treningowych. Ten oczywisty deficyt transparentności po stronie dostawców technologii podważa pożądany balans pomiędzy prawami twórców a interesem firm technologicznych w promowaniu innowacji, którą dyrektywa DSM miała przecież zachować. W efekcie autorzy dowiadują się już po fakcie, że ich dzieła zasilały systemy AI. Rodzi o gigantyczne poczucie niesprawiedliwości wynikające ze świadomości budowania mechanizmu biznesowego w oparciu o cudzy dorobek intelektualny (*free-riding*). Proces pozyskiwania utworów z sieci internetowej trwa od wielu lat. Środowisko twórców oraz producentów dóbr kultury oczekuje zatem nie tylko ustanowienia efektywnego mechanizmu wynagrodzenia zorientowanego na przyszłość, ale także **rzetelnego rozliczenia szkody wynikłej w przeszłości z powodu bezumownego korzystania z cudzej twórczości**.

W świetle powyższego nie mogą dziwić coraz częściej podejmowane w środowiskach twórców akcje protestacyjne, a w niektórych jurysdykcjach wytaczane procesy prawne z tego tytułu. Przykładem

konkretnych działań jest Hiszpania, gdzie zaproponowano mechanizm licencjonowania zbiorowego z rozszerzonym skutkiem (*ECL*) dla trenowania AI, tak aby twórcy mieli zapewnione wynagrodzenie, a masowe korzystanie z danych odbywało się legalnie. To pierwsza tego rodzaju inicjatywa w Europie, możliwa dzięki przepisom dyrektywy DSM (art. 12) pozwalającym na tworzenie licencji zbiorowych w sytuacjach, gdy indywidualne umowy z tysiącami autorów są niewykonalne.

Podsumowanie

Podsumowując, chociaż przedłożony projekt dostrzega wiele z powyższych wyzwań i sygnalizuje ogólnie kierunek rozwiązań, w niniejszym stanowisku proponujemy doprecyzowanie i wzmocnienie brzmienia polityki w obszarach: etyki AI, nadzoru człowieka, prawa autorskiego i interesu twórców.

II. Uzupełnienia polityki AI

Mając na uwadze powyższe postulaty wnosimy o dokonanie następujących uzupełnień dokumentu.

1. Wzmocnienie zasady *human-centric AI* (nadzór człowieka nad AI)

Zarówno standardy wypracowane na forum międzynarodowym – w szczególności UNESCO i OECD – jak i rozporządzenie w sprawie sztucznej inteligencji (AI Act), podkreślają konieczność umiejscowienia człowieka oraz podstawowych wartości społecznych w centrum procesów projektowania i wdrażania systemów AI. W naszym przekonaniu wyzwania związane z rozwojem sztucznej inteligencji wymagają jednoznacznej deklaracji zasady: **to człowiek, a nie algorytm, ponosi ostateczną odpowiedzialność za decyzje podejmowane z udziałem AI**. Zasada ta powinna zostać wprost sformułowana w treści polityki AI jako ogólna motyw przewodni lub jako część rozdziałów dotyczących wizji oraz celów polityki AI. Jej jednoznaczne zapisanie wzmocni przekaz, że Polska zobowiązuje się do rozwoju **sztucznej inteligencji godnej zaufania i zorientowanej na człowieka (*human-centric and trustworthy AI*)**.

Niejako w połączeniu z powyższym zwracamy uwagę, że narodowa kultura i oparta o potęgę ludzkiego umysłu kreatywność stanowią fundament tożsamości narodowej, a jednocześnie istotny sektor gospodarki. W kontekście rozwoju AI konieczne jest objęcie tych obszarów szczególną troską i rozwijanie ich **równolegle, a nie alternatywnie** względem postępu technologicznego. Polityka powinna w sposób zdecydowanie wyraźniejszy zaakcentować, że **wspieranie twórców i rozwój technologii nie są celami sprzecznymi**, lecz wzajemnie się uzupełniającymi. Wsparcie dla AI powinno iść w parze ze wsparciem dla kultury – w sposób **symetryczny i zorientowany na dobro człowieka**.

Sztuczna inteligencja może z powodzeniem wspierać procesy twórcze – np. poprzez automatyzację zadań technicznych lub personalizację narzędzi – lecz nie może to prowadzić do **redukcji roli człowieka jako podmiotu twórczego** ani do obniżenia wartości utworów ludzkich w przestrzeni publicznej. Zasada ta znajduje wyraz m.in. w dokumentach UNESCO dotyczących różnorodności kulturowej w erze cyfrowej, gdzie podkreśla się znaczenie ochrony wkładu ludzkiego w komunikacji i produkcji symbolicznej.

