



25 sierpnia 2026

Sz.P. Wicepremier Krzysztof Gawkowski
Minister ds. Cyfryzacji
ul. Królewska 27
00-060 Warszawa

Szanowny Panie Premierze,

dziękujemy za list z dnia 18 sierpnia 2026 r. Traktujemy poruszone przez Pana kwestie z najwyższą powagą i zdajemy sobie sprawę, że mają one pilny charakter. Cenimy naszą współpracę z Polską w budowaniu bezpiecznego środowiska cyfrowego.

Zwalczanie oszukańczych reklam jest jednym z najbardziej złożonych wyzwań stojących obecnie przed ekosystemami cyfrowymi, wymagającym stałej współpracy między platformami, rządzącymi i organami ścigania. Doceniamy możliwość dalszego wyjaśnienia szeroko zakrojonych działań, które podejmujemy – i nieustannie udoskonalamy – w celu zapobiegania nadużywaniu naszych platform do rozpowszechniania oszukańczych reklam.

Na początek chcielibyśmy odnieść się do najważniejszej kwestii: zaufanie naszych użytkowników i reklamodawców, zarówno w Polsce, jak i na całym świecie, ma kluczowe znaczenie dla działalności firmy Meta i jej długoterminowego sukcesu. Nieuczciwe działania na naszych platformach podważają to zaufanie, a my nie dążymy do czerpania z nich zysku; nasze warunki, zasady i systemy moderacji treści mają na celu ograniczenie rozpowszechniania nieuczciwych treści, a co za tym idzie – wszelkich związanych z nimi przychodów. Mówiąc wprost, Meta nie jest zainteresowana oszukańczą reklamą: nielegalne treści tego rodzaju szkodzą osobom korzystającym z Facebooka i Instagrama, uczciwie działającym firmom w Polsce i na całym świecie, które się u nas reklamują, a także samej Mecie. Nasze motywacje są ukierunkowane na wykrywanie i usuwanie takich treści z naszych platform. Dlatego z zadowoleniem przyjmujemy możliwość współpracy z organami w Polsce, w całej Europie i w innych częściach świata w zwalczaniu prób nadużywania naszych platform przez nieuczciwych graczy.

W niniejszym liście szczegółowo opisujemy nasze podejście do tego złożonego tematu, w tym wdrażane przez nas systemy, sposób ich dostosowywania w miarę ewolucji nieuczciwych taktów oraz rezultaty, które dzięki naszym działaniom osiągamy. Jak przedstawiono poniżej, publikujemy raporty dotyczące przejrzystości naszych platform internetowych zgodnie z Aktem o usługach

Meta Platforms Ireland Limited

Registered Office:

Merrion Road

Dublin 4, D04 X2K5

Registered in Ireland as a private company limited
by shares

Company No. 462932

Directors: Majella Goss, Yvonne Cunnane, Anne O'Leary,
Genevieve Hughes

www.meta.com

cyfrowych („DSA”), z których wynika, że w okresie od lipca do grudnia 2025 r. liczba podjętych z własnej inicjatywy działań polegających na usunięciu treści dotyczących „oszustw i wyłudzeń” w Unii Europejskiej wyniosła łącznie 12 656 725 na Facebooku (z czego 86% przeprowadzono automatycznie) oraz 9 257 089 na Instagramie (z czego 88% przeprowadzono automatycznie). Możemy również ujawnić, z zachowaniem ścisłej poufności, że w okresie od lipca 2025 r. do czerwca 2026 r. usunięto 380 000 pochodzących z Polski przykładów treści na Facebooku, Instagramie i Threads naruszających nasze zasady Standardy społeczności dotyczące oszustw i innych oszukańczych praktyk oraz Politykę niedopuszczalnych praktyk biznesowych. Spośród nich ponad 93% zostało usunięte, zanim zgłosili je użytkownicy. W tym samym okresie usunęliśmy 137 000 pochodzących z Polski *reklam* z powodu naruszenia naszych zasad dotyczących oszustw i wyłudzeń. Spośród nich ponad 88% zostało usuniętych, zanim zgłosili je użytkownicy.

Przedstawiamy również plan działania Meta, który zakłada podjęcie dalszych kroków poprzez rozszerzenie weryfikacji reklamodawców, aby do końca 2026 r. 90% przychodów Meta z reklam pochodziło od zweryfikowanych reklamodawców na całym świecie, co ma na celu promowanie większej przejrzystości i ograniczenie prób fałszywego przedstawiania tożsamości reklamodawców.

Biorąc pod uwagę złożoność zagrożeń i ich wieloaspektowość, bardzo chętnie spotkamy się z Panem Wicepremierem i Pana zespołem, aby przedstawić dodatkowy kontekst wykraczający poza to, co można zawrzeć w pisemnej odpowiedzi: pełniejszy obraz tego, w jaki sposób współpraca międzysektorowa – w tym z polskimi władzami – może wzmocnić naszą wspólną obronę. Uważamy, że tego rodzaju wymiana poglądów przysłuży się dyskusji i pomoże nam zidentyfikować konkretne obszary do ściślejszej współpracy.

Zgodność działań Meta z Aktem o usługach cyfrowych

Zapewniamy, że Meta regularnie i konstruktywnie współpracuje z Komisją Europejską oraz irlandzką Coimisiún na Meán w zakresie przestrzegania przez nas przepisów DSA, w tym naszych obowiązków dotyczących oceny ryzyka i ograniczania ryzyka wynikających z art. 34 i 35 DSA. Zarówno Komisji, jak i Coimisiún na Meán, przekazujemy roczne sprawozdania z oceny ryzyka systemowego, które przygotowujemy zgodnie z art. 42 DSA, a także publikujemy te sprawozdania w naszym [Centrum Transparentności](#). Nasza ocena ryzyka uwzględnia różnorodność regionalną i językową w całej UE, a działania ograniczające ryzyko opierają się na informacjach otrzymywanych od organów w państwach członkowskich. Poniżej przedstawiamy szeroki zakres środków ograniczających ryzyko, które wdrożyliśmy w całej UE, w tym w Polsce.

Oprócz raportów dotyczących oceny ryzyka systemowego, publikujemy również raporty dotyczące przejrzystości naszych platform zgodnie z DSA. W sprawozdaniach tych przedstawiamy, w odniesieniu do całej UE oraz w podziale na kategorie nielegalnych treści lub naruszeń naszego regulaminu (w szczególności „oszustwa i wyłudzenia”), liczbę interwencji podejmowanych z własnej inicjatywy oraz stopień ich automatyzacji, liczbę zgłoszeń otrzymywanych od użytkowników za pośrednictwem naszego kanału zgłaszania nielegalnych treści, medianę czasu potrzebnego na podjęcie działań w odpowiedzi na te zgłoszenia, liczbę otrzymanych skarg i ich wyniki, a także medianę czasu potrzebnego na wykonanie nakazów pochodzących od organów państw członkowskich.

Zgodnie z naszymi zobowiązaniami wynikającymi z art. 39 DSA udostępniamy również publicznie – za pośrednictwem narzędzia umożliwiającego wyszukiwanie ([Ad Library](#)) oraz [interfejsu API](#) – treść,

reklamodawcę, płatnika, parametry targetowania oraz zasięg każdej reklamy wyświetlanej w UE (w tym w Polsce), w okresie wyświetlania reklamy oraz przez rok po jego zakończeniu.

Nasze zabezpieczenia przed oszukańczymi reklamami

Wykrywanie i egzekwowanie

Wskaźniki wykrywania i usuwania treści określone w naszych raportach dotyczących przejrzystości są osiąganе dzięki naszym ciągłym inwestycjom w technologie, egzekwowanie przepisów i współpracę, mające na celu ochronę osób fizycznych i prawnych przed oszustwami internetowymi oraz ochronę naszych platform przed nadużyciami ze strony nieuczciwych graczy. Rezultatem tych inwestycji są systemy, w ramach których wszystkie tworzone reklamy są przed opublikowaniem sprawdzane pod kątem zgodności z naszymi [standardami reklamowymi](#) automatycznie, a w razie potrzeby ręcznie; nasze modele sztucznej inteligencji i uczenia maszynowego proaktywnie wykrywają treści oszukańcze oraz podejrzane zachowania reklamodawców, a reklamy, które możemy jednoznacznie zidentyfikować jako oszustwa, są blokowane, zanim w ogóle trafią do aukcji reklamowej.

Zakaz stosowania deepfake'ów, fałszywych rekomendacji i nielegalnie wykorzystywanych zdjęć

Choć nie możemy omawiać spraw dotyczących indywidualnych kont użytkowników ani związanych z toczącymi się postępowaniami, podkreślamy, że takie praktyki jak wykorzystywanie deepfake'ów, fałszywych rekomendacji i nielegalnie wykorzystywanych zdjęć są zabronione przez naszą [Politykę niedopuszczalnych praktyk biznesowych](#). Gdy jednak dowiemy się o takich treściach, podejmujemy zarówno natychmiastowe działania w celu ich usunięcia, jak i odpowiednie kroki wobec kont i prowadzonej przez nie działalności. Zdajemy sobie sprawę, że osoby stojące za tego rodzaju działaniami dysponują zaawansowanymi technologiami i działają z dużą determinacją, dlatego stale udoskonalamy nasze narzędzia i metody, aby wyprzedzać stosowane przez nie taktyki.

Wobec osób publicznych, które wyraziły na to zgodę, stosujemy [technologię rozpoznawania twarzy](#) w celu porównania reklamy ze zdjęciami profilowymi danej osoby, jeśli mamy podejrzenie, że reklama w sposób nieuprawniony wykorzystuje jej wizerunek. Jeśli potwierdzimy zgodność wizerunku, ale stwierdzimy, że reklama stanowi oszustwo, zostanie ona odrzucona przed wyświetleniem (lub zostanie usunięta, jeśli wykryto ją później). Uprawnionym podmiotom, w tym polskim bankom i dużym przedsiębiorstwom, udostępniamy również nasze narzędzie [ochrony praw marki](#), umożliwiając im identyfikowanie i zgłaszanie reklam oraz innych treści nadużywających ich znaków towarowych lub innych oznaczeń marki, które my następnie weryfikujemy i usuwamy. W ciągu ostatnich dwóch lat intensywnie promowaliśmy korzystanie z tego narzędzia poprzez liczne inicjatywy edukacyjne realizowane we współpracy z polskimi organami (w szczególności z Urzędem Komisji Nadzoru Finansowego) oraz organizacjami branżowymi, takimi jak Związek Banków Polskich.

Ciągłe egzekwowanie przepisów

Gdy dowiemy się, że na naszych platformach zostały opublikowane treści naruszające zasady, stale reagujemy, usuwając niezgodne z przepisami reklamy i blokując konta stojące za organizacjami oszustów. Walka z tym zagrożeniem musi być prowadzona nieprzerwanie: sprawcy nieustannie

modyfikują swoje taktyki i aktywnie poszukują nowych sposobów obchodzenia wprowadzanych przez nas zabezpieczeń.

Nasze zautomatyzowane systemy wykrywania ciągle się uczą i są na bieżąco aktualizowane, aby przeciwdziałać pojawiającym się technikom omijania zabezpieczeń. Na przykład, nasza technologia dopasowywania automatycznie wykrywa i eliminuje nowe oraz historyczne kopie treści, które już zidentyfikowaliśmy jako naruszające zasady, w tym obrazy, filmy, teksty i adresy URL, podczas gdy nasze klasyfikatory oparte na uczeniu maszynowym i modelach językowych oceniają wzorce wizualne, język i kontekst wcześniej nieznanych treści reklamowych, aby przewidzieć, czy istnieje prawdopodobieństwo, że są one oszukańcze. Uzupełnieniem tych rozwiązań są reguły heurystyczne, które łączą wiele sygnałów (takich jak wyniki klasyfikatorów, słowa kluczowe i pochodzenie geograficzne treści), a każda z nich jest nieustannie ponownie szkolona i udoskonalana w miarę pojawiania się nowych taktyk omijania zabezpieczeń.

Kampanie przeciwko oszustwom

Działając na naszych platformach, zdajemy sobie sprawę, że oszustwa i wyłudzenia stanowią wyzwanie dla całej branży, z którym żaden podmiot nie jest w stanie poradzić sobie samodzielnie. Dlatego aktywnie współpracujemy z partnerami branżowymi – w tym ze Związkiem Banków Polskich – a także z osobami publicznymi. Poprzez wspólne kampanie edukacyjne dla konsumentów podnosimy świadomość społeczną i pomagamy użytkownikom w Polsce rozpoznawać oszustwa, chronić się przed nimi oraz je zgłaszać. Zapewniamy również użytkownikom solidne, łatwo dostępne narzędzia do zgłaszania podejrzanych reklam, kont i treści, a także oferujemy wskazówki dotyczące odzyskiwania kont i bezpieczeństwa za pośrednictwem dedykowanych kanałów wsparcia. W szczególności uruchomiliśmy specjalne [Centrum Ochrony przed Oszustwami](#), które stanowi dla użytkowników miejsce, gdzie mogą zgłaszać oszustwa popełniane na naszych platformach, uzyskać dostęp do wskazówek dotyczących rozpoznawania i unikania oszustw oraz lepiej zrozumieć nasze działania mające na celu ich zwalczanie.

