



stowarzyszenie
żydowskie
czulent

Badanie usuwania treści nielegalnych w internecie 2023/2024

/Pierwsza tura ewaluacji



Publikacja powstała w ramach projektu:

„Zabezpieczenie naszej społeczności, ochrona naszej demokracji: zwalczanie antysemityzmu poprzez zintegrowane podejście do rzecznictwa i bezpieczeństwa” (projekt PROTEUS), współfinansowanego przez Unię Europejską.



Co-funded by
the European Union

G | M | F
Transatlantic
Foundation

Wyrażane poglądy i opinie są wyłącznie poglądami autora(-ów) i niekoniecznie odzwierciedlają stanowisko Unii Europejskiej lub Europejskiej Agencji Wykonawczej ds. Edukacji i Kultury. Ani Unia Europejska, ani organ przyznający grant nie ponoszą za nie odpowiedzialności.

Projekt współfinansowany był także ze środków Fundacji Pamięć, Odpowiedzialność i Przyszłość (EVZ).



Niniejsza publikacja nie prezentuje stanowiska i opinii Fundacji Pamięć, Odpowiedzialność i Przyszłość (EVZ).

ISBN 978-83-972977-1-5



stowarzyszenie
żydowskie
czulent

Żydowskie Stowarzyszenie Czulent
ul. św. Sebastiana 36/1, 31-051 Kraków
www.czulent.pl

Kraków 2024

Wesprzyj Żydowskie Stowarzyszenie Czulent
Nr konta bankowego:
42 1600 1462 1880 9351 6000 0001

Wspierając nasze działania,
pomagasz chronić społeczność
żydowską i przeciwdziałać
antysemityzmowi.

Spis Treści

- 4 Wstęp
- 5 Słowniczek
- 7 Ramy prawne
- 9 Metodologia
- 10 Etapy MRE
- 11 Kluczowe dane
- 13 Wskaźniki usuwalności zgłoszeń
- 17 Wnioski i rekomendacje
- 19 Spis rycin
- 19 Bibliografia

Wstęp

W ostatnich latach obserwujemy silny wzrost treści antysemitycznych w internecie. Antysemityzm stanowi narzędzie dezinformacji i polaryzacji społeczeństwa, będące zagrożeniem dla demokracji i bezpieczeństwa publicznego. Jest to globalny problem, co potwierdzają analizy i raporty opracowane przez międzynarodowe organizacje, takie jak Institute for Strategic Dialogue (ISD)¹² oraz Instytut Kościuszki, pokazujące także skalę tego zjawiska w Polsce.³

Monitoring antysemitycznych incydentów prowadzony przez Żydowskie Stowarzyszenie Czulent dodatkowo wskazuje, że antysemityczne przestępstwa są w przeważającej mierze popełniane w internecie oraz wpływają negatywnie na poczucie bezpieczeństwa członkiń i członków społeczności żydowskiej oraz żydowskich organizacji.⁴

W ramach półrocznego badania podjęliśmy działania mające na celu sprawdzenie, czy i kiedy nienawistne treści o charakterze antysemitycznym są usuwane przez międzynarodowe i polskie serwisy IT po otrzymaniu zgłoszenia od użytkowniczek i użytkowników o konieczności ich usunięcia. Chcieliśmy także sprawdzić, czy istnieje różnica w usuwaniu nienawistnych treści zgłaszanych przez zwykłych użytkowników a tzw. zaufane podmioty sygnalizujące. W tym celu przeprowadziliśmy badanie usuwania treści nielegalnych w internecie (zwane także MRE – *Monitoring and Reporting Exercise*), testując międzynarodowe platformy internetowe: Facebook, Instagram, YouTube, TikTok, platforma X (dawniej Twitter) oraz dostawców polskich usług pośrednich: Agora, wp.pl, onet.pl, natemat.pl, dorzeczy.pl i wykop.pl pod kątem stosowania krajowych i unijnych przepisów nakazujących usunięcie lub uniemożliwienie dostępu do treści nielegalnych, w tym mowy nienawiści. Badanie zostało przeprowadzone w momencie wchodzenia w życie nowej unijnej regulacji dotyczącej poprawy bezpieczeństwa w przestrzeni cyfrowej, znanej jako Rozporządzenie 2022/2065 lub Akt o usługach cyfrowych.

1 Institute for Strategic Dialogue, B'nai B'rith International, [Online Antisemitism: A Toolkit for Civil Society](#), 2022 (dostęp: 19.08.2024).

2 Institute for Strategic Dialogue, [Narratives of Hate: Post-7 October Antisemitism and Anti-Muslim Hate on Social Media](#), 2024 (dostęp: 19.08.2024).

3 Civic Resilience Initiative, Detector Media i Instytut Kościuszki, [Odporność na dezinformacje w krajach trójkąta lubelskiego](#), 2022 (dostęp: 19.08.2024).

4 Żydowskie Stowarzyszenie Czulent, [Antysemityczne incydenty w 2023 roku. Badanie doświadczenia społeczności żydowskiej](#), 2024 (dostęp: 19.08.2024).

Słowniczek

Badanie przygotowane i zrealizowane przez Żydowskie Stowarzyszenie Czulent używa pojęć stosowanych w unijnym Akcie o usługach cyfrowych oraz innych dokumentach w odniesieniu do przeciwdziałania rozpowszechnianiu nielegalnych treści w internecie.

antysemityzm – “to określone postrzeganie Żydów, które może wyrażać się jako nienawiść wobec Żydów. Retoryczne i fizyczne przejawy antysemityzmu są skierowane wobec żydowskich lub nieżydowskich osób i/lub ich własności, wobec instytucji społeczności żydowskiej oraz obiektów religijnych”.⁵ Żydowskie Stowarzyszenie Czulent stosuje operacyjną definicję antysemityzmu IHRA. Polski rząd poparł definicję w 2016 roku, a jej przyjęcie zostało potwierdzone przez Ministerstwo Kultury i Dziedzictwa Narodowego w 2021 roku;⁶

DSA lub **Rozporządzenie 2022/2065** – Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/2065 z dnia 19 października 2022 r. w sprawie jednolitego rynku usług cyfrowych oraz zmiany dyrektywy 2000/31/WE (Akt o usługach cyfrowych);⁷

eskalacja – powiadomienia o nielegalnych treściach przesłanych przez podmioty posiadające status zaufanego podmiotu sygnalizującego lub zaufanego partnera sygnalizującego przesłane do bardzo dużych platform internetowych przy użyciu konkretnych narzędzi komunikacyjnych;

IHRA – Międzynarodowy Sojusz na rzecz Pamięci o Holokauście (International Holocaust Remembrance Alliance);⁸

KK – ustawa z dnia 6 czerwca 1997 r. – Kodeks Karny;⁹

moderowanie treści – „działania, bez względu na to czy są one zautomatyzowane, podejmowane przez dostawców usług pośrednich, których celem jest, w szczególności, wykrywanie, identyfikowanie i zwalczanie nielegalnych treści lub informacji niezgodnych z warunkami korzystania z ich usług, przekazywanych przez odbiorców usługi, w tym wdrażane środki, które wpływają na dostępność, widoczność i osiągalność takich nielegalnych treści lub informacji, takie jak depozycjonowanie takich treści lub informacji, demonetyzacja, uniemożliwienie dostępu do nich lub ich usunięcie, lub które wpływają na możliwość przekazywania takich informacji przez odbiorców usługi, takie jak zamknięcie lub zawieszenie konta odbiorcy”;¹⁰

5 Międzynarodowy Sojusz na rzecz Pamięci o Holokauście, *Robocza definicja antysemityzmu* (dostęp: 19.08.2024).

6 Ministerstwo Kultury i Dziedzictwa Narodowego, *Oświadczenie w sprawie definicji IHRA* (dostęp: 19.08.2024).

7 Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/2065 z dnia 19 października 2022 r. w sprawie jednolitego rynku usług cyfrowych oraz zmiany dyrektywy 2000/31/WE – akt o usługach cyfrowych (zwane dalej *Rozporządzeniem 2022/2065* lub *DSA*), Dz. Urz. UE L 277/1 z 27.10.2022 (dostęp: 19.08.2024).

8 Międzynarodowy Sojusz na rzecz Pamięci o Holokauście (dostęp: 19.08.2024).

9 Ustawa z dnia 6 czerwca 1997 r. - Kodeks Karny (zwana *Kodeksem karnym*), Dz.U. 1997 nr 88 poz. 553, (dostęp: 19.08.2024).

10 Rozporządzenie 2022/2065, art. 3 (t).

mowa nienawiści — nie istnieje prawnie wiążąca definicja mowy nienawiści w Polsce. W raporcie używamy definicji przyjętej przez Radę Europy w 1997 r., według której mowę nienawiści należy rozumieć jako „wszelkie formy wypowiedzi, które szerzą, podżegają, promują lub usprawiedliwiają nienawiść rasową, ksenofobię, antysemityzm lub inne formy nienawiści opartej na nietolerancji, w tym nietolerancji wyrażonej przez agresywny nacjonalizm i etnocentryzm, dyskryminację oraz wrogość wobec mniejszości, migrantów i osób pochodzenia imigranckiego”;¹¹

MRE — Monitoring and Reporting Exercise;

nielegalne treści — „informacje, które same w sobie lub przez odniesienie do działania, w tym sprzedaży produktów lub świadczenia usług, nie są zgodne z prawem Unii lub z prawem jakiegokolwiek państwa członkowskiego, które jest zgodne z prawem Unii, niezależnie od konkretnego przedmiotu lub charakteru tego prawa”;¹²

platforma internetowa — „usługa hostingu, która na żądanie odbiorcy usługi przechowuje i rozpowszechnia publicznie informacje, chyba że takie działanie jest nieznaczną lub wyłącznie poboczną cechą innej usługi lub nieznaczną funkcją głównej usługi, i ze względów obiektywnych i technicznych nie można z niej skorzystać bez takiej innej usługi, a włączenie takiej cechy lub funkcji w taką inną usługę nie jest sposobem na obejście stosowania niniejszego rozporządzenia”;¹³

usługa pośrednia — „usługa „hostingu” polegająca na przechowywaniu informacji przekazanych przez odbiorcę usługi oraz na jego żądanie”;¹⁴

VLOP — bardzo duża platforma internetowa (very large online platform); termin występujący w unijnym Rozporządzeniu 2022/2065;¹⁵

VLOSE — bardzo duża wyszukiwarka internetowa (very large online search engine); termin występujący w unijnym Rozporządzeniu 2022/2065;¹⁶

zaufany partner sygnalizujący — podmioty, którym bardzo duże platformy internetowe przyznały specjalny status umożliwiający zgłaszanie nielegalnych treści lub materiałów naruszających zasady platform internetowych. Status ten jest przyznawany na zasadach dobrowolnych i zgodnie z warunkami platformy;

zaufany podmiot sygnalizujący — status nadawany przez krajowych koordynatorów ds. usług cyfrowych podmiotom posiadającym szczególną wiedzę ekspercką oraz kompetencje w identyfikowaniu i zgłaszaniu nielegalnych treści;¹⁷

zgłoszenie zautomatyzowane — zgłoszenie dokonywane poprzez wciśnięcie ikony „zgłoś” lub „kosz”.

11 CoE, Recommendation No. R (97) 20 of the Committee of Ministers to member states on “hate speech” (dostęp: 19.08.2024) Nowa definicja mowy nienawiści została przyjęta przez Radę Europy w 2022 r.: CoE, Recommendation CM/Rec (2022) 16 of the Committee of Ministers to Member States on combating hate speech, (dostęp:19.08.2024).

12 Rozporządzenie 2022/2065, art. 3 (h).

13 Ibidem, art. 3 (i).

14 Ibidem, art. 3 (g) (iii).

15 Komisja Europejska, Nadzór nad wyznaczonymi bardzo dużymi platformami internetowymi i wyszukiwarkami w ramach aktu o usługach cyfrowych, (dostęp: 19.08.2024)

16 Ibidem.

17 Rozporządzenie 2022/2065, art. 22.

Ramy prawne

DSA ma na celu harmonizację zasad dotyczących dostawców usług pośrednich. Reguluje ich odpowiedzialność w UE z zamiarem „zapewnienia bezpiecznego, przewidywalnego i budzącego zaufanie środowiska internetowego, przeciwdziałania rozpowszechnianiu nielegalnych treści w internecie oraz zagrożeniom społecznym, jakie może stwarzać rozpowszechnianie dezinformacji lub innych treści”.¹⁸ DSA nie dostarcza jednoznacznej definicji, co stanowi „nielegalną działalność” lub „nielegalne treści”. Podkreśla, że koncepcja nielegalności powinna być oceniana zgodnie z prawem unijnym i krajowym.¹⁹ Pojęcie „nielegalnych treści” powinno, według Rozporządzenia 2022/2065, odzwierciedlać środowisko pozainternetowe. Innymi słowy, należy je rozumieć szeroko i niezależnie od formy, czy są to treści, produkty, usługi czy działania.²⁰

Zgodnie z art. 1(1)(a) Decyzji ramowej Rady UE 2008/913/WSiSW, państwa członkowskie UE są zobowiązane zapewnić karalność „publicznego nawoływania do przemocy lub nienawiści skierowanej przeciwko grupie osób, którą definiuje się według rasy, koloru skóry, wyznawanej religii, pochodzenia albo przynależności narodowej lub etnicznej, lub przeciwko członkowi takiej grupy”.²¹

Polski Kodeks karny nie zawiera definicji prawnej mowy nienawiści. Jednak posiada zbiór przepisów prawnych mających na celu zwalczanie zjawiska mowy nienawiści, które są również stosowane w środowisku internetowym. Przepisy te zabraniają m.in.: stosowania bezprawnych gróźb wobec osoby lub grupy osób (art. 119 KK), propagowania totalitarnego ustroju państwa lub nawoływania do nienawiści (art. 256 KK) oraz znieważania określonej grupy ludności lub jednostki (art. 257 KK). Niemniej przepisy te można stosować tylko wtedy, gdy wymienione przestępstwa są motywowane narodowością, pochodzeniem etnicznym, rasą, przynależnością polityczną lub religijną atakowanej osoby lub grupy. Oznacza to, że tylko niektóre kategorie mowy nienawiści są chronione (wśród nich antysemicka mowa nienawiści). Inne (jak homofobiczna oraz trans-fobiczna mowa nienawiści) nie są chronione w ramach tego specjalnego systemu i wymagają stosowania ogólnych przepisów (na przykład zakazujących znieważania innych osób), które jednak przewidują dużo niższe sankcje, a często przerzucają także obowiązek ścigania sprawcy na pokrzywdzonego. Oddzielnie, ustawa o Instytucie Pamięci Narodowej (w art. 55) zakazuje publicznego zaprzeczania zbrodniom nazistowskim oraz innym zbrodniom popełnionym w latach 1917-1990 penalizując m.in. tzw. kłamstwo oświęcimskie.²²

¹⁸ Ibidem, pkt 9.

¹⁹ Ibidem, art. 3 (h).

²⁰ Ibidem, pkt 12.

²¹ Decyzja ramowa Rady 2008/913/WSiSW z dnia 28 listopada 2008 r. w sprawie zwalczania pewnych form i przejawów rasizmu i ksenofobii za pomocą środków prawno-karnych (zwana [Decyzją ramową Rady 2008/913/WSiSW](#)), Dz. Urz. UE L 328/55 z 6.12.2008 (dostęp: 19.08.2024).

²² Ustawa z dnia 18 grudnia 1998 r. o Instytucie Pamięci Narodowej – Komisji Ścigania Zbrodni przeciwko Narodowi Polskiemu (zwana [ustawą o IPN](#)), Dz. U. z 2023 r. poz. 102 (dostęp: 19.08.2024).

Wśród polskich przepisów krajowych ważne jest również wskazanie na art. 14 ustawy o świadczeniu usług drogą elektroniczną, który stanowi, że dostawcy usług pośrednich nie ponoszą odpowiedzialności za przechowywanie nielegalnych treści, jeśli usuną je po otrzymaniu wiarygodnej informacji na ten temat.²³ Przepis ten powinien mieć zastosowanie do moderacji treści w Polsce, szczególnie gdy przepisy DSA nie są w pełni stosowane w kraju.

Od końca sierpnia 2023 r. DSA ma zastosowanie do bardzo dużych platform internetowych i bardzo dużych wyszukiwarek internetowych, czyli platform z ponad 45 milionami użytkowników w UE. Przepisy Rozporządzenia 2022/2065 znajdują pełne zastosowanie od 17 lutego 2024 r. Oznacza to, że dostawcy usług pośrednich, oferujący swoje usługi na jednolitym rynku, niezależnie od miejsca ich ustanowienia, muszą przestrzegać nowych zasad.²⁴

Dostawcy usług cyfrowych, nie są zobowiązani do moderowania treści, które są przez nich przekazywane lub przechowywane, pod kątem nielegalnej działalności. DSA wyraźnie stwierdza w art. 6 (1), że dostawcy usług pośrednich nie będą odpowiedzialni za przechowywanie informacji, jeśli nie są świadomi nielegalnej działalności lub nielegalnych treści, lub po uzyskaniu takiej wiedzy, działają bezzwłocznie w celu usunięcia lub zablokowania do nich dostępu.

Każda użytkowniczka oraz użytkownik ma prawo i możliwość zgłosić serwisom IT wykorzystanie ich usług do szerzenia treści nielegalnych. Na mocy art. 22 DSA, szczególna rola w tym zakresie przypada zaufanym podmiotom sygnalizującym. Status zaufanego podmiotu sygnalizującego jest przyznawany przez Koordynatora ds. usług cyfrowych państwa członkowskiego podmiotom posiadającym wiedzę fachową w określonej dziedzinie. Jak stwierdzono w art. 22 (2) (a) Rozporządzenia 2022/2065, podmiot ubiegający się o taki status powinien posiadać „szczególną wiedzę i kompetencje w zakresie wykrywania, identyfikacji i zgłaszania nielegalnych treści”.

Polska nie implementowała przepisów DSA do swojego porządku prawnego do dnia daty publikacji niniejszego raportu. Tym samym, nie została jeszcze wyznaczona instytucja, która pełniłaby rolę Koordynatora ds. usług cyfrowych odpowiedzialnego za nadzór, egzekwowanie i monitorowanie wdrażania DSA, w tym przyznawanie statusu zaufanego podmiotu sygnalizującego. Pomimo braku implementacji Aktu o usługach cyfrowych do polskiego porządku prawnego, Rozporządzenie 2022/2065 jest wiążące na terenie Unii Europejskiej.

²³ Ustawa z dnia 18 lipca 2002 r. o świadczeniu usług drogą elektroniczną, Dz.U. 2002 nr 144 poz. 1204 (dostęp: 19.08.2024).

²⁴ Komisja Europejska, Unijny akt o usługach cyfrowych. Bezpieczeństwo i odpowiedzialność uczestników interakcji w internecie (dostęp: 19.08.2024).

MRE przedstawia wyniki monitorowania i zgłaszania nielegalnych treści w internecie zgodnie z założeniami DSA. **Wszystkie zgłoszone antysemickie treści naruszały przepisy polskiego kodeksu karnego i spełniały przesłanki operacyjnej definicji antysemityzmu IHRA.** Raport przedstawia najważniejsze fakty dotyczące usuwania treści przez zagraniczne i polskie firmy IT w ramach DSA, w tym liczby i proporcje dokonanych zgłoszeń, czas ich oceny, wskaźniki usunięcia oraz prawną kwalifikację nielegalnych treści. MRE było inspirowane ewaluacją Unijnego kodeksu postępowania w sprawie zwalczania nielegalnej mowy nienawiści w internecie przeprowadzanym przez Komisję Europejską w latach 2016-2022.²⁵

Metodologia

- Wyniki MRE są pochodną zagregowanych i wysłanych zgłoszeń oraz metodologii badania. Przedstawione w badaniu dane nie odzwierciedlają rzeczywistej ilości oraz rodzajów nielegalnej mowy nienawiści w przestrzeni cyfrowej.
- Dla zachowania rzetelności i wiarygodności badania firmy IT nie zostały poinformowane o trwającym badaniu. Analizowane były treści wyłącznie w języku polskim.
- MRE było przeprowadzane od 1 listopada 2023 r. do 7 kwietnia 2024 roku (6 miesięcy) z zastosowaniem mechanizmów określonych w unijnym akcie o usługach cyfrowych oraz części metodologii przyjętej w ewaluacji Unijnego kodeksu postępowania w zakresie zwalczania nielegalnego nawoływania do nienawiści w internecie. Badanie zostało poszerzone o analizę serwisów polskojęzycznych działających na polskim rynku. MRE było realizowane zgodnie z wewnętrznymi procedurami organizacji niezbędnymi do wykrywania, identyfikacji i zgłaszania nielegalnych treści do 5 bardzo dużych platform internetowych (**Facebook, Instagram, YouTube, TikTok i X**) oraz 6 polskich dostawców usług pośrednich (**Agora, wp.pl, onet.pl, natemat.pl, dorzeczy.pl, wykop.pl**).
- Wszystkie zgłoszone treści wypełniały kryteria roboczej definicji antysemityzmu IHRA oraz definicji mowy nienawiści Rady Europy.
- Wszystkie zgłoszone treści zostały poddane ocenie prawnej i uznane za nielegalną mowę nienawiści na podstawie przepisów polskiego Kodeksu karnego oraz Decyzji ramowej Rady UE 2008/913/JHA
- Wszystkie zgłoszone treści naruszały także regulaminy ustalone przez serwisy IT.

²⁵ Komisja Europejska, [Unijny kodeks postępowania w zakresie zwalczania nielegalnego nawoływania do nienawiści w internecie](#) (dostęp: 19.08.2024).

- Zgłoszenia były przekazywane za pośrednictwem kanałów raportowania dostępnych dla wszystkich użytkowników, a w niektórych przypadkach również przez kanały dedykowane wyłącznie zaufanym podmiotom sygnalizującym i zaufanym partnerom, co pozwoliło nam porównać reakcje firm IT. Mechanizm raportowania dostępny dla zaufanych podmiotów sygnalizujących lub zaufanych partnerów działa obecnie tylko w odniesieniu do bardzo dużych platform internetowych. Polscy dostawcy usług pośrednich nie posiadają odrębnych mechanizmów raportowania dla zaufanych podmiotów sygnalizujących lub zaufanych partnerów. Korzystają jedynie ze zautomatyzowanych narzędzi raportowania dostępnych dla ogólnych użytkowników, a te same zgłoszenia przesyłane im drogą elektroniczną nie były klasyfikowane jako eskalacja.

Stowarzyszenie Żydowskie Czulent wysłało trzy wnioski do pięciu bardzo dużych platform internetowych z prośbą o dołączenie do programu dedykowanego zaufanym partnerom. Tylko jeden VLOP udzielił odpowiedzi organizacji – odmownej, kierując Stowarzyszenie do przepisów DSA. Eskalacje były więc przekazywane VLOP-om za pośrednictwem podmiotów trzecich.

Etapy MRE

- a. wykrywanie, identyfikacja i agregacja nielegalnych treści w internecie,
- b. zgłaszanie nielegalnych treści bardzo dużym platformom internetowym oraz polskim dostawcom usