Warszawa, 11 czerwca 2021 r.

Kancelaria Prezesa Rady Ministrów
Departament Rozwiązań
Innowacyjnych

Szanowni Państwo,

Związek Pracodawców Branży Internetowej IAB Polska dziękuje za możliwość wzięcia udziału w konsultacjach publicznych opublikowanego przez Komisję Europejską pakietu projektów przepisów dotyczących sztucznej inteligencji (AI). Poniżej wskazujemy wstępne uwagi dot. kwestii jakie zwróciły naszą uwagę po zapoznaniu się z treścią w/w projektu. Jednocześnie deklarujemy chęć czynnego uczestnictwa w konsultacjach i wsparcie podczas całego procesu legislacyjnego pozostawiając sobie możliwość aktualizacji przedstawionych uwag.

KLUCZOWE PROBLEMY/ OBSZARY DO WYJAŚNIENIA

1. **Zbyt szeroka definicja sztucznej inteligencji:** Jeśli utrzymamy taką definicję, to obawiamy się, że rozszerzająca interpretacja AI może doprowadzić do objęcia nią np. konwencjonalne funkcje informatyczne oparte na formule „jeśli to".

   - Definicja Sztucznej Inteligencji (art. 3 pkt 1) nie ma charakteru uniwersalnego, lecz została powiązana z kazuistycznie wymienionymi "techniques and approaches" w aneksie I. W Rozporządzeniu przyjmuje się (art. 4), iż wskazana lista winna być aktualizowana na bieżąco, ale tego rodzaju technika legislacyjna zupełnie nie odpowiada szybkości zmian technologicznych w zakresie Sztucznej Inteligencji. Wszystko to sprawia to, że przyjęte rozwiązanie całkowicie mija się z założeniami wskazanymi w pkt 6 preambuły, wedle którego pojęcie Sztucznej Inteligencji musi być "clearly defined to ensure legal certainty, while providing the flexibility to accomodate future technological developments".

     W tym zakresie lepszą definicję zaoferowała High Level Expert Group on Artificial Intelligence: "*Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals*".

   - Proste kontrolowane przez człowieka uczenie maszynowe powinno pozostać wyraźnie poza zakresem proponowanej regulacji - nie każdą automatyzację IT należy uznać za AI, bo nie składa się na nią automatyzacja analityki predykcyjnej, a jedynie podstawowe algorytmy.

2. **Wyjaśnienia dotyczące art. 5 (zabronione praktyki sztucznej inteligencji)** w szczególności litery a) i b) - zaproponowany język jest wciąż niejasny – jaka jest definicja technik podprogowych? Co to jest „szkoda psychiczna”? Co oznacza sformułowanie: „wykorzystuje dowolną lukę w zabezpieczeniach określonej grupy osób”?

- Z uwagi na fakt, iż niektóre systemy Sztucznej Inteligencji są całkowicie zakazane, niezwykle istotne jest precyzyjne wytyczenie granic takiego zakazu. Kryteria tego nie spełnia art. 5 ust. 1 pkt a), który zakazuje korzystania z systemów, które mają "podświadomie" ("subliminal") oddziaływać na odbiorców. Rozumienie tego pojęcia może być przyczyną sporów, które w rezultacie mogą zmniejszyć innowacyjność. Z drugiej strony należy także założyć, że każdy kraj unijny posiada odpowiednie przepisy ochronne, w tym chroniące konsumentów, jeśli idzie o "podprogowy" przekaz. Inny problem jest związany z pojęciem "psychological harm"; w tym zakresie należałoby się raczej ograniczyć do rozumienia szkody jako szkody materialnej ("physical harm"). Treść punktu 16 preambuły nie eliminuje wątpliwości w tym zakresie.
- Odwołanie do pojęcia *fundamental rights* w art. 7.1.b jest dość nieostre i może być interpretowane rozszerzająco. Przyjęcie pełnego katalogu praw podstawowych oznaczałoby konieczność stosowania bardzo szerokiej perspektywy (Karta praw podstawowych Unii Europejskiej czy Europejska konwencja o ochronie praw człowieka i podstawowych wolności). **W takiej sytuacji pewność prawa doznaje zauważalnego uszczerbku, pomijając już sam fakt, że w praktyce zagrożenie godności człowieka wywołuje zupełnie inne skutki, niż ograniczenie jego wolności artystycznej. Ww. termin powinien zostać istotnie uściślony poprzez doprecyzowanie tekstu przepisu lub dodanie motywu, który zacieśniałby kierunki interpretacyjne.**
- **Pojęcia “bias” oraz “discriminatory effect” powinny zostać doprecyzowane.** Możliwym rozwiązaniem byłoby “wyłączenie pozytywne” tj. wskazanie np. kiedy “bias” jest dopuszczalny. Można także „bias” (art. 10.2.f; motyw 33) zdefiniować jako „dyskryminację rozumianą jako błąd statystyczny (odgórne przypisywanie cech niezgodnych z rzetelnie uzyskanymi statystykami) lub odgórne wprowadzanie założeń szkodliwych dla jednostki ”.
- Sugerujemy także, aby **pojęcie “child”** (w ramach risk management- art. 9.9) zostało doprecyzowane. Po pierwsze, wskazanie konkretnej granicy wiekowej rozwiązałoby problem rozbieżności na poziomie ustaw krajowych. Po drugie, ocena wpływu na dziecko w ramach risk management powinna opierać się o konkretne granice wiekowe z uwagi na istotne różnice poznawcze i emocjonalne dzieci.

3. **Wyjaśnienia dotyczące obowiązków przejrzystości (art. 52):** czym są „interakcje z osobami fizycznymi”?

- Art. 52 nakłada obowiązek powiadomienia użytkownika, że wchodzi on w interakcję z systemem AI, jeśli nie jest to oczywiste. KE podaje przykład chatbota, który miałby na przykładzie wyjaśnić ten obowiązek. Jednak język w obecnym brzmieniu jest zbyt niejasny, biorąc pod uwagę, że sztuczna inteligencja jest zintegrowana z wieloma systemami skierowanymi do użytkownika używanymi w celu uzyskania rekomendacji, wyszukiwania informacji, udzielania wskazówek, prognoz — jak zdefiniować „interakcję z osobami fizycznymi” i jak szeroko/wąsko należy to postrzegać? Skoro sama definicja systemu SI (art. 3 pkt 1) została oparta o wejście systemu SI w "interakcję" z użytkownikiem, to w gruncie rzeczy każdy system SI będzie spełniał tę przesłankę, bowiem jest ona częścią definicji SI. Mając na uwadze postępujące zaawansowanie systemów SI, pozostaje zatem problem dalszej pracy nad zdefiniowaniem „interakcji z osobami fizycznymi”.

- Art. 52 ust. 3 nakłada obowiązek informowania o dokonaniu przez user of AI zmiany obiektywnej rzeczywistości, oznaczając to działanie jako “deep fake”. Użycie w tym kontekście sformułowania “deep fake” nie wydaje się zasadne, ponieważ pojęcie to sugeruje celowe wprowadzanie odbiorcy w błąd co do tożsamości przedstawionego obiektu, często z niskich pobudek. Oczywiście takie praktyki powinny byc napiętnowane. Wprowadzanie zmian do przedstawianej rzeczywistości może mieć jednak różnorodną motywację - artystyczną, użytkową, wyjaśniającą, edukacyjną, cytującą, polemizującą etc. Należy jednocześnie założyć, że właśnie te sytuacje stanowią większość przypadków “manipulowania” rzeczywistym obrazem. Nałożenie w każdym przypadku obowiązku informowania o wprowadzonej zmianie, nawet w najmniejszym wymiarze, mogłoby znacząco utrudnić prowadzenie działalności dziennikarskiej, artystycznej, czy szerzej - twórczej. Drugi ustęp wskazanego punktu jedynie w niewielkim stopniu eliminuje to ryzyko. Ustawodawca powinien sprecyzować, czy każda zmiana winna być oznaczona.

4. Zaburzona równowaga obowiązków między dostawcami, wdrażającymi i użytkownikami sztucznej inteligencji wysokiego ryzyka

- W obecnym brzmieniu przepisy nie rozróżniają między obowiązkami nakładanymi na użytkownika sztucznej inteligencji, jeśli pełni on rolę wdrażającego dane zastosowanie AI, a obowiązkami „dostawcy AI” wobec klienta.

- Wdrażający zastosowania AI powinni ostatecznie być głównym podmiotem oceny, ponieważ przedsiębiorstwa oferujące narzędzia AI ostatecznie nie są w stanie zweryfikować zastosowań końcowych, do których wykorzystywane są

ich systemy, ani dodatkowych danych, które mogą być wprowadzane do systemu. Dostawcy rozwiązań AI mogą i powinni dostarczać wszystkie informacje niezbędne wdrażającym do przeprowadzenia samooceny. Jest to bardzo ważne dla dostawcy rozwiązań/interfejsów API, nad którymi dostawca rozwiązań AI nie przejmuje kontroli, gdy użytkownik wyraża zgodę na umożliwienie klientom/użytkownikom dostępu do rozwiązania według własnego uznania.

5. **Zwolnienia dotyczące wielozadaniowych systemów/narzędzi typu open source:**

- obowiązki przestrzegania wymogów dotyczących systemów AI powinny spoczywać na podmiotach prawnych lub osobach fizycznych korzystających z narzędzi typu open source, takich jak TensorFlow, czy AutoML, ponieważ mają one ostateczną kontrolę nad celem i wykorzystaniem zastosowań sztucznej inteligencji. Nałożenie obowiązków na dostawcę narzędzi open source w dużej mierze zniechęciłoby do udostępniania takich technologii, które wspierają całe ekosystemy innowacji.
- Zwolnienie z obowiązku publikacji badań podstawowych: wyjaśnienie, że publikacja badań podstawowych nie kwalifikuje się jako „wprowadzanie na rynek" lub „oddawanie do użytku"
- Wymagane wyjaśnienia/zabezpieczenia np. zagwarantowanie wolnych od błędów zbiorów danych, lub publikacja kodu źródłowego w celu nadzoru rynku, nie zawsze mogą być możliwe i mogą doprowadzić do tzw. efektu mrożącego. Zasadnym może się wydawać wprowadzenie takich obowiązków dla zastosowań wysokiego ryzyka, jednak nie widzimy racjonalności dla innego rodzaju zastosowań. Czy np. błędne tłumaczenie wynikające z niepełnej/niereprezetatywnej bazy danych ma taki sam negatywny efekt jak np. interpretacja badania medycznego?
- Niejasne, czym jest „element zabezpieczający" (safety component) z art. 3 ust. 14, zwłaszcza w związku z dyrektywą w sprawie urządzeń radiowych. Czy np. system Android jest "elementem bezpieczeństwa" urządzenia mobilnego? Czy obowiązek oznacza, że samo urządzenie lub jej system musi spełniać jakąś funkcję krytyczną dla bezpieczeństwa?

6. **Możliwość szkolenia danych nawet w przypadku niektórych aplikacji wysokiego ryzyka**

- **Testowanie niektórych systemów sztucznej inteligencji, takich jak ocena ryzyka kredytowego, na mniejszą skalę, bez konieczności traktowania ich jako wysokiego ryzyka**, może nadal być korzystne dla konsumentów przy jednoczesnym ograniczaniu ryzyka związanego z AI. Jest to powszechna

praktyka pozwalająca na dobór najlepszych rozwiązań AI do konkretnych zastosowań.

7. Kontrola zgodności

- Przepis art 64 projektowanego rozporządzenia nakazujący ujawniać i udostępniać dane oraz dokumentację wymaga wyraźnego sprecyzowania pod kątem relacji z tajemnicą przedsiębiorstwa. Informacje te mogą mieć bowiem charakter poufny, a nierzadko stanowić główny o ile nie jedyny czynnik przewagi konkurencyjnej danego rozwiązania. Uważamy, że prawa przyznane krajowym organom nadzoru rynku uprawnień jak **żądanie dostępu do zbiorów danych, interfejsów API i kodów źródłowych są zbyt daleko idące.** W szczególności brak precyzyjnych definicji kluczowych ryzyk (takich jak dyskryminacja, stronniczość) nie zwiększa obiektywności oceny nadzorczej.
- Dodatkowo, **na poziomie krajowym powinny być wprowadzone rzeczywiste gwarancje procesowe chroniące tajemnicę przedsiębiorstwa zarówno przed nieuprawnionym dostępem jak i nadmiarowym, czy nieuzasadnionym dostępem osób lub podmiotów do tego typu informacji.**
- Podsumowując, podejście oparte na **zasadzie proporcjonalności** powinno być zawarte w systemach zgodności aplikacji AI wysokiego ryzyka.

Z poważaniem,

Włodzimierz Schmidt
Prezes Zarządu