Współpraca z polskimi organami

Od wielu lat ściśle współpracujemy z kluczowymi polskimi jednostkami organów ścigania oraz szerokim gronem polskich instytucji publicznych. Za pośrednictwem naszego dedykowanego zespołu ds. współpracy z organami ścigania (Law Enforcement Outreach) przeprowadziliśmy specjalistyczne szkolenia dla organów ścigania, prokuratorów i sędziów dotyczące procesów Meta oraz kanałów składania wniosków, umożliwiając im żądanie danych od Meta zgodnie z obowiązującymi procedurami. Pozostajemy zaangażowani w dalsze udostępnianie tego wsparcia w przyszłości. Utrzymujemy bezpośredni i stały kontakt z kluczowymi organami, w tym z Centralnym Biurem Zwalczania Cyberprzestępczości, w celu wymiany informacji na temat najnowszych wydarzeń, trendów i procesów, a Urząd Komisji Nadzoru Finansowego oraz NASK mają dostęp do naszych dedykowanych kanałów eskalacji i z nich korzystają.

Odpowiadamy również na właściwe wnioski od organów ścigania zgodnie z obowiązującym prawem i Regulaminem Mety. Meta oferuje również organom ścigania dostęp do internetowego systemu składania wniosków (*Law Enforcement Online Requests*), który umożliwia łatwe i szybkie składanie oraz przetwarzanie wniosków organów ścigania dotyczących danych użytkowników oraz wniosków o usunięcie treści.

Współpraca z organami ścigania na całym świecie

Współpracujemy również z organami ścigania na całym świecie, aby ścigać wysoce wyspecjalizowanych i dysponujących znacznymi zasobami sprawców. Jesteśmy zaangażowani w zwalczanie oszustw na naszych platformach oraz podejmowanie działań przeciwko przestępcom stojącym za nimi. W ciągu ostatnich trzech lat zbudowaliśmy zespoły i systemy, które - zgodnie z naszymi zasadami - pomagają wykrywać i zwalczać ogniska oszustw na całym świecie. Nasze zespoły nieustannie monitorują pojawianie się nowych takich ośrodków i rutynowo zakłócają ich działalność. W 2025 r. usunęliśmy na całym świecie ponad 159 milionów oszukańczych reklam, które naruszały nasze zasady, a także zlikwidowaliśmy 10,9 miliona kont na Facebooku i Instagramie powiązanych z przestępczymi ośrodkami oszustw.

Środki weryfikacji reklam

W liście zwrócił się Pan o szczegółowe przedstawienie naszych procedur potwierdzania tożsamości reklamodawców. Aby zapewnić bezpieczeństwo społeczności, Meta stosuje standardy dotyczące reklamodawców, w tym zasady i wymagania dotyczące weryfikacji danych dotyczących ich działalności gospodarczej lub tożsamości w przypadku osób fizycznych. Meta stosuje podejście oparte na ocenie ryzyka w zakresie weryfikacji reklamodawców, wykorzystując sygnały do identyfikacji i blokowania reklamodawców przed wyświetlaniem reklam do czasu zweryfikowania ich tożsamości. Oferujemy również wiarygodnym reklamodawcom możliwość proaktywnej weryfikacji. Wymagania weryfikacyjne są bardziej rygorystyczne w przypadku reklamodawców i rodzajów reklam, co do których podejrzewamy, że stanowią większe ryzyko dla użytkowników – na przykład tych publikujących treści w kategoriach reklamowych bardziej narażonych na nadużycia, takich jak kryptowaluty, hazard, farmaceutyki oraz polityka, wybory czy kwestie społeczne. Obecnie reklamodawcy mogą zostać zobowiązani do weryfikacji w zależności od takich czynników, jak rynek, na którym kierują reklamy, oraz to, czy wykazali potencjalnie ryzykowne zachowania (np. czy prowadzą określone rodzaje reklam lub czy prowadzone przez nich reklamy są bardziej podatne na nadużycia). Oceniamy takie czynniki, jak od kiedy konto jest aktywne, wykorzystywane przezeń funkcje oraz rodzaje wyświetlanych reklam. W przypadku wyższego ryzyka reklamodawca może zostać poproszony o wykonanie dodatkowych zadań weryfikacyjnych, począwszy od *captcha* (w celu wyeliminowania kont zautomatyzowanych lub botów), aż po weryfikację dokumentów rejestracyjnych firmy czy dokumentów tożsamości. Przed opublikowaniem reklam reklamodawcy mogą być zobowiązani do posiadania zweryfikowanego numeru telefonu powiązanego z kontem reklamowym. Wymaga to od administratora konta potwierdzenia kodu wysłanego przez Meta za pomocą wiadomości SMS lub połączenia głosowego w celu zweryfikowania numeru telefonu, zanim reklamy będą mogły zostać opublikowane. Ten proces weryfikacji wymaga od reklamodawców potwierdzenia danych dotyczących ich samych lub organizacji korzystającej z konta reklamowego. Jest to jeden ze środków, które pomagają nam zapobiegać nadużyciom i oszustwom ze strony nieautentycznych podmiotów.

Wprowadzamy te wymagania stopniowo, skupiając się przede wszystkim na reklamodawcach wyświetlających reklamy w kategoriach najbardziej narażonych na ataki oszustów – takich jak reklamy dotyczące usług i inwestycji finansowych – oraz na tych, których konta wykazują potencjalnie podejrzane zachowania. Jak wcześniej wspomnieliśmy, naszym celem jest, aby do końca 2026 r. zweryfikowani reklamodawcy generowali 90% globalnych przychodów Meta z reklam, w porównaniu z obecnym poziomem 70%. Rozszerzenie tych wymogów sprzyja większej przejrzystości, ograniczając próby fałszywego przedstawiania tożsamości reklamodawców. Stanowi to ważny element naszego wielopoziomowego podejścia do ochrony użytkowników naszych platform przed oszustwami.

Dane dotyczące usuwania treści

Poniżej wskazujemy najbardziej aktualne dane pochodzące z naszych raportów dotyczących przejrzystości od lipca do grudnia 2025 r. Chcemy być jak najbardziej pomocni i odpowiadamy na Państwa pytania w dobrej wierze, działając w duchu współpracy.

Dla kontekstu wskazujemy, że w naszych procesach dotyczących zaufania i bezpieczeństwa oraz moderacji treści, w tym w różnych formach sprawozdawczości regulacyjnej wykorzystujemy ustalone miary, wskaźniki i dane. Dostarczanie szczegółowych i rozdrobnionych danych jest na ogół trudne, a nasze metody pomiaru w sprawozdawczości zewnętrznej ograniczają się zazwyczaj do pewnych kluczowych wskaźników, w tym tych, do prowadzenia których jesteśmy zobowiązani na mocy prawa. Niektóre dane odpowiadające na Państwa pytania są gromadzone i generowane przez nasze systemy oraz zespoły w celu uwzględnienia ich w raportach dotyczących przejrzystości, które publikujemy zgodnie z DSA, jednak dane te są przygotowywane zgodnie z ramami określonymi przez Komisję Europejską. W rezultacie nasze raporty dotyczące przejrzystości zawierają pewne dane na poziomie państw członkowskich (takie jak np. wnioski otrzymane od organów krajowych), ale ramy sprawozdawcze nie wymagają, aby wszystkie dane były rozbite według wszystkich określonych przez Państwa filtrów.

Chcielibyśmy jednak zwrócić Państwa uwagę na następujące istotne wskaźniki dotyczące UE, pochodzące z naszych raportów za okres sprawozdawczy od lipca do grudnia 2025 r.:

- W tym okresie nasze systemy podjęły na Facebooku łącznie 171,8 mln działań moderacyjnych z własnej inicjatywy, z czego 148,9 mln zostało wykrytych za pomocą środków automatycznych.
- W kategorii oszustw i nadużyć podjęliśmy 12 656 725 działań z własnej inicjatywy na Facebooku oraz 9 257 089 na Instagramie. Obejmowało to usunięcie 978 643 przykładów treści związanych z phishingiem na Facebooku i 66 567 na Instagramie, a także usunięcie 1 438 887 fałszywych kont na Facebooku i 1 361 394 na Instagramie.
- Otrzymaliśmy 788 745 zgłoszeń na podstawie art. 16 DSA na Facebooku oraz 353 997 na Instagramie. Spośród otrzymaliśmy 1 732 zgłoszeń od zaufanych podmiotów sygnalizujących dotyczących oszustw i wprowadzania w błąd na Facebooku i 95 na Instagramie, przy medianie czasu rozpatrzenia wynoszącej odpowiednio 9,8 godz. i 18,32 godz.

Wskazujemy, że są to najbardziej aktualne dane, jakie jesteśmy w stanie przedstawić, ponieważ nasze raporty dotyczące przejrzystości za okres od stycznia do czerwca 2026 r. zostaną opublikowane dopiero w tym miesiącu. Chętnie przedstawimy bardziej aktualne i istotne wskaźniki przed naszym spotkaniem.

Podsumowanie

Chcemy powiedzieć wprost: nie ma jednego rozwiązania, które pozwoliłoby całkowicie wyeliminować oszukańcze reklamy. Właśnie dlatego nasze podejście, opisane powyżej, jest wielopoziomowe i nieustannie ewoluuje: kontrola przed publikacją, ciągłe działania mające na celu usuwanie opublikowanych treści, weryfikacja reklamodawców, blokowanie sieci oraz partnerstwa międzysektorowe – każde z tych działań dotyczy innego wymiaru problemu. Żadne pojedyncze działanie nie jest wystarczające samo w sobie, ale razem podnoszą one koszty działalności nieuczciwych podmiotów na każdym etapie – sprawiając, że wykorzystywanie naszych platform

staje się coraz trudniejsze i mniej opłacalne. Chcemy również jasno powiedzieć: zagrożenie to ma z natury charakter ciągły – w miarę jak wzmacniamy nasze zabezpieczenia, podmioty stojące za tymi oszustwami dostosowują swoje taktyki i poszukują nowych sposobów na ich obejście. Dlatego nasze działania w tej kwestii nigdy się nie kończą. Środki opisane w niniejszym piśmie nie są wyczerpujące, a nasze zespoły ciężko pracują i są w pełni zaangażowane w ciągłe udoskonalanie i rozszerzanie naszych działań w miarę ewolucji zagrożeń. Doceniamy możliwość dalszej współpracy w tym zakresie z polskimi władzami, w tym z Państwem.

Mamy nadzieję, że niniejszy list świadczy zarówno o tym, jak poważnie podchodzimy do opisanych wyzwań, jak i o naszym zaangażowaniu na rzecz przejrzystości. Podzielamy Państwa dążenie do ochrony użytkowników w Polsce przed oszukańczymi reklamami i mamy świadomość, że jego realizacja wymaga stałej współpracy między platformami, rządami i organami ścigania. Z przyjemnością spotkamy się z Panem Wicepremierem i Pana zespołem w dogodnym dla Państwa terminie, aby omówić nasze dotychczasowe działania i plany na przyszłość, a także określić konkretne sposoby zacieśnienia naszej współpracy.

Z wyrazami szacunku,

Meta Platforms Ireland Limited

ZAŁĄCZNIK

1. Skala oszukańczych reklam kierowanych do użytkowników w Polsce, w szczególności reklam wykorzystujących bezprawnie wizerunek osób publicznych i przedsiębiorców.

Zgodnie z naszymi zobowiązaniami wynikającymi z art. 39 DSA, który nakłada obowiązek publicznego udostępniania – za pośrednictwem narzędzia umożliwiającego wyszukiwanie ([Ad Library](#)) oraz interfejsu [API](#) – treści, danych reklamodawcy, płatnika, parametrów targetowania oraz zasięgu każdej reklamy wyświetlanej w UE (w tym w Polsce), w okresie wyświetlania reklamy oraz przez rok po jej zakończeniu. Możemy ujawnić, z zachowaniem ścisłej poufności, że w okresie od lipca 2025 r. do czerwca 2026 r. usunęliśmy około 380 000 przykładów treści pochodzących z Polski z serwisów Facebook, Instagram i Threads z powodu naruszenia naszych Standardów społeczności dotyczących oszustw i wyłudzeń oraz Polityki niedopuszczalnych praktyk biznesowych. Spośród nich ponad 93% zostało usunięte, zanim zgłosił je użytkownik. W tym samym okresie usunięto 137 000 pochodzących z Polski reklam z powodu naruszenia tych samych zasad i polityk. Spośród nich ponad 88% zostało usuniętych, zanim zgłosił je użytkownik.

2. Mechanizmy stosowane przez Meta do automatycznego wykrywania i blokowania reklam wykorzystujących deepfake'i, fałszywe rekomendacje oraz bezprawnie wykorzystany wizerunek.

Wszystkie reklamy tworzone przez użytkowników są sprawdzane automatycznie, a w razie potrzeby ręcznie, pod kątem zgodności z naszymi standardami reklamowymi przed ich opublikowaniem. Nasze modele sztucznej inteligencji i uczenia maszynowego proaktywnie wykrywają treści oszukańcze oraz podejrzane zachowania reklamodawców, a reklamy, które możemy jednoznacznie zidentyfikować jako oszustwa, są blokowane, zanim trafią do aukcji reklamowej. Wobec osób publicznych, które wyraziły na to zgodę, stosujemy technologię rozpoznawania twarzy w celu porównania reklamy ze zdjęciami profilowymi tej osoby, jeśli mamy podejrzenie, że reklama w sposób nieuprawniony wykorzystuje jej wizerunek; w przypadku potwierdzenia dopasowania i stwierdzenia, że reklama stanowi oszustwo, jest ona odrzucana przed wyświetleniem (lub usuwana, jeśli zostanie wykryta później). Jak opisano powyżej, uprawnionym podmiotom, w tym polskim bankom i dużym przedsiębiorstwom, udostępniamy również nasze narzędzie ochrony praw marki, umożliwiając im identyfikację i zgłaszanie treści, które nadużywają ich znaków towarowych lub innych oznaczeń marki, a my następnie je weryfikujemy i usuwamy.

3. Procedury weryfikacji reklamodawców oraz sposób potwierdzania ich tożsamości przed dopuszczeniem reklam do emisji.

Obecnie od reklamodawców może być wymagana weryfikacja w zależności od takich czynników, jak miejsce wyświetlania reklam, historia potencjalnie ryzykownych zachowań lub to, czy rodzaje wyświetlanych przez nich reklam są bardziej podatne na nadużycia. Aby określić poziom ryzyka, bierzemy pod uwagę takie czynniki, jak czas od kiedy konto jest aktywne, wykorzystywane funkcje oraz rodzaje wyświetlanych reklam. Ten proces weryfikacji oznacza, że reklamodawcy muszą potwierdzić dane dotyczące siebie lub organizacji stojących za ich kontami reklamowymi. Pomaga nam to zapobiegać nadużyciom i oszustwom ze strony nieuczciwych podmiotów.

Kontynuujemy nasze wysiłki rozszerzając wymagania weryfikacyjne dla reklamodawców. Rozszerzenie to jest wdrażane etapowo, począwszy od reklamodawców wyświetlających reklamy na wybranych rynkach oraz w kategoriach reklam, które są bardziej narażone na ataki oszustów, takich jak reklamy inwestycji finansowych. Rozszerzamy weryfikację reklamodawców, aby zapewnić, że do końca 2026 r. zweryfikowani reklamodawcy będą generować 90% naszych przychodów z reklam, w porównaniu z obecnym poziomem 70%. Skupiamy się na weryfikacji reklamodawców o najwyższym ryzyku, natomiast pozostałe 10% będzie pochodzić od przedsiębiorstw o niskim ryzyku. Proces weryfikacji pomaga promować większą przejrzystość, ograniczając próby fałszywego przedstawiania tożsamości reklamodawcy.

4. Średni czas od zgłoszenia oszukańczej reklamy do jej usunięcia lub zablokowania dostępu do niej.

W okresie od lipca do grudnia 2025 r. mediana czasu reakcji na wszystkie treści zgłoszone przez użytkowników za pośrednictwem naszego kanału zgłaszania treści niezgodnych z prawem na podstawie art. 16 DSA wyniosła 14,8 godziny w przypadku serwisu Facebook i 48,18 godziny w przypadku Instagrama. W przypadku zgłoszeń od zaufanych podmiotów sygnalizujących, przesłanych za pośrednictwem tego samego kanału, mediana czasu reakcji wyniosła 11,7 godziny w przypadku serwisu Facebook i 10,71 godziny w przypadku Instagrama (dla wszystkich kategorii treści). Mediana czasu potrzebnego na wykonanie nakazów organów państw członkowskich dotyczących podjęcia działań przeciwko treściom uznanym za niezgodne z prawem wyniosła 33,6 godziny w przypadku serwisu Facebook i 58,5 godziny w przypadku serwisu Instagram. Mediany czasu reakcji przedstawione w naszych raportach dotyczących przejrzystości nie są podzielone według tego, czy dana treść została uznana za oszukańczą reklamę.

5. Liczba zgłoszeń dotyczących oszukańczych reklam otrzymanych od użytkowników w Polsce w ostatnich 12 miesiącach oraz liczba reklam usuniętych w następstwie tych zgłoszeń.

Nie jesteśmy w stanie przedstawić tego wskaźnika w odniesieniu do konkretnego kraju, niemniej podkreślamy – jak wskazaliśmy już wyżej –, że na poziomie UE, w okresie od lipca do grudnia 2025 r., otrzymaliśmy 788 745 zgłoszeń na podstawie art. 16 DSA na Facebooku i 353 997 na Instagramie. Spośród zgłoszeń od zaufanych podmiotów sygnalizujących, dotyczących konkretnie oszustw i wyłudzeń, otrzymaliśmy 1 732 na Facebooku i 95 na Instagramie, a mediana czasu rozpatrzenia wyniosła odpowiednio 9,8 godziny i 18,32 godziny.

6. Sposób realizacji przez Meta obowiązków wynikających z art. 34 i 35 DSA w zakresie identyfikowanie i ograniczania ryzyka systemowego związanego z oszustwami reklamowymi.

Przekazujemy nasz roczny raport z oceny ryzyka systemowego Komisji Europejskiej oraz Coimisiún na Meán i publikujemy go w naszym Centrum Transparentności. Raporty te są przygotowywane na podstawie ocen ryzyka, które przeprowadzamy zgodnie z art. 34 DSA, w ramach których starannie identyfikujemy, analizujemy i oceniamy ryzyko systemowe wynikające z konstrukcji i funkcjonowania Facebooka i Instagrama, w tym

rozpowszechnianie treści niezgodnych z prawem (co w tym kontekście obejmuje również oszukańcze reklamy) oraz negatywne skutki w zakresie ochrony konsumentów. Zgodnie z art. 35 DSA wprowadzamy następnie rozsądne, proporcjonalne i skuteczne środki ograniczające ryzyko, dostosowane do tych zagrożeń, jak szczegółowo opisano powyżej. Środki te działają na pięciu poziomach i obejmują sprawdzanie każdej reklamy przez sztuczną inteligencję przed publikacją, technologię rozpoznawania twarzy, która odrzuca oszukańcze reklamy nadużywające wizerunku osób publicznych (o ile wyraziły na to zgodę), stale ulepszane klasyfikatory przeciwdziałające ewoluującym taktikom omijania zabezpieczeń po publikacji, bezpośrednie kanały eskalacji do polskich organów (CBZC, UKNF, NASK, UOKiK) oraz plan weryfikacji reklamodawców, którego celem jest osiągnięcie do końca 2026 r. 90% globalnych przychodów z reklam pochodzących od zweryfikowanych reklamodawców. Nasze środki ograniczające ryzyko podlegają niezależnemu audytowi zgodnie z art. 37 DSA.

7. Informacje, czy w ramach oceny ryzyka systemowego Meta uwzględnia specyfikę polskiego rynku, języka polskiego oraz charakterystyczne dla Polski schematy oszustw.

Zgodnie z wymogami DSA nasze oceny ryzyka uwzględniają różnorodność językową i rynkową w całej UE, a środki ograniczające ryzyko opierają się – między innymi – na informacjach otrzymanych od organów państw członkowskich. W praktyce obejmuje to zasoby służące do moderacji treści w języku polskim, systemy wykrywania oparte na schematach oszustw charakterystycznych dla polskiego rynku oraz opisane powyżej dedykowane kanały eskalacji spraw do polskich organów.

8. Działania podjęte przez Meta w następstwie spraw dotyczących oszukańczych reklam wykorzystujących wizerunek Rafała Brzosińskiego oraz innych polskich przedsiębiorców i osób publicznych.

Nie możemy ujawniać informacji dotyczących indywidualnych kont użytkowników ani spraw związanych z toczącymi się postępowaniami sądowymi. Jednak w ujęciu ogólnym, w przypadku podejrzenia, że reklama wykorzystuje w sposób nieuprawniony wizerunek osoby publicznej, stosujemy technologię rozpoznawania twarzy w celu porównania reklamy ze zdjęciami profilowymi tej osoby – o ile ta osoba wyraziła na to zgodę. Jeśli potwierdzimy zgodność i stwierdzimy, że reklama ma charakter oszukańczy, jest ona odrzucana przed jej wyświetleniem (lub usuwana, jeśli zostanie wykryta później). Udostępniamy również nasze narzędzie do ochrony praw do marki poszkodowanym właścicielom praw, w tym polskim bankom i dużym przedsiębiorstwom, umożliwiając im identyfikację i zgłaszanie treści, które nadużywają ich znaków towarowych lub innych oznaczeń marki.

9. Informacje na temat współpracy Meta z polskimi organami państwa, organami ścigania oraz właściwymi instytucjami odpowiedzialnymi za cyberbezpieczeństwo w zakresie przeciwdziałania tego typu oszustwom.

Od 2022 r. współpracujemy z UKNF w celu zwalczania oszukańczych i wprowadzających w błąd praktyk; na dzień dzisiejszy UKNF, NASK oraz UOKiK oraz mają dostęp do naszych dedykowanych kanałów eskalacji i korzystają z nich. Bardzo cenimy sobie naszą stałą współpracę z tymi organami i chętnie przyjmujemy ich opinie i spostrzeżenia, które mogą pomóc w usprawnieniu naszych działań.

Meta współpracuje również z organami ścigania w ramach prowadzonych przez nie dochodzeń i ujawnia dane w odpowiedzi na ważne wnioski prawne zgodnie z obowiązującym prawem, naszym regulaminem oraz politykami. Organizujemy specjalistyczne szkolenia dla organów ścigania dotyczące naszych procesów i kanałów składania wniosków, aby umożliwić im żądanie od Meta danych, które mogą pomóc w prowadzonych przez nie dochodzeniach.