



## Stanowisko Fundacji Panoptykon w sprawie Aktu prawnego o usługach cyfrowych (Digital Services Act)<sup>1</sup>

Fundacja Panoptykon jest organizacją pozarządową, której celem jest ochrona praw człowieka, w szczególności prawa do prywatności, w kontekście rozwoju nowych technologii. W naszej działalności zajmujemy się m.in. problemami związanymi z rosnącą dominacją platform internetowych.

Akt prawny o usługach cyfrowych (Digital Services Act, dalej: projekt, DSA) jest niezwykle ważną i potrzebną regulacją, która zdefiniuje zasady funkcjonowania platform internetowych na wiele lat. Projekt rozporządzenia zaproponowanego przez Komisję Europejską dotyka w istocie wszystkich kluczowych obszarów działania platform, które wymagają regulacyjnej interwencji, w tym moderacji treści, reklam internetowych i algorytmów rekomendujących treści.

Jednak w obecnym kształcie w wielu miejscach projekt jest **zbyt ogólny** i daje platformom internetowym zbyt szerokie pole do interpretacji. Dodatkowo, projekt **skupia się wyłącznie na zasadach przejrzystości działania platform, zasadniczo nie wprowadzając konkretnych ograniczeń** (np. co do możliwości targetowania reklam). Przejrzystość jest niewątpliwie niezbędnym punktem wyjścia do zapewnienia rozliczalności platform, ale sama w sobie, bez konkretnych ograniczeń i systemu kontroli regulatora (np. nad algorytmami platform), nie wzmocni praw i wolności użytkowników ani nie poprawi warunków konkurencji.

W niniejszym stanowisku skupiamy się na przepisach dotyczących platform internetowych i bardzo dużych platform internetowych w trzech najistotniejszych z perspektywy praw człowieka obszarach: moderacji treści, przejrzystości targetowania reklam internetowych i kontroli nad algorytmami zarządzającymi treścią, licząc na uwzględnienie naszych uwag w stanowisku Rządu.

Biorąc pod uwagę znaczenie propozycji Komisji i to, jak wielu zagadnień dotyka, spodziewamy się, że stanowisko Rządu będzie ewoluowało wraz z kolejnymi spotkaniami Rady. W tym kontekście liczymy na bieżącą współpracę Rządu z interesariuszami, w tym z organizacjami społecznymi.

### 1. Zasady dotyczące moderacji treści

Zaproponowane w projekcie rozwiązania dotyczące moderacji treści przez hostingodawców (w tym platformy internetowe) w dużej części zasługują na aprobatę i powinny zostać utrzymane w toku dalszych prac nad DSA.

#### 1.1. Podstawowe zasady

Przede wszystkim dotyczy to utrzymania obowiązującej obecnie zasady warunkowego wyłączenia odpowiedzialności hostingodawców za treści dodawane przez użytkowników oraz zakazu wprowadzenia obowiązku generalnego monitorowania takich treści. Istotne jest również,

---

<sup>1</sup> Stanowisko przygotowane przez Karolinę Iwańską i Dorotę Głowacką.

że przewidziana w projekcie procedura „notice and takedown” dotyczy wyłącznie bezprawnych treści i nie obejmuje innego typu „szkodliwych” czy „niebezpiecznych” materiałów, które jednocześnie nie są nielegalne. Objęcie taką procedurą innych niż bezprawne kategorii treści stwarzałoby naszym zdaniem zbyt duże ryzyko nadmiarowego usuwania publikacji, które powinny być dopuszczalne w debacie publicznej. Ważne jest ponadto, że procedura „notice and take down”, w tym przewidziane w niej gwarancje dla użytkowników oraz odpowiadające im obowiązki hostingodawców/platform (o których mowa niżej), mają dotyczyć zarówno treści zgłoszonych do moderacji, jak i materiałów, które tacy usługodawcy internetowi samodzielnie zidentyfikują jako bezprawne, ale także – co bardzo istotne – sprzeczne z ich wewnętrznymi regulaminami, w tym przy użyciu automatycznych narzędzi. W tym kontekście zgadzamy się z Komisją Europejską, że **nie ma żadnego uzasadnienia, aby różnicować sytuację osób zablokowanych w zależności od podstawy usunięcia treści, czy tym bardziej sposobu pozyskania przez platformę wiedzy o jej bezprawnym charakterze.** Zarówno blokady nałożone w ramach procedury „notice and takedown”, jak i w wyniku stosowania wewnętrznych regulaminów, powinny podlegać takiej samej kontroli.

## 1.2. Wykorzystanie algorytmów w moderacji treści

Co więcej, taka kontrola jest szczególnie ważna, gdy usunięcie treści było efektem wskazania jej przez automatyczne narzędzia. Użycie algorytmów w moderacji generuje bowiem wyjątkowo wysokie ryzyko „fałszywie pozytywnych trafień” (*false positives*)<sup>2</sup>. Z tych względów pozytywnie oceniamy propozycję pewnych obowiązków i obostrzeń związanych z wykorzystaniem algorytmów przez platformy internetowe w procesie oceny dopuszczalności i usuwania treści (m.in. zakaz rozpoznawania odwołań od decyzji o usunięciu jedynie w sposób zautomatyzowany czy obowiązki wzmacniające przejrzystość funkcjonowania takich algorytmów). Jest to tym bardziej istotne, że – jak wynika z publikowanych raportów przejrzystości platform – zdecydowana większość niedopuszczalnych treści identyfikowana jest właśnie za pomocą algorytmów (np. na You Tube jest to 95 %<sup>3</sup>). Jednoczesne wprowadzenie w art. 6 tzw. „klauzuli Dobrego Samarytanina” (która sama w sobie jest słusznym rozwiązaniem, pod warunkiem obwarowania jej odpowiednimi gwarancjami) może dodatkowo utrwalić bądź nawet wzmocnić tę praktykę. Za **niewystarczające** uznajemy jednak zaproponowane przez Komisję **gwarancje dotyczące wykorzystania algorytmów w szerszym procesie zarządzania treścią** na platformach internetowych (zob. uwagi w pkt. 3).

## 1.3. Przejrzystość usuwania treści i zasady sprawiedliwej procedury (*due process*)

Za kluczowe propozycje, które zasługują na aprobatę, uznajemy rozwiązania dotyczące wprowadzenia regulacji zapewniających większą przejrzystość samego procesu moderacji treści oraz rozwiązania zmierzające do **ograniczenia arbitralności podejmowanych przez platformy decyzji** o usunięciu treści (art. 12, 15, 17). Odnosi się to m.in. do obowiązku odpowiedniego uzasadnienia decyzji o usunięciu treści, wyjaśnienia, w jaki sposób została podjęta, obowiązku umożliwienia użytkownikowi przedstawienia własnych argumentów, wprowadzenia wspomnianych ograniczeń dotyczących wykorzystania algorytmów w procesie rozpoznawania odwołań, obowiązku szczegółowego opisanie ww. gwarancji w wewnętrznych regulaminach i innych elementach, które **zwiększą uprawnienia zablokowanych**

---

<sup>2</sup> Zob. E. Llanso, J. van Hoboken, P. Leerssen, J. Harambam (2020), Artificial Intelligence, Content Moderation, and Freedom of Expression’, Transtaltic Workin Group, <https://www.ivir.nl/publicaties/download/AI-Llanso-Van-Hoboken-Feb-2020.pdf>

<sup>3</sup> <https://transparencyreport.google.com/youtube-policy/removals>

**użytkowników do weryfikowania zasadności decyzji moderacyjnych.** Fakt, że obecnie obowiązujące regulacje nie przewidują żadnych gwarancji w tym zakresie uznajemy za jeden z ich największych mankamentów, a uzupełnienie tej luki - jeden z najważniejszych celów, które powinno realizować DSA.

#### 1.4. Zewnętrzny, niezależny nadzór nad moderacją treści

Bardzo ważnym elementem uzupełniającym ww. gwarancje przejrzystości i sprawiedliwej procedury po stronie platform jest **utworzenie systemu niezależnej, zewnętrznej kontroli nad ich ostatecznymi decyzjami.** Projekt przewiduje w tym zakresie powołanie specjalnych organów pozasądowych do rozpoznawania sporów o treści. W naszej ocenie organy te będą pełnił bardzo ważną funkcję (istotnie poprawiającą dostępność i efektywność niezależnego, zewnętrznego mechanizmu nadzoru w kontekście spodziewanej dużej liczby wpływających spraw), ale **ich wprowadzenie nie może w pełni zastąpić możliwości uruchomienia kontroli sądowej.** W tym zakresie zaproponowane przepisy są w naszej ocenie niejasne i wymagają doprecyzowania bądź uzupełnienia, które wprost wprowadzi możliwość i określi warunki złożenia przez osoby, które czują się niesłusznie zablokowane, odwołania również do sądu.

#### 1.5. Raporty przejrzystości

Popieramy ponadto wprowadzenie obowiązku prawnego przygotowywania i publikowania przez platformy internetowe cyklicznych „raportów przejrzystości”, obejmujących informacje o usuwaniu treści, ale także – co szczególnie warto podkreślić – **o innych działaniach powodujących ograniczenie widoczności publikowanych materiałów oraz wykorzystaniu algorytmów w moderacji** (art. 13, 23). Za istotne uznajemy również precyzyjne określenie w rozporządzeniu elementów składowych tych raportów, co przyczyni się nie tylko do wprowadzenia odpowiedniego standardu podawanych w nich danych, ale także ujednolicenia ich formuły dla poszczególnych platform. To z kolei ułatwi analizę tych dokumentów i zwiększy ich skuteczność jako narzędzia kontroli nad działaniami usługodawców internetowych.

#### 1.6. Niewystarczające gwarancje dla użytkowników

Jednocześnie, w przedstawionej propozycji dotyczącej zasad moderacji treści dostrzegamy kilka obszarów, które naszym zdaniem wymagają pewnych modyfikacji.

Opisane gwarancje sprawiedliwej procedury w ramach platform i mechanizmu zewnętrznego nadzoru nad działaniami platform dotyczą, mówiąc ogólnie, wyłącznie decyzji o usunięciu treści lub usunięciu/zawieszeniu konta (art. 15 i 17). Tymczasem, **powinny one objąć każde działanie platform, które prowadzi do zmniejszenia widoczności danej treści bądź w inny sposób ogranicza możliwość szerokiego docierania z nią do odbiorców** – dotyczy to np. tzw. „ograniczania zasięgów” (*shadow bans*), utraty możliwości promowania treści, czy zarabiania na nich. Uważamy, że przy stosowaniu wszelkich tego rodzaju środków powinny być **spełnione te same standardy przejrzystości i rozliczalności, jak w przypadku usuwania treści.** Choć działania te mogą wydawać się mniej inwazyjne, w praktyce prowadzą często do równie poważnej ingerencji w swobodę wypowiedzi, zwłaszcza że – choćby w przypadku ograniczania zasięgów – odbywa się ona często bez jakiegokolwiek notyfikacji użytkownika, powodując, że nie tylko nie ma on możliwości zakwestionować takiego działania, ale może nawet o nim nie wiedzieć. Jednocześnie, przewidziany w projekcie obowiązek przekazywania informacji statystycznych na temat działań powodujących ograniczenie widoczności treści (art. 13 ust. 1 lit. c) nie stanowi wystarczającej gwarancji przed nadużywaniem tych środków, gdyż **nie chroni dostatecznie interesów indywidualnych użytkowników, wobec których je zastosowano.**

Dodatkowo, zaproponowana procedura w przypadku usunięcia treści nie przewiduje konieczności powiadomienia i wysłuchania argumentów użytkownika **przed** usunięciem treści. Notyfikacja i możliwość odniesienia się przez użytkownika do takiej decyzji zostały przewidziane dopiero *post factum*. Tymczasem, w naszej ocenie wprowadzenie takiego środka **przed podjęciem działania przez platformę** byłoby rozwiązaniem, które dodatkowo zwiększyłoby gwarancje sprawiedliwości proceduralnej w zakresie moderacji.

Jak sygnalizowaliśmy wyżej, dostrzegamy **potrzebę jasnego zagwarantowania w rozporządzeniu możliwości i warunków uruchomienia kontroli sądowej** dla użytkowników, którzy kwestionują zasadność decyzji o usunięciu treści lub konta (lub zastosowania innej formy moderacji, takiej jak zmniejszenie widoczności treści). Kwestia ta nie powinna być uzależniona od istnienia bądź swobody wprowadzenia odpowiedniego mechanizmu ochrony sądowej na poziomie poszczególnych państw członkowskich UE. Nie kłóci się to oczywiście z powołaniem przewidzianych w projekcie rozporządzenia organów pozasądowych, które mogą istotnie wesprzeć efektywność niezależnego mechanizmu nadzoru nad decyzjami moderacyjnymi. Jednocześnie w przypadku obu tych ścieżek rozporządzenie powinno przewidywać jak najbardziej konkretne rozwiązania zapewniające realną dostępność i szybkość takich postępowań.

## 2. Przejrzystość targetowania reklam

### 2.1. Uwagi ogólne

Model biznesowy wiodących platform internetowych opiera się na sprzedaży reklam **targetowanych do użytkowników w oparciu o profilowanie wykorzystujące dane na ich temat oraz dane na temat innych użytkowników**. Ten proces jest w pełni zautomatyzowany: algorytmy ustalają cechy i profile użytkowników w oparciu o dane zebrane przez platformy, dopasowują te profile do kryteriów targetowania wybranych przez reklamodawców oraz optymalizują proces dostarczania reklamy, czyli decydują o tym, która reklama w danym momencie powinna się wyświetlić konkretnej osobie<sup>4</sup>.

Na przestrzeni ostatnich kilku lat ujawniono i opisano wiele problemów i zagrożeń, które łączą się z targetowaną reklamą: od naruszania prywatności użytkowników (w tym wykorzystywania do targetowania ich wrażliwych danych), przez dyskryminację pewnych grup osób, po społeczne konsekwencje związane z manipulacją i dezinformacją<sup>5</sup>. Wiele z tych problemów wynika z braku przejrzystości tego, jakie kryteria targetowania wybierają reklamodawców i jak reklama jest następnie dopasowywana przez platformy. W odpowiedzi na tę kwestię Komisja Europejska zaproponowała w DSA wprowadzenie **podstawowych zasad przejrzystości targetowania reklam**.

Zgodnie z art. 24 projektu wszystkie platformy internetowe będą zobowiązane do przedstawienia użytkownikom, w odniesieniu do każdej reklamy, **indywidualnych wyjaśnień** zawierających m.in. „istotne informacje o głównych kryteriach wykorzystanych do ustalenia odbiorcy reklamy”

<sup>4</sup> Więcej na temat technicznego funkcjonowania reklamy internetowej na platformach internetowych w raporcie Fundacji Panoptykon „Kto (naprawdę) Cię namierzył. Facebook w polskiej kampanii wyborczej”:

<https://panoptykon.org/ktocienamierzyl-raport#part-1-4>

<sup>5</sup> Por. np. zestawienie problemów przygotowane przez Europejską Radę Ochrony Danych:

[https://edpb.europa.eu/sites/edpb/files/consultation/edpb\\_guidelines\\_202008\\_onthetargetingofsocialmediausers\\_en.pdf](https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_202008_onthetargetingofsocialmediausers_en.pdf)

(art. 24 lit. c). Dodatkowo, zgodnie z art. 30, bardzo duże platformy będą zobowiązane do utrzymywania **publicznych repozytoriów (bibliotek) reklam**, w których znajdą się zbiorcze informacje o każdej reklamie, w tym główne kryteria wykorzystane do targetowania (lit. d).

W praktyce, przepisy te **sankcjonują już istniejące mechanizmy przejrzystości**, które wiodące platformy internetowe wdrożyły po przystąpieniu w 2018 r. do unijnego Kodeksu postępowania w zakresie zwalczania dezinformacji<sup>6</sup>. Funkcjonowanie zarówno indywidualnych wyjaśnień, jak i bibliotek reklam było przez ostatnie dwa lata przedmiotem badań i oceny, zarówno ze strony Komisji Europejskiej i państw członkowskich UE (w Polsce efektywność mechanizmów przejrzystości platform internetowych badała Krajowa Rada Radiofonii i Telewizji), jak i badaczy, dziennikarzy oraz organizacji społecznych. Również w Fundacji Panoptykon – z pomocą ponad 6200 polskich użytkowników Facebooka – badaliśmy skuteczność mechanizmów przejrzystości reklam w ramach projektu dotyczącego mikrotargetowania politycznego w mediach społecznościowych<sup>7</sup>. Te doświadczenia pozwalają zatem na lepszą, opartą na empirycznych obserwacjach, ocenę rozwiązań zaproponowanych przez Komisję w DSA. W tym kontekście:

- pozytywnie oceniamy uregulowanie zasad przejrzystości reklam **w powszechnie obowiązującym rozporządzeniu**. Na niedociągnięcia instrumentów samoregulacyjnych wskazywała sama Komisja Europejska w opublikowanej we wrześniu 2020 r. ocenie dwóch lat funkcjonowania Kodeksu postępowania<sup>8</sup>;
- na pozytywną ocenę zasługuje również **rozszerzenie obowiązku prezentowania w bibliotece reklam wszystkich reklam, a nie tylko reklam politycznych** (jak miało to miejsce dotychczas). Dotychczasowe doświadczenia pokazują, że ograniczenie zasad przejrzystości do reklamy politycznej przyniosło rozbieżności definicyjne oraz trudności w określeniu, jaka treść jest polityczna, a jaka nie, co niejednokrotnie prowadziło do nieujmowania w bibliotekach niektórych reklam. Obowiązek publikacji informacji na temat wszystkich reklam odpowiada również na problem możliwych nadużyć w obszarze reklamy komercyjnej, np. dyskryminację określonych grup, naruszanie praw konsumentów, czy wykorzystywanie do targetowania reklamy wrażliwych danych;
- jednocześnie, dotychczasowe obserwacje związane z funkcjonowaniem narzędzi przejrzystości prowadzą do wniosku, że wymogi sformułowane przez Komisję w DSA są **zbyt ogólne i w obecnym kształcie mogą nie być wystarczające do zapewnienia efektywnej przejrzystości i rozliczalności procesu targetowania reklam**. Poniżej przedstawiamy szczegółowe uwagi dotyczące tej kwestii.

## 2.2. Zidentyfikowane problemy istniejących mechanizmów przejrzystości

W kontekście wielu badań poświęconych funkcjonowaniu istniejących narzędzi przejrzystości, pragniemy zwrócić uwagę na dwa kluczowe problemy, których w naszej opinii projekt w obecnym kształcie nie rozwiąże:

- **niekompletne i wprowadzające w błąd informacje o parametrach wybranych przez reklamodawcę**

---

<sup>6</sup> [https://ec.europa.eu/commission/presscorner/detail/pl/IP\\_18\\_3370](https://ec.europa.eu/commission/presscorner/detail/pl/IP_18_3370)

<sup>7</sup> Raport „Kto (naprawdę) Cię namierzył. Facebook w polskiej kampanii wyborczej” dostępny na <https://panoptykon.org/ktocienamierzyl-raport>

<sup>8</sup> <https://ec.europa.eu/digital-single-market/en/news/assessment-code-practice-disinformation-achievements-and-areas-further-improvement>

Empiryczne badania<sup>9</sup> pokazały, że platformy internetowe, w szczególności Facebook, nie pokazują użytkownikom wszystkich kryteriów targetowania reklam. Indywidualne wyjaśnienia (dostępne na Facebooku w ramach funkcji „Dlaczego to widzę?”) zawierają tylko jedno lub dwa kryteria, podczas gdy reklamodawcy mają ich do dyspozycji dużo więcej i zazwyczaj z nich korzystają. Badacze ustalili też, że platforma nie informuje użytkowników o wykluczeniu ich z grupy odbiorców danej reklamy (ani tym bardziej o podstawie takiego działania), mimo że taka opcja jest dostępna dla reklamodawców.

Dodatkowo, badacze ustalili, że wyjaśnienia prezentowane przez Facebooka mogą wprowadzać w błąd. Platforma ujawnia bowiem wybrane przez reklamodawcę kryterium, które dotyczy potencjalnie największej grupy odbiorców, a nie to, które odegrało największą rolę w pokazaniu reklamy konkretnej osobie. Przykładowo, jeśli reklamodawca wybrał kryteria „medycyna” (potencjalny zasięg 668 mln użytkowników) i „ciąża” (316 mln), użytkownicy dowiedzą się, że zobaczyli reklamę, gdyż interesują się medycyną, nawet jeśli w ich przypadku większe znaczenie w doborze reklamy miało to, że są w ciąży<sup>10</sup>.

- **brak wyjaśnienia optymalizacji targetowania przeprowadzonej przez samą platformę**

W procesie targetowania reklamy platformy internetowe nie pełnią jedynie roli biernego pośrednika pomiędzy reklamodawcami a użytkownikami, ale **odgrywają kluczową rolę i wprost kształtują wyniki targetowania**. Gdy reklamodawca wybierze kryteria targetowania, algorytmy platform ustalają, którzy użytkownicy je spełniają. Na tym rola platformy się nie kończy – ponieważ grupa potencjalnych odbiorców (użytkowników, którzy spełniają kryteria reklamodawców) jest zazwyczaj dużo większa niż grupa, do której reklamodawca może dotrzeć przy określonym budżecie reklamowym, **platformy optymalizują targetowanie pod kątem celów, które chce osiągnąć reklamodawca** (np. kliknięcia) i wybierają spośród wszystkich potencjalnych odbiorców te osoby, które z większym prawdopodobieństwem zareagują w oczekiwany sposób. Badania przeprowadzone na Northeastern University w Bostonie<sup>11</sup> wskazują na to, że **same algorytmy optymalizacyjne mogą być odpowiedzialne za dyskryminację określonej grupy osób** lub wspieranie baniek informacyjnych i – w efekcie – polaryzacji politycznej poprzez pobieranie większych opłat za dotarcie z reklamą do osób o innych poglądach.

Tymczasem, platformy nie informują publicznie (np. w regulaminie) o tym, jakie parametry są brane pod uwagę w procesie optymalizacji i jaki jest jej cel. Również indywidualne wyjaśnienia prezentowane użytkownikom nie pozwalają im ustalić, jakie konkretnie cechy zadecydowały o uznaniu ich za skłonnych do zareagowania na konkretną reklamę w sposób, jaki oczekuje tego reklamodawca.

### **2.3. Niewystarczająco precyzyjne przepisy i brak dodatkowych ograniczeń dla platform**

---

<sup>9</sup> A. Andreou i in., Investigating Ad Transparency Mechanisms in Social Media: A Case Study of Facebook’s Explanations, <https://hal.archives-ouvertes.fr/hal-01955309/document>

<sup>10</sup> Autentyczny przykład takiej sytuacji opisaliśmy w raporcie „Kto (naprawdę) Cię namierzył. Facebook w polskiej kampanii wyborczej”: <https://panoptykon.org/ktocienamierzyl-raport#part-2-4>

<sup>11</sup> M. Ali i in., Discrimination Through Optimization: How Facebook’s Ad Delivery Can Lead to Skewed Outcomes: <https://arxiv.org/abs/1904.02095>; oraz M. Ali i in., Ad Delivery Algorithms: The Hidden Arbiters of Political Messaging: <https://arxiv.org/abs/1912.04255>



W kontekście opisanych wyżej problemów obecne brzmienie art. 24 i 30 jest zdecydowanie zbyt ogólne. Przepisy te wspominają wyłącznie o „głównych parametrach” targetowania, co pozostawia platformom **zbyt szerokie pole do interpretacji**. To platformy dokonywałyby oceny, które kryteria zostaną uznane za najważniejsze, przy jednoczesnym braku możliwości zewnętrznej weryfikacji, czy ta ocena była właściwa. Jak wyjaśniamy powyżej, wiodące platformy pokazują wyłącznie niekontrowersyjne, demograficzne kryteriów targetowania. Bardziej szczegółowe kryteria, zwłaszcza te, których użytkownicy bezpośrednio nie ujawnili, a które zostały przez platformę na ich temat wyinterpretowane, nie są udostępniane. W związku z tak ogólnie sformułowanym brzmieniem przepisu, który nie odnosi się wprost do tego, jakie parametry targetowania powinny zostać uznane za „główne” lub chociażby jakie kryteria powinny zostać zastosowane do przeprowadzenia takiej oceny, istnieje ryzyko utrzymania tego złego standardu. Uniemożliwi to efektywną weryfikację działań platform i reklamodawców, a tym samym – wykrycie ewentualnych nadużyć.

**Dlatego rekomendujemy, by Rząd w swoim stanowisku uwzględnił następujące postulaty:**

- (1) przepisy powinny zostać doprecyzowane poprzez zobowiązanie platform do ujawnienia – zarówno w bibliotekach reklam, jak i indywidualnych wyjaśnieniach – wszystkich, a nie tylko „głównych” kryteriów targetowania wybranych przez reklamodawców, łącznie z – jeśli ma to zastosowanie – wybranym przez reklamodawcę celem optymalizacji (np. kliknięcia, odtworzenia filmu);
- (2) przepisy powinny ustanowić konkretne zasady przejrzystości i rozliczalności dla algorytmów samych platform wykorzystywanych w procesie targetowania reklam, zwłaszcza algorytmów optymalizacyjnych (por. uwagi w pkt. 3 poniżej).

Jednocześnie, pragniemy zwrócić uwagę na to, że **projekt ogranicza się do wprowadzenia wymogów przejrzystości reklam, nie przewiduje natomiast żadnych ograniczeń**, np. co do niedopuszczalnych technik lub kryteriów targetowania. Tym samym propozycja Komisji jest o wiele bardziej zachowawcza niż chociażby stanowisko Parlamentu Europejskiego, który w ramach własnej inicjatywy prawodawczej postulował wprowadzenie ograniczeń, a nawet zakazu dla targetowania reklam w oparciu o dane osobowe na rzecz reklam niewykorzystujących danych<sup>12</sup>. Postulujemy, aby Rząd w swoim stanowisku odniósł się do propozycji Parlamentu i wypracował własne propozycje ograniczeń.

### **3. Systemy rekomendacji i inne metody algorytmicznego zarządzania treścią**

Projekt DSA w art. 29 wprowadza – w odniesieniu do bardzo dużych platform – minimalny standard przejrzystości dla algorytmów, które wykorzystują one do personalizowania i rekomendowania treści. Zgodnie z tym przepisem, platformy będą musiały poinformować swoich użytkowników o głównych parametrach, które te systemy wykorzystują. Muszą one również umożliwić użytkownikom wyłączenie personalizacji oraz mogą opcjonalnie wprowadzić inne możliwości modyfikacji parametrów algorytmu.

Pozytywnie oceniamy chęć zapewnienia przejrzystości algorytmów rekomendacji treści oraz zwiększenia kontroli użytkowników nad treściami, które do nich docierają. Jednocześnie, podobnie jak w przypadku przejrzystości reklam, zaproponowane przepisy w naszej opinii są zbyt

<sup>12</sup> <https://www.europarl.europa.eu/news/en/press-room/20201016IPR89543/digital-eu-must-set-the-standards-for-regulating-online-platforms-say-meps>

ogólne, by umożliwić realną weryfikację zasad działania algorytmów stosowanych przez platformy i wykrycie ewentualnych nadużyć. W tym kontekście przedstawiamy następujące uwagi:

### 3.1. Rozszerzenie zakresu algorytmów, których działanie powinno być wyjaśniane

Zakres przedmiotowy art. 29 powinien zostać rozszerzony o **algorytmy optymalizacji reklam** (por. pkt 2.2) i **algorytmy ograniczające widoczność konkretnych treści** (tzw. „shadow bans”, por. pkt 1.6). Nie ma powodu, dla którego platformy nie powinny wyjaśniać działania tych algorytmów na podobnych zasadach, jak systemów rekomendacyjnych. Podobnie jak systemy rekomendacyjne te algorytmy pełnią bowiem istotną rolę w tym, jakie treści docierają (lub nie docierają) do użytkowników.

### 3.2. Precyzyjne zasady przejrzystości i rozliczalności algorytmów

Celem art. 29 jest umożliwienie użytkownikom zrozumienia podstawowych zasad działania algorytmów personalizujących treści. W naszej opinii przedstawienie wyłącznie głównych parametrów działania systemu nie będzie wystarczające do spełnienia tego celu. Wyjaśnienie będzie przydatne wówczas, gdy pozwoli użytkownikom **zrozumieć cele i logikę systemu oraz ryzyka, jakie się z nim wiążą**. W związku z tym, poza głównymi parametrami, platformy powinny ujawnić informacje o tym, **jak te parametry wzajemnie ze sobą oddziałują** (np. jakie wagi zostały im przypisane, które parametry mają największe znaczenie i w jakich sytuacjach). Platformy powinny również szczegółowo wyjaśnić cele działania algorytmu, przy czym cele te powinny obejmować nie tylko efekt doświadczany przez użytkownika (szeroko pojęta personalizacja), ale **prawdziwą intencję i założenia samego właściciela systemu** – czyli np. cel optymalizacji jakim jest zwiększenie zaangażowania lub czasu spędzanego na platformie. Dodatkowo, platformy powinny wyjaśnić, jakie ryzyka mogą się wiązać z przyjętymi przez nie algorytmami (np. ryzyko błędów lub uprzedzeń) i wyjaśnić, **jakie kroki podjęły do zminimalizowania tych ryzyk** (np. jakie techniczne zasady sprawiedliwości zostały przyjęte lub na jakich danych algorytm został wytrenowany). Biorąc pod uwagę również to, że algorytmy personalizacji treści są oparte o uczenie maszynowe, w związku z czym ich działanie podlega ciągłym zmianom, informacje przedstawiane użytkownikom powinny być na bieżąco aktualizowane, a istotne zmiany parametrów lub celów systemów publicznie ogłaszane przed ich wprowadzeniem.

### 3.3. Interoperacyjność jako warunek efektywnej kontroli użytkowników nad algorytmami

Art. 29 przewiduje, że platformy mogą umożliwić użytkownikom zmianę parametrów algorytmu rekomendacji treści. Jednak to, czy użytkownicy taki przywilej uzyskają, zależy wyłącznie od woli platform. Biorąc pod uwagę wysoki koszt wprowadzenia takich rozwiązań technicznych spodziewamy się, że w praktyce platformy nie będą po tę opcję sięgać. Dodatkowo, pozostawienie do uznania platform tego, jakie parametry można zmieniać, a jakich nie, nie rozwiązuje problemu kontroli platform nad treścią, która użytkownikom się wyświetla, i nie daje użytkownikom wystarczających narzędzi do tego, by samodzielnie decydować o swoim sposobie korzystania z usług.

Narzędziem, które w praktyce mogłoby przywrócić kontrolę nad treścią użytkownikom byłoby wprowadzenie obowiązku zapewnienia przez platformy interoperacyjności określonych warstw swoich usług (np. usług personalizowania treści). Obecnie obowiązek zapewnienia interoperacyjności został wprowadzony wobec tzw. strażników dostępu (gatekeeperów)



w projekcie Digital Markets Act (akt o rynkach cyfrowych), ale dotyczy on wyłącznie „ancillary services”, czyli usług pobocznych, a nie ma zastosowania do kluczowych usług platform („core services”. Tymczasem, wprowadzenie pełnej interoperacyjności w miejsce zamkniętych ekosystemów, kontrolowanych przez właścicieli platform, umożliwiłoby powstanie otwartego rynku technologicznych rozwiązań. W kontekście systemów rekomendacji doprowadziłoby to do usług zarządzania treścią, które funkcjonują na platformach, ale są od nich niezależne i oddają kontrolę w ręce użytkowników lub wybranych przez nich pośredników. Takie rozwiązania są niezbędne nie tylko dla zapewnienia rozwoju gospodarki opartej na danych i jej konkurencyjności, ale też podniesienia standardu ochrony praw użytkowników platform.

\*\*\*

**Podsumowując, nasze najważniejsze postulaty obejmują:**

1. Zwiększenie gwarancji dla użytkowników w procesie moderacji treści poprzez: wprowadzenie obowiązku notyfikacji przed zablokowaniem treści, rozszerzenie obowiązku notyfikacji i innych zasad sprawiedliwej procedury o przypadki ograniczenia widoczności treści, a nie tylko jej usunięcia, doprecyzowanie, że ścieżka pozasądowa przewidziana w DSA nie wyklucza możliwości skorzystania z sądowej drogi odwoławczej.
2. Doprecyzowanie zasad przejrzystości targetowania reklam internetowych poprzez obowiązek ujawnienia wszystkich, a nie tylko głównych kryteriów targetowania oraz ujawnienia roli platformy w procesie optymalizacji reklamy.
3. Wprowadzenie dodatkowych ograniczeń co do możliwości wykorzystywania niektórych metod i kryteriów targetowania.
4. Doprecyzowanie i rozszerzenie kryteriów przejrzystości systemów rekomendacji oraz wprowadzenie przejrzystości innych algorytmów zarządzania treścią wykorzystywanych przez platformy (algorytmy ograniczające widoczność postów, algorytmy optymalizacji reklam).
5. Wprowadzenie obowiązku zapewnienia pełnej interoperacyjności danych i protokołów przez bardzo duże platformy internetowe.