

Debaty

o współczesnym świecie

MUZEUUM
WARSZAWY 90 LAT
1936—2026



POLSKA DEBATUJE

MOWA NIENAWIŚCI

część 1.

Platformy społecznościowe powinny ponosić odpowiedzialność prawną za treści publikowane przez użytkowników



WSTĘP

Współczesna komunikacja opiera się na szybkości. Informacje rozchodzą się dziś nieporównywalnie szybciej niż kiedykolwiek wcześniej, a przestrzeń, w której je odbieramy, zmieniła się z miejsc fizycznych w strumienie treści dostępne z każdego urządzenia. Jeszcze niedawno informacje miały swój rytm – pojawiały się w gazetach, w telewizji, w rozmowach z ludźmi, których znaliśmy. Dzisiaj docierają do nas z wielu źródeł jednocześnie, często zanim zdążymy zadać sobie pytanie, skąd właściwie pochodzą i dlaczego widzimy akurat to, co widzimy.

Ten nowy sposób obiegu informacji sprawia, że łatwiej niż dawniej jest coś przekazać, ale trudniej zrozumieć, jak powstaje to, co do nas dociera. Treść nie funkcjonuje już w oderwaniu od formy – to, jak materiał wygląda, jak szybko krąży i jakie wzbudza emocje, stało się równie ważne jak to, co właściwie mówi. W efekcie informacje zyskują nie tylko własną treść, ale także własną dynamikę. Jedne rozplývają się bez echa, inne – często te najbardziej sugestywne albo najbardziej poruszające – potrafią zdobyć ogromną widoczność w bardzo krótkim czasie.

Z tego względu coraz trudniej jest oddzielić fakty od interpretacji, a interpretacje od treści przygotowanych tak, by wyglądały przekonująco, nawet jeśli rozmijają się z rzeczywistością. Wiele zjawisk, które dziś uważamy za nowe, w rzeczywistości mają długą historię – zmieniły się jedynie środki przekazu. Od wieków ludzie wykorzystywali narracje, obrazy i symbole, by wpływać na wyobraźnię innych. Dziś te same mechanizmy działają w środowisku cyfrowym, wzmocnione przez tempo obiegu treści i możliwości technologiczne, które pozwalają tworzyć materiały trudne do odróżnienia od autentycznych.

Współczesne treści funkcjonują w środowisku, w którym emocje, obrazy i natychmiastowość mogą nadać im znaczenie nieproporcjonalne do ich wiarygodności. Łatwość udostępniania i powtarzania sprawia, że nawet materiały niezweryfikowane mogą zyskać status „powszechnie znanych”, a odbiorcy nie zawsze mają czas lub możliwość, by przeprowadzić ocenę, która dawniej odbywała się bardziej naturalnie. Informacja staje się więc czymś więcej niż komunikatem – jest elementem szerszego krajobrazu, który kształtuje społeczne przekonania i nastroje.

Dlatego w dyskusji o odpowiedzialności za treści ważne jest nie tylko to, co zostaje opublikowane, ale także warunki, w jakich funkcjonuje współczesna komunikacja. Wyobrażenia odbiorców mogą być kształtowane przez materiały, które wyglądają jak relacje

z rzeczywistości, ale rzeczywistości nie przedstawiają. Treści mogą przybierać formy trudne do odróżnienia od autentyku. Strumienie informacji mogą być selektywne w sposób, którego użytkownicy nie widzą, a który wpływa na to, jak odbierają świat.

To wszystko powoduje, że pytania o odpowiedzialność – zarówno twórców, jak i platform – nie pojawiają się w próżni. Są reakcją na zmieniające się środowisko, w którym odbiorcy mają coraz mniej narzędzi, by samodzielnie ocenić wiarygodność materiałów, a jednocześnie coraz większą ekspozycję na treści, które mogą oddziaływać na ich decyzje, opinie czy emocje.

Warto przy tym pamiętać, że równie istotny fragment tego tematu dotyczy języka, który rani, wyklucza i podsycia przemoc symboliczną. Szczegółowe omówienie tego obszaru znajduje się w odrębnej publikacji: *„Należy karać za użycie mowy nienawiści w Internecie”*. To tam opisano mechanizmy mowy nienawiści, jej formy i konsekwencje, które stanowią ważne uzupełnienie niniejszego materiału.

FAKE NEWSY W HISTORII

Kiedy ludzie zaczęli utrwalać swoje opowieści – najpierw na ścianach jaskiń, później w piśmie – zmienił się nie tylko sposób komunikowania się, ale też wpływania na innych. Zapis dawał władzę nad pamięcią: pozwalał decydować, która wersja wydarzeń przetrwa, a która zniknie. Od samego początku z tej możliwości korzystali głównie rządzący, dlatego zapis historii rzadko był neutralnym odbiciem rzeczywistości.

Już w starożytności widzimy, że opowieść mogła być narzędziem propagandy. Ramzes II po bitwie pod Kadesz przedstawił siebie jako wielkiego zwycięzcę, choć konflikt prawdopodobnie zakończył się remisem. To jednak jego wersja, utrwalona w kamieniu, przetrwała. Podobnie w Rzymie Oktawian skutecznie zniszczył reputację Marka Antoniusza, stale powtarzając narrację o jego zdradzie i moralnym upadku. Fakty miały mniejsze znaczenie niż zasięg i powtarzalność przekazu.

Najbardziej dramatyczne skutki fake newsów pojawiały się jednak nie wtedy, gdy dotyczyły władców, lecz gdy uderzały w zwykłych ludzi. W 1475 roku w Trydencie zaginął chłopiec imieniem Simonino. Bez dowodów, niemal natychmiast, lokalny kaznodzieja Bernardino da Feltre oskarżył społeczność żydowską o rytualne morderstwo. Ta opowieść idealnie trafiała w istniejące uprzedzenia i lęki, dlatego rozprzestrzeniła się błyskawicznie.

Na podstawie plotki, która nigdy nie została potwierdzona, aresztowano niewinnych ludzi; część z nich stracono. Nawet gdy papieżstwo próbowało interweniować i podważyć oskarżenia, lokalne władze wolały podtrzymywać fałszywą narrację niż przyznać, że doszło do tragicznego błędu. Aby ją wzmocnić, dokładano kolejne „dowody”: opowieści o cudach, rzekome zeznania i legendy. Fake news stał się samonapędzającą się historią, której prawda nie była już potrzebna.

Ten przypadek pokazuje mechanizm, który pozostaje aktualny do dziś: emocjonalna, prosta opowieść wygrywa z faktami, zwłaszcza wtedy, gdy potwierdza wcześniejsze lęki i przekonania. Fake newsy nie narodziły się wraz z internetem. Istnieją od momentu, kiedy ludzie odkryli, że odpowiednio opowiedziana historia może stać się potężnym narzędziem władzy. Zmieniła się jedynie skala i prędkość ich rozprzestrzeniania – mechanizm pozostał ten sam.

OBRAZY, KTÓRE UDAJĄ RZECZYWISTOŚĆ

Manipulowanie obrazem istniało na długo przed tym, zanim komukolwiek śniły się komputery. Ludzie od dawna próbowali „poprawiać” rzeczywistość tak, żeby wyglądała trochę lepiej, trochę wygodniej albo po prostu bardziej po ich myśli. Władcy usuwali z fotografii lub portretów osoby, które popadły w niełaskę, artyści dodawali do rodzinnych zdjęć członków rodziny, których już z nimi nie było, a nożyczki, klej i cierpliwość potrafiły zrobić z obrazu coś, co wyglądało jak autentyk. To był taki analogowy Photoshop, zanim ktokolwiek znał to słowo.

Potem do gry weszły komputery, a wraz z nimi zupełnie nowy rozdział w historii fałszowania obrazu. Cyfrowe narzędzia sprawiły, że retusz zdjęć stał się łatwiejszy niż kiedykolwiek. Ludzie zaczęli masowo tworzyć przeróbki, memy, fotomontaże, a Internet tylko przyspieszył ten proces. Można było wrzucić obrazek w sieć i w ciągu kilku godzin zobaczyć, jak krąży po grupkach, forach i serwisach społecznościowych. W pewnym momencie wydawało nam się nawet, że zdjęcie to już „za mało”, żeby mieć pewność, że patrzymy na coś prawdziwego. Pozostawało wideo – ostatnie medium, które wydawało się odporne na oszustwa. Do czasu.

Technologia dogoniła wyobraźnię i pojawiły się deepfake’i. To filmy lub nagrania głosowe, w których twarz albo głos jednej osoby zostaje zastąpiony twarzą lub głosem kogoś innego w sposób, który ma wyglądać jak najbardziej autentycznie. Kiedyś te próby były nieudane i łatwe do wykrycia, ale wraz z rozwojem sztucznej inteligencji zaczęły wyglądać coraz bardziej przekonująco.

Co ciekawe, deepfake'i wcale nie są nowym wynalazkiem – pierwsze koncepcje pojawiły się już w latach 90. Na początku były jednak tak toporne, że trudno było je traktować poważnie. Dopiero połączenie ogromnej mocy obliczeniowej, nowoczesnych algorytmów uczenia maszynowego i gigantycznych baz danych doprowadziło do jakości, którą dziś widzimy w sieci. I choć wciąż istnieją drobne wskazówki pozwalające odróżnić fałsz od autentyku, wiele osób daje się na nie nabrać.

Zasada działania tej technologii jest w gruncie rzeczy zaskakująco podobna do filtrów, które znamy z Instagrama czy Snapchata. Program analizuje twarz: rozpoznaje oczy, nos, usta, obserwuje, jak się poruszają, jak mrugasz, kiedy unosisz brwi, a kiedy mówisz. To jest punkt wyjścia do późniejszej podmiany. Algorytm ogląda ogromne ilości nagrań osoby, której wizerunek będzie „nałożony”, żeby nauczyć się jej mimiki. Jeżeli aktor mrugnie, deepfake mrugnie tak samo. Jeżeli coś powie, sztuczna inteligencja dobierze taki ruch ust, jaki realnie wykonywałaby dana osoba. Sam obraz to jednak dopiero połowa problemu.

Drugą częścią jest głos. Kiedyś deepfake'i musiały korzystać z aktorów głosowych, którzy próbowali kogoś naśladować, ale teraz tę pracę również przejęła sztuczna inteligencja. Po przeanalizowaniu nagrań głosowych potrafi odtworzyć sposób mówienia, intonację, tempo, a nawet charakterystyczne pauzy. W efekcie powstaje coś, co brzmi jak czyjaś wypowiedź – mimo że ta osoba nigdy tego nie powiedziała.

Łącząc obraz i głos, otrzymujemy deepfake'i, które potrafią być naprawdę przekonujące. Widzieliśmy już fałszywe przemówienia polityków, filmy udające materiały newsowe czy przeróbki pokazujące celebrytów robiących rzeczy, których nigdy nie zrobili. W wielu przypadkach twórcy od razu wyjaśniają, że to eksperyment – ale wyobraźmy sobie sytuację, że ktoś publikuje wideo przedstawiające przywódcę kraju ogłaszającego atak zbrojny. Nawet jeśli prawda wyszłaby na jaw po godzinie, ta jedna godzina mogłaby wystarczyć, żeby wywołać panikę.

Konsekwencje nie ograniczają się jednak tylko do fake newsów. Deepfake'i są wykorzystywane także do szantażu, oszustw finansowych, a nawet do tworzenia fałszywej pornografii, często wbrew woli osób, które są na niej „umieszczane”. W niektórych przypadkach samo podejrzenie deepfake'a potrafi wywołać chaos. W Gabonie ludzie zaczęli wierzyć, że prezydent nie żyje, kiedy na opublikowanym nagraniu wyglądał „inaczej niż zwykle”. Była to zwykła zmiana wynikająca ze stanu zdrowia po udarze, ale plotki wystarczyły, aby wzniecić polityczne napięcia i doprowadzić do próby zamachu stanu.

Największy problem polega jednak na czymś innym – nie na tym, że coś może być fałszywe, ale że wszystko może być uznane za fałszywe. Jeżeli nagranie przestaje być wiarygodnym dowodem, to osoba przyłapana na czymś złym może po prostu powiedzieć, że film jest deepfake’iem. To zjawisko nazywa się „wiarygodnym zaprzeczeniem”. W takim świecie nie wystarczy już zobaczyć, żeby uwierzyć. Potrzebujemy nowych sposobów weryfikacji.

Pierwszym krokiem jest świadomość. Im bardziej ludzie znają tę technologię, tym mniej ufają każdej treści na pierwszy rzut oka. To element nowoczesnej edukacji medialnej. Drugim krokiem jest rozwój narzędzi, które wykrywają deepfake’i. One też korzystają z uczenia maszynowego i ścigają się z twórcami fałszywek w wyścigu przypominającym grę w kotka i myszkę. Trzecim możliwym rozwiązaniem może być znakowanie oryginalnych nagrań w momencie ich tworzenia, np. za pomocą technologii blockchain. Jeśli każde prawdziwe wideo miałoby swój cyfrowy „znak wodny”, łatwiej byłoby odróżnić je od fałszywek.

Deepfake’i zmieniają więc nie tylko sposób tworzenia informacji, ale również sposób, w jaki musimy te informacje oceniać. W czasach, gdy wszystko można podrobić, paradoksalnie największym wyzwaniem staje się nie oszukanie ludzi, ale odzyskanie zaufania do jakiegokolwiek formy dowodu.

JAK DZIAŁAJĄ SOCIAL MEDIA

Kiedy otwieramy aplikację na telefonie, rzadko myślimy o tym, co dzieje się „pod spodem”. Widzimy posty, filmiki, komentarze znajomych i nieznajomych – tak jakby po prostu pojawiały się same, jedna po drugiej. To wygodne złudzenie. W rzeczywistości social media nie działają jak tablica ogłoszeń, na której wszystko układa się chronologicznie. Bardziej przypominają ogromną, dynamiczną maszynę, która w każdej sekundzie decyduje, co zobaczymy jako pierwsze, co pominie, a co pokaże nam dopiero wtedy, gdy uzna, że jesteśmy na to „gotowi”.

Platformy analizują każdy nasz ruch – zatrzymanie na filmiku, przewinięcie posta, polubienie zdjęcia, komentarz pod memem. Nie po to, by nas śledzić w złej wierze, ale po to, by jak najlepiej przewidzieć, co nas zatrzyma na dłużej. Jeśli wcześniej obejrzelismy trzy filmiki o gotowaniu, dostaniemy więcej gotowania. Jeśli raz weszliśmy w ostry polityczny post, mamy szansę zobaczyć całą serię podobnych sporów. Rzadko zastanawiamy się, jak bardzo te drobne kliki zmieniają to, co później uznajemy za „co się dzieje na świecie”.

Z biegiem czasu każdy z nas zaczyna oglądać trochę inną wersję rzeczywistości. To, co widzi jedna osoba, może być zupełnie inne od tego, co widzi jej znajomy – mimo że obie osoby patrzą na „ten sam Internet”. W ten sposób powstają bańki informacyjne. Nie dlatego, że ktoś specjalnie je tworzy, ale dlatego, że algorytmy robią to, do czego zostały zaprojektowane: pokazują nam to, na co najprawdopodobniej zareagujemy. A że najłatwiej reagujemy na treści mocne, wyraziste i emocjonalne – to właśnie one mają najczęściej największą widoczność.

To wszystko jest jeszcze bardziej widoczne, kiedy spojrzymy na sposób działania wiralowych treści. Internet kocha szybkość. Post może obiecać cały świat w kilka minut, tylko dlatego że wywołał w kimś krótkie „wow”, „ale skandal” albo „muszę to wysłać dalej”. Nie trzeba go nawet dokładnie przeczytać. Wystarczy chwila. W efekcie treści, które budzą silne emocje – oburzenie, śmiech, szok, strach – mają przewagę nad tymi, które wymagają namysłu czy wyjaśnienia. To nie przypadek, tylko mechanizm, na którym platformy opierają swoją konstrukcję.

Do tego dochodzi kolejna warstwa, która często pozostaje niewidoczna: nie wszystkie treści są tworzone przez ludzi. W sieci działają boty, automatyczne konta i całe grupy profili zaprogramowanych po to, by wzmacniać określone przekazy albo sprawiać wrażenie, że jakiś temat jest bardziej popularny, niż jest w rzeczywistości. Kiedy przewijamy komentarze pod postem, trudno czasem zauważyć, że część reakcji czy odpowiedzi nie pochodzi od realnych osób. Dla użytkownika to wszystko wygląda tak samo – a jednak kształtuje odbiór świata w sposób, którego nie jesteśmy świadomi.

Social media nie są tylko miejscem rozmów – są także miejscem reklamy. Ale reklama w dzisiejszym wydaniu nie polega już na wyświetlaniu tego samego każdemu. Platformy potrafią dopasować przekaz do bardzo małych grup, a nawet do pojedynczych osób. Mogą pokazać jedną wersję komunikatu jednemu użytkownikowi, a zupełnie inną jego sąsiadowi, bo każdy z nich ma inny zestaw danych, zainteresowań i reakcji. To daje ogromne możliwości – i ogromną siłę oddziaływania. Jeden produkt, jedno hasło, a tysiące wersji dopasowanych pod odbiorców tak precyzyjnie, że często nawet nie zauważamy, że ktoś mówi do nas w sposób specjalnie przygotowany.

W tle tego wszystkiego jest jeszcze jeden ważny element: moderacja. Social media są ogromne. Każdego dnia pojawiają się miliardy nowych treści – zdjęć, filmów, komentarzy. Nie da się wszystkiego zobaczyć, przejrzeć i ocenić ręcznie. Część pracy wykonują algorytmy, które wychwytyują podejrzane wzorce. Część robią ludzie, próbując w ułamku

sekundy zdecydować, czy coś jest szkodliwe, czy po prostu kontrowersyjne. Część treści umyka ich uwadze albo zostaje uznana za nieszkodliwą. W efekcie Internet staje się miejscem, które żyje własnym tempem – szybszym, niż jakakolwiek grupa moderatorów czy systemy potrafią w pełni ogarnąć.

A kiedy połączymy to wszystko – algorytmy, emocje, wiralowość, personalizowane reklamy, boty, ogromną skalę i ograniczenia moderacji – dostajemy środowisko, które ma realny wpływ na to, jakie treści widzimy, w co wierzymy, o czym rozmawiamy i jak reagujemy. Social media stały się czymś więcej niż platformą. Są przestrzenią, która kształtuje naszą percepcję świata, nasze emocje, a czasem nawet nasze decyzje. I żeby zrozumieć współczesne dyskusje o odpowiedzialności, najpierw musimy dobrze zrozumieć właśnie to – jak działa ten mechanizm i jak wiele potrafi zmienić, zanim jeszcze ktokolwiek zacznie mówić o prawie czy regulacjach.

PRÓBY REGULACJI

Wraz z rosnącą rolą platform społecznościowych coraz więcej państw i instytucji próbuje odpowiedzieć na pytanie, jak ograniczyć szkody związane z obiegiem informacji – od dezinformacji, przez szkodliwe treści, po manipulację obrazem i głosem. Nie istnieje jeden uniwersalny model. Poszczególne kraje sięgają po różne rozwiązania, czasem bardzo stanowcze, a czasem ostrożne. Wspólny jest jednak kierunek: próba dostosowania systemów prawnych do środowiska, w którym informacja działa szybciej, mocniej i szerzej niż kiedykolwiek wcześniej.

W Unii Europejskiej najważniejszą próbą uporządkowania zasad jest Digital Services Act (DSA). To regulacja, która określa obowiązki platform w zakresie reagowania na treści nielegalne, przejrzystości działania systemów rekomendacji i oceny ryzyk związanych z wpływem platform na debatę publiczną oraz społeczeństwo. To podejście zakłada, że wpływ platform nie wynika wyłącznie z treści użytkowników, ale również z architektury ich działania.

Jeszcze zanim DSA stało się punktem odniesienia dla Europy, Niemcy wprowadziły własne rozwiązanie – NetzDG. To jedna z pierwszych ustaw na świecie nakazujących platformom szybkie usuwanie „oczywiście nielegalnych treści” w ciągu 24 godzin, pod groźbą bardzo wysokich kar. Choć regulacja została skrytykowana za ryzyko nadmiernego usuwania treści, wpłynęła na globalną dyskusję i stała się inspiracją dla wielu kolejnych inicjatyw.

Podobną ścieżką próbowała pójść Francja, wprowadzając tzw. ustawę Avia, nakazującą szybkie usuwanie treści nienawistnych. Jednak kluczowe przepisy zostały uznane przez francuski sąd konstytucyjny za nieproporcjonalne względem wolności wypowiedzi, co pokazało, jak trudne jest znalezienie równowagi między ochroną użytkowników a ochroną praw obywatelskich.

W Wielkiej Brytanii dyskusja skupiła się na bezpieczeństwie użytkowników w sieci. Online Safety Act nałożył na platformy obowiązki związane z przeciwdziałaniem treściom nielegalnym oraz ochroną dzieci przed szkodliwymi treściami, w tym konieczność stosowania skutecznej weryfikacji wieku. Brytyjskie podejście mocniej akcentuje ochronę użytkownika niż przejrzystość algorytmów, choć oba elementy pojawiają się w regulacji.

Stany Zjednoczone prowadzą debatę bardziej skoncentrowaną na fundamentalnych zasadach Internetu. Kluczowym punktem sporu jest Section 230 ustawy Communications Decency Act – przepis, który od lat chroni platformy przed odpowiedzialnością za treści użytkowników. W ostatnich latach pojawia się coraz więcej głosów podważających jego dotychczasową interpretację, a sądy oraz Kongres debatują nad tym, czy i jak powinien zostać ograniczony. Ta dyskusja jest trudna, bo dotyczy podstaw funkcjonowania amerykańskiej wolności słowa i równowagi między moderacją a odpowiedzialnością.

W Ameryce Południowej szczególnie ważny jest przypadek Brazylii, gdzie projekt tzw. „PL das Fake News” przewiduje szeroką odpowiedzialność platform za dezinformację oraz obowiązki związane z przejrzystością treści i identyfikacją użytkowników. Propozycje budzą kontrowersje – zwłaszcza ze względu na obawy o możliwe nadużycia i zbyt szerokie definicje dezinformacji – ale pokazują, że kraje mierzące się z kampaniami manipulacyjnymi szukają bardziej stanowczych narzędzi.

Równie zdecydowaną ścieżkę obrały Indie, gdzie w ramach zmian w przepisach dotyczących pośredników cyfrowych rząd zyskał możliwość wskazywania treści dotyczących działań władz jako „fałszywych”, co obliguje platformy do ich usunięcia pod groźbą utraty statusu chronionego podmiotu. Dyskusja wokół tych przepisów pokazuje napięcie między walką z dezinformacją a ryzykiem nadmiernego wpływu państwa na obieg informacji.

Choć poszczególne państwa podchodzą do tematu inaczej, istnieje kilka wspólnych wniosków płynących z tych prób regulacji. Po pierwsze – tempo obiegu informacji i siła, z jaką oddziałują treści, sprawiają, że dotychczasowe modele prawne w wielu miejscach okazują się niewystarczające. Po drugie – każda ingerencja w treści musi ważyć dwie

wartości: ochronę użytkowników i wolność wypowiedzi. Po trzecie – sama treść to tylko część zagadnienia. Równie ważne są mechanizmy jej rozpowszechniania: algorytmy, rekomendacje, personalizacja i zasięgi.

Próby regulacji nie są więc jedynie reakcją na pojedyncze zjawiska, ale próbą odpowiedzi na szerszą zmianę tego, jak funkcjonuje komunikacja. Państwa starają się uchwycić moment, w którym technologia zaczyna wpływać na rzeczywistość społeczną szybciej, niż instytucje są w stanie nadążyć. Nie ma jednego rozwiązania, ale każdy z powyższych przykładów pokazuje, że poszukiwanie równowagi między bezpieczeństwem, przejrzystością i wolnością słowa stało się jednym z kluczowych wyzwań współczesnego świata.

“Platformy społecznościowe powinny ponosić odpowiedzialność prawną za treści publikowane przez użytkowników.”

DRUŻYNA PROPOZYCJI (“ZA”)

Poniższa sekcja to materiał do pracy dla całej drużyny propozycji. Jej celem jest uporządkowanie przygotowań i ułatwienie budowy spójnego wystąpienia: od zdefiniowania problemu, przez uzasadnienie celu karania, po wskazanie skutecznych i proporcjonalnych rozwiązań. Lista ta nie wyczerpuje tematu - wskazuje kierunek pracy, podsuwa możliwe konstrukcje argumentów i ma ułatwić własne doprecyzowanie linii argumentacyjnej, przykładów i definicji.

Zaczynając przygotowanie do debaty zastanów się nad trzema podstawowymi aspektami

1. **Cel** - po co społeczeństwo ma reagować: jakie szkody realnie ograniczamy i jakie dobra wspólne wzmacniamy.
2. **Granica** - gdzie kończy się rola „biernego pośrednika”, a zaczyna **aktywny wpływ platform** na widoczność, zasięg i rozpowszechnianie treści.
3. **Skuteczność i proporcja** – jakie środki wobec platform są adekwatne i możliwe do wdrożenia, aby osiągnąć efekt bez nadmiernej kontroli nad komunikacją.

Pytanie 1. Jakie szkody uzasadniają wprowadzenie odpowiedzialności platform?

Pomyśl o trzech poziomach: indywidualnym (oszustwa, nękanie, naruszenia wizerunku), grupowym (wyciszanie mniejszości, polaryzacja), społecznym (destabilizacja informacji publicznej, erozja zaufania, wpływ na wybory).

Pytanie 2. Co chcemy osiągnąć poprzez odpowiedzialność?

Prewencję (zmniejszenie zasięgu szkodliwych treści), ochronę (szybsze zatrzymywanie szkody), przejrzystość (wiadomo, „co i dlaczego” widać w feedzie), sprawiedliwość proceduralną (realne ścieżki odwołań).

Pytanie 3. Co się dzieje, gdy nie reagujemy?

„Efekt bezkosztowości” – treści szkodliwe szerzą się viralowo, bo są opłacalne zasięgowo → spada zaufanie do informacji w ogóle → rosną nierówności informacyjne (różne grupy społeczne widzą całkowicie odmienne wersje świata).

Pytanie 4. Dlaczego odpowiedzialność platform nie musi naruszać wolności słowa?

Zastanów się nad różnicą między usuwaniem treści a **odpowiedzialnością za środowisko**, które platforma sama tworzy: projekt algorytmu, mechanizmy wzmacniania, rekomendacje, mikro-targetowanie.

To nie jest cenzura treści użytkowników, ale wymóg kontroli nad własnym wpływem.

Pytanie 5. Gdzie przebiega granica między neutralnym hostowaniem a współtworzeniem przekazu?

Weź pod uwagę: ranking treści zamiast układu chronologicznego, rekomendacje oparte na emocjach, promowanie virali, dopasowanie reklam do mikro-grup.

To platforma nadaje treściom tempo i skalę – warto ustalić, od którego momentu powinna za to odpowiadać.

Pytanie 6. Co zrobić z treściami „legalnymi, ale szkodliwymi”?

Nie wszystko, co jest dozwolone prawnie, jest neutralne w skali.

Rozważ, czy odpowiedzialność platform dotyczy tylko samego usunięcia treści, czy także zarządzania jej zasięgiem.

Pytanie 7. Jak definiować „wiedzę” platformy o bezprawności treści?

Zgłoszenia użytkowników i wiarygodnych podmiotów, orzeczenia i nakazy, powtarzalne wzorce naruszeń, „oczywista bezprawność” wprost z treści.

Pytanie 8. Za co dokładnie ma odpowiadać platforma wprost wobec treści użytkowników?

Za utrzymywanie i dalsze rozpowszechnianie treści bezprawnych po uzyskaniu wiarygodnej wiedzy o ich bezprawności; za zaniechanie podjęcia działań naprawczych; za niewdrożenie procedur pozwalających skutecznie zgłaszać i przeglądać takie treści.

Pytanie 9. Co z recydywą i kontami „seryjnych naruszcycieli”?

Czy platforma ma obowiązek identyfikować powtarzające się naruszenia i stosować gradację środków (ostrzeżenia, ograniczenia funkcji, zawieszenia)?

DRUŻYNA OPOZYCJI (“ZA”)

Poniższa sekcja to materiał do pracy dla całej drużyny propozycji. Jej celem jest uporządkowanie przygotowań i ułatwienie budowy spójnego wystąpienia: od zdefiniowania problemu, przez uzasadnienie celu karania, po wskazanie skutecznych i proporcjonalnych rozwiązań. Lista ta nie wyczerpuje tematu - wskazuje kierunek pracy, podsuwa możliwe konstrukcje argumentów i ma ułatwić własne doprecyzowanie linii argumentacyjnej, przykładów i definicji.

Pytanie 1. Jakie ryzyko niesie przypisanie platformom odpowiedzialności za treści stworzone przez użytkowników?

Zastanów się, czy oczekiwanie od platform prawnej odpowiedzialności nie wymusi nadmiernego usuwania treści na wszelki wypadek – co może tłumić debatę, satyrę, ostrą krytykę i głosy mniejszości.

Pytanie 2. Czy możliwe jest precyzyjne i sprawiedliwe określenie tego, które treści wymagają reakcji prawnej?

Weź pod uwagę trudność rozróżnienia: krytyka vs atak, ironia vs zniesławienie, mem vs manipulacja. Platformy podejmują tysiące decyzji na minutę – margines błędu jest nieunikniony.

Pytanie 3. Jakie mogą być niezamierzone konsekwencje takiej regulacji dla wolności słowa?

Czy realnie istnieje ryzyko „efektu mrozącego”, w którym platformy – bojąc się odpowiedzialności – zaczną kasować nie tylko treści bezprawne, ale też te po prostu kontrowersyjne.

Pytanie 4. Czy technicznie możliwe jest monitorowanie i ocenianie każdej publikowanej treści?

Treści przyrastają szybciej, niż ludzie czy algorytmy są w stanie je ocenić. Rozważ, czy wymóg odpowiedzialności nie tworzy oczekiwań niewykonalnych w praktyce.

Pytanie 5. Czy odpowiedzialność platform nie zniechęci mniejszych firm do konkurowania?

Zastanów się, czy wysokie wymagania prawne, koszty moderacji i ryzyko sankcji nie doprowadzą do tego, że tylko największe firmy będą w stanie funkcjonować – co wzmacnia monopol, a nie bezpieczeństwo.

Pytanie 6. Jakie istnieje ryzyko nadużycia narzędzi zgłaszania treści?

Pomyśl o masowych zgłoszeniach wykorzystywanych jako forma cichej cenzury – grupy mogą organizować się, by usuwać treści polityczne, społeczne lub światopoglądowe, z którymi się nie zgadzają.

Pytanie 7. Czy platforma powinna być prawnie odpowiedzialna za treści, których sama nie stworzyła i których tworzenia nie może w pełni kontrolować?

Zastanów się, czy odpowiedzialność za cudze słowa nie stoi w sprzeczności z podstawową logiką komunikacji internetowej, gdzie to użytkownicy są nadawcami.

Pytanie 8. Czy skupienie się na odpowiedzialności platform nie odciąga uwagi od rzeczywistych sprawców szkody?

Można argumentować, że to autor treści – osoba publikująca – wie najlepiej, z jaką intencją działa, a platforma jedynie „udostępnia półkę”. Czy przenoszenie odpowiedzialności nie osłabia skuteczności prawa?

Pytanie 9. Jak chronić proces demokratyczny przed ryzykiem politycznego nacisku na platformy?

Wprowadzenie odpowiedzialności może stworzyć presję ze strony rządów, żeby usuwać treści niewygodne, ale legalne – pod pozorem „ryzyka”. Rozważ, czy nie jest to zagrożenie większe niż brak regulacji.

Pytanie 10. Czy użytkownicy nie utracą zaufania do otwartości platform, jeśli te będą działać jak redakcje?

Opozycja może podnieść argument, że platformy mają zapewniać przestrzeń wymiany myśli, a nie odgrywać rolę wydawców treści. Odpowiedzialność podobna do odpowiedzialności redakcji może zmienić charakter medium.

Organizator:



Partner:



Współfinansowanie:



Partner Edukacji
Muzeum Warszawy:



Patron
medialny:



TVP | info_