W tym kontekście postulujemy dopisanie w polityce, np. w rozdziale „Wizja i cele Polityki AI" lub w rozdziale 6.2:

„*Systemy AI wdrażane w Polsce muszą być projektowane i stosowane w sposób zapewniający pełny nadzór człowieka oraz możliwość ingerencji w ich działanie, w szczególności w obszarach wiążących się z wysokim ryzykiem. Rozwój sztucznej inteligencji w Polsce powinien służyć wzmacnianiu potencjału ludzkiej kreatywności i ekspresji, nie ich zastępowaniu. Wdrażanie AI w sektorze kultury, edukacji i mediach musi odbywać się z poszanowaniem godności pracy twórczej*".

2. Ochrona praw autorskich w dobie AI – konieczne doprecyzowanie działań

W rozdziale 6.5 polityki AI opisano kwestię ochrony praw autorskich i interesów twórców. Żałujemy jednak, że jedynie w tak krótki sposób odniesiono się do tej fundamentalnej kwestii. W dokumencie wskazano na braku potrzeby ochrony dla treści syntetycznych, z czym zgadzamy się w pełni. Z zadowoleniem odnotowaliśmy wyraźne zaakcentowanie poszkodowania kręgów twórczych poprzez dowolne korzystanie z ich zasobów („*Obecna rewolucja w AI nie byłaby możliwa bez korzystania z zasobów kultury, przy czym istnieje duże prawdopodobieństwo, że w wielu przypadkach odbyło się z to bez odpowiedniej podstawy prawnej. Pogodzenie zapewnienia innowatorom dostępu do wysokiej jakości danych umożliwiających trenowanie modeli powinno iść w parze z ochroną praw autorskich i interesów, które za nimi stoją*, s. 57). **Fragment ten świadczy o pełnej świadomości naszego państwa w zakresie krzywdy dziejącej się podmiotom uprawnionym.**

Choć doceniamy to spostrzeżenie, oczekiwalibyśmy konsekwentnego rozwiązania zasygnalizowanego wątku tej swoistej grabieży własności intelektualnej. Dlatego też proponujemy uzupełnić politykę o konkretne kierunki zmian legislacyjnych i systemowych wzmacniających pozycję twórców wobec rozwoju AI.

Po pierwsze, Polska powinna aktywnie forsować stanowisko, wedle którego **eksploracja tekstów i danych nie powinna mieć zastosowania do trenowania modeli generatywnej AI.** W związku w tym nasz kraj powinien zabiegać na forum UE o przyjęcie rozwiązań przywracającym twórcom prawo do decydowania o wykorzystaniu ich utworów do trenowania AI. Samodzielne decydowanie podmiotów autorsko uprawnionych wobec ich utworów stanowi przecież fundament majątkowych praw autorskich. Oczekiwalibyśmy poparcia dla zmian w prawie autorskim (na poziomie unijnym i krajowym) dla zasady, wedle której **korzystanie z utworu powinno być możliwe wyłącznie po udzielonej zgodzie lub uiszczeniu uczciwego wynagrodzenia (*fair remuneration*)**. Takie podejście zyskuje coraz większe poparcie w Europie. Mamy tu na myśli nie tylko jednoznaczny raport JURI, jak i głosy w doktrynie prawa autorskiego (zob. np. D. Flisak, The Holy Grail of Copyright Law: Text and Data Mining. Selected Issues, Zeszyty Naukowe UJ 1/2024, s. 24 - 26, *passim*; T. Dornis, The Training of Generative AI Is Not Text and Data Mining, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4993782).

Po drugie, polityka AI powinna należycie odnotować konieczność pełnego wdrożenia w Polsce obowiązków z AI Act dotyczących praw autorskich. Rozporządzenie nakłada na dostawców generatywnych modeli AI obowiązek posiadania polityki poszanowania praw autorskich (m.in. respektowania zastrzeżeń opt-out) oraz ujawnienia szczegółowego streszczenia danych użytych do treningu narzędzi AI. Innymi słowy, unijne prawo wymusza większą przejrzystość i respektowanie praw autorskich. Już na poziomie polityki Polska powinna zadeklarować wsparcie dla tych zasad i gotowość ich pełnego egzekwowania.

Mając powyższe na uwadze w rozdziale 6.5 sugerujemy dodanie akapitu:

„Polska opowiada się za rozwiązaniami prawnymi gwarantującymi twórcom kontrolę nad wykorzystaniem ich utworów w procesie trenowania sztucznej inteligencji oraz zapewniającymi godziwe wynagrodzenie w przypadku takiego wykorzystania”.

Sugerujemy także wsparcie stanowiska Polski następującym fragmentem:

„Polska podejmie działania mające na celu wzmocnienie pozycji twórców, tak aby ich prawo do wyrażenia zgody lub sprzeciwu wobec wykorzystania utworu w ramach trenowania AI było realnie wykonywane, a w przypadku wykorzystania utworów do uczenia AI – by twórcy uzyskiwali sprawiedliwe wynagrodzenie za takie wykorzystanie.”

3. Instrumenty proceduralne ułatwiające egzekwowanie praw

Powszechnie wiadomo, że nawet najlepiej napisane prawo nie spełni swojej roli, jeśli twórcy nie będą mogli go wyegzekwować w praktyce. Dlatego w polityce konieczne jest wskazanie kierunku wzmocnienia pozycji prawnej twórców poprzez wprowadzenie odpowiednich instrumentów proceduralnych. Z naszego punktu widzenia kluczowe są następujące elementy:

a) **odwrócenie ciężaru dowodu**, jeżeli wygenerowany *output* AI jest identyczny lub podobny do utworu określonego twórcy, to na dostawcy narzędzia powinien spoczywać obowiązek wykazania, że jego produkt nie był trenowany na utworze danego autora. Z tą konstrukcją wiąże się **domniemanie korzystania z utworu** określonego autora, gdy znajduje się w zbiorach treningowych.

   Tego rodzaju mechanizmy faktycznego domniemania są powszechne w prawie i wynikają z logicznego zakazu obciążania strony ciężarem dowodowym, którego przeprowadzenie jest po prostu niemożliwe. W kontekście AI takie mechanizmy złagodziłyby asymetrię informacyjną: twórca z reguły nie ma dostępu do tego, co działo się we „wnętrzu” modelu AI podczas gdy dostawca ma – bądź powinien mieć - pełnię wiedzy w tym zakresie.

b) wprowadzenie na poziomie unijnym **mechanizmu skutecznego arbitrażu** opartego na obowiązku złożenia przez dostawcę generatywnego modelu AI oferty odzwierciedlającej rynkową wartość udostępniających danych treningowych. Dotychczasowa praktyka kontaktów między podmiotami uprawnionymi – w tym wydawcami, organizacjami reprezentującymi twórców – a globalnymi dostawcami generatywnych modeli AI pokazuje jednoznacznie, że nie istnieje realna możliwość nawiązania równoprawnego dialogu, a tym bardziej wynegocjowania uczciwych warunków wykorzystania utworów w procesie trenowania modeli. Próby wystosowywania zapytań, inicjowania rozmów lub proponowania stawek za licencjonowanie danych kończą się albo milczeniem, albo ofertami rażąco nieodpowiadającymi rzeczywistej wartości ekonomicznej udostępnianych treści.

   Ta skrajna asymetria siły negocjacyjnej, wynikająca z globalnej pozycji platform technologicznych, przewagi kapitałowej oraz braku realnego obowiązku podejmowania negocjacji, prowadzi do strukturalnej nierównowagi pomiędzy twórcami a dostawcami AI. Wobec tego wsparcie państwa w procesie dochodzenia uzasadnionych roszczeń i uczestnictwa w negocjacjach – zarówno na poziomie krajowym, jak i unijnym – jest nie tylko pożądane, ale i konieczne.

W tym kontekście **uzasadniony jest postulat wprowadzenia na poziomie unijnym mechanizmu skutecznego arbitrażu**, który stanowiłby odpowiedź na nieefektywność obecnych prób negocjacyjnych. Postulowany system powinien opierać się na:

- obowiązku po stronie dostawcy generatywnego modelu AI złożenia konkretnej oferty wynagrodzenia za korzystanie z danych objętych prawem autorskim (np. tekstów, obrazów, nagrań) w procesie trenowania modelu,
- ocenie tej oferty przez niezależny organ arbitrażowy z udziałem przedstawicieli twórców, organizacji zbiorowego zarządzania i regulatorów,
- możliwości ustalenia stawki odzwierciedlającej realną wartość rynkową danych,
- obowiązkowym charakterze decyzji arbitrażowej

c) **Kolejnym filarem odpowiedzialnego wdrażania generatywnej sztucznej inteligencji powinno być zapewnienie przejrzystości zarówno na etapie trenowania modeli, jak i na etapie generowania treści.** Oznacza to po pierwsze obowiązek ujawniania wykorzystywanych zbiorów danych treningowych (w szczególności źródeł pochodzenia treści), a po drugie – jednoznaczne oznaczanie treści syntetycznych w sposób umożliwiający ich odróżnienie od materiałów pochodzących od ludzi.

W odniesieniu do etapu trenowania, Polska powinna zadeklarować wolę aktywnego monitorowania realizacji obowiązku przejrzystości źródeł, o którym mowa w art. 53 ust. 1 lit. d AI Act. Zobowiązanie to ma szczególne znaczenie z perspektywy ochrony praw autorskich, rzetelności źródeł oraz kontroli nad wykorzystywaniem dorobku twórców w procesach treningowych.

W odniesieniu do etapu generowania treści (outputu), polityka publiczna powinna jednoznacznie zobowiązywać państwo do monitorowania przestrzegania obowiązku oznaczania treści syntetycznych, jak przewiduje art. 50 AI Act. Odbiorcy muszą być w sposób jasny i zrozumiały informowani, że mają do czynienia z materiałem wygenerowanym przez sztuczną inteligencję.

Niezbędne jest również uregulowanie kwestii autorstwa lub afiliacji wygenerowanych treści – przynajmniej poprzez wskazanie pochodzenia materiału oraz informacji o wykorzystanych źródłach. Takie oznaczenia są istotne z punktu widzenia ochrony integralności informacji, praw twórców oraz zaufania społecznego do treści funkcjonujących w przestrzeni publicznej.

d) W odniesieniu do powyższych rekomendacji o charakterze procesowym proponujemy następujące uzupełnienie polityki AI w rozdziale traktujących o ochronie praw autorskich i interesów twórców:

- domniemanie korzystania z utworu i odwrócenie ciężaru dowodu

„W przypadku istotnego podobieństwa między utworem a treścią wygenerowaną przez system AI, uznaje się domniemanie, że utwór został wykorzystany w procesie trenowania. Ciężar dowodu wykazania braku takiego wykorzystania spoczywa na dostawcy modelu AI."

- mechanizm unijnego arbitrażu w sprawach o wynagrodzenie

„Polska poprze na poziomie unijnym wprowadzenie mechanizmu arbitrażowego w sprawach wynagrodzenia za wykorzystanie utworów podczas trenowania modeli AI, opartego na obowiązku złożenia przez dostawcę oferty odzwierciedlającej rynkową wartość treści."

- obowiązek przejrzystości źródeł danych treningowych oraz treści syntetycznych

„Polska będzie monitorować wykonanie obowiązku transparentności źródeł danych treningowych generatywnych modeli AI oraz oznaczania treści wygenerowanych przez AI, jak i informowania odbiorców o ich syntetycznym pochodzeniu”

- afiliowanie źródeł syntetycznych treści

„Treści syntetyczne powinny być opatrzone informacją o źródłach materiałów wykorzystanych do ich wygenerowania”

4. Naruszenie równowagi w ekosystemie informacyjnym przez koncerny technologiczne

Tematem, który wymaga odrębnego zasygnalizowania, jest dewastujący wpływ dominujących platform internetowych na ekosystem informacyjny. Dla jasności, **nie tracimy z pola widzenia licznych cywilizacyjnych korzyści wynikających z rozwoju technologicznego technologii sztucznej inteligencji.** Od dawna mocno kontestujemy propagowaną przez koncerny technologiczne narrację, wedle której środowisko wydawców i dziennikarzy miałoby hamować postęp technologiczny wskutek zgłaszanych postulatów. Zwracamy po prostu uwagę na postępującą w bezprecedensowym tempie degradację wartości demokratycznych wynikających z braku regulacyjnych ograniczeń dla globalnych koncernów technologicznych w sferze wykorzystania cudzych treści.

Od ponad dekady koncerny technologiczne przejmują rosnącą część rynku reklamy cyfrowej. **To podmioty agregujące treści i pośredniczące w ich dystrybucji paradoksalnie stały się głównymi beneficjentami korzyści ekonomicznych płynących z wytwarzania** medialnego kontentu – i to bez udziału w samym procesie twórczym. Zjawisko to nastąpiło oczywistym kosztem innych uczestników ekosystemu informacyjnego, w szczególności mediów, wydawców, a wreszcie – samych dziennikarzy. W odróżnieniu jednak od tych ostatnich, platformy internetowe nie zatrudniają dziennikarzy, nie ponoszą kosztów redakcyjnych, ani innych kosztów związanych z wytworzeniem wartościowych informacji. Generatywna sztuczna inteligencja (np. chatboty, inteligentne wyszukiwarki typu Google AI-Overview) jedynie **przyspieszyła proces monetyzowania wyników generowanych na bazie cudzych treści**. Językowe modele AI (LLM) czerpią z treści źródłowych, streszczając, czasami memoryzując, parafrazując, reinterpretując je bez wskazania autorstwa, bez licencjonowania, bez odesłania do źródła.

Podsumowując kluczowe konsekwencje obserwowane już na poziomie ogólnoświatowym obejmują one:

- asymetrię ekonomiczną (*value gap*) polegającą na tym, że koncerny technologiczne czerpią olbrzymie korzyści z pracy wydawców i dziennikarzy, nie uczestnicząc w żaden sposób w kosztach wytworzenia wartościowego kontentu,
- spadek ruchu do treści źródłowych (model tzw. *zero-click AI*): generatywne funkcje wdrażane w wyszukiwarkach (typu AI Overviews) skutkują faktyczny zanikiem kliknięć w hyperlinki do stron wydawców, co stanowi fundamentalne zagrożenie dla modeli biznesowych mediów, szczególnie operujących lokalnie czy poruszających niszowe tematy,
- zwiększenie zagrożenia dezinformacją: znaczna część outputu generatywnych modeli językowych zawiera błędy merytoryczne lub tzw. halucynacje. W połączeniu z naturalnie brzmiącym językiem, stylem imitującym wypowiedzi eksperckie czy brakiem oznaczeń źródła, zjawisko to stanowi realne zagrożenie dla edukacji, demokratycznego dyskursu,

zaufania obywateli do instytucji państwa. Dezinformujące treści wypływające z modeli AI mogą prowadzić do utrwalania fałszywych przekonań, wzmacniania uprzedzeń (*bias*), wzmagania polaryzacji w społeczeństwie, której szczególnie dotkliwie doświadczamy w naszym kraju. Ryzyka te zasługują na szczególną uwagę w newralgicznych obszarach funkcjonowania państwa: w sferze informacyjnej, w środowisku edukacyjnym, w kontekście kampanii wyborczych.

Mając na uwadze powyższe, proponujemy dodanie do rozdział 6 pt. „Sztuczna inteligencja godna zaufania" nowy podrozdział nr 6.6 zatytułowany „*Pluralizm informacyjny i jakość wiedzy w dobie sztucznej inteligencji*"

„Należy zapewnić, że rozwój i wdrażanie sztucznej inteligencji w Polsce, zwłaszcza modeli generatywnych, wspiera dostęp obywateli do wiarygodnych, zróżnicowanych i transparentnych źródeł informacji, nie narusza warunków funkcjonowania niezależnych mediów, instytucji naukowych i edukacyjnych, oraz przeciwdziała rozprzestrzenianiu treści dezinformujących, nieprawdziwych lub zmanipulowanych. Cel ten obejmuje zarówno ochronę niezależnych źródeł informacji i instytucji edukacyjnych, jak i wprowadzenie rozwiązań umożliwiających uczciwy podział korzyści z wykorzystania treści tworzonych przez ludzi w systemach AI. Obejmuje on również rozwój umiejętności krytycznego myślenia i świadomego korzystania z mediów w społeczeństwie oraz wspieranie wysokich standardów jakości, przejrzystości źródeł i odpowiedzialności za treści generowane cyfrowo.

Co umożliwi realizację celu:

• Edukacja medialna i cyfrowa w szkołach i w administracji, w tym rozpoznawanie treści syntetycznych, mechanizmów wpływu informacyjnego oraz źródeł wiedzy publicznej;

• Zapewnienie użytkownikom mechanizmów przekierowania do pełnych treści źródłowych;

• Promowanie systemu zgłaszania i korygowania halucynacji".

III. Podsumowanie

Strategia rozwoju AI powinna nie tylko wyznaczać kierunki technologicznego postępu, ale też dawać czytelny sygnał, że rozwój tej technologii w Polsce będzie przebiegać z poszanowaniem człowieka, jego twórczości i wartości, które budują wspólnotę. W zgłoszonych postulatach chcieliśmy, aby wybrzmiała potrzeba zachowania większej równowagi pomiędzy innowacją a kulturą, między potencjałem technologii a prawami tych, których praca i dorobek są dziś wykorzystywane do trenowania modeli AI. Dlatego też wskazujemy na konieczność większej przejrzystości działania systemów AI, jasnego oznaczania treści tworzonych przez algorytmy i realnej kontroli człowieka nad tym, co AI robi i na jakiej podstawie. Zwracamy też uwagę na ogromną nierówność sił między twórcami a globalnymi firmami technologicznymi oraz powstałą wskutek tej asymetrii potrzebę, by państwo aktywnie włączyło się w wyrównywanie nierównowagi, przede wszystkim na poziomie europejskim.

Tylko wtedy AI będzie naprawdę godna zaufania i rzeczywiście służąca człowiekowi. ■

Warszawa, 25 czerwca 2025 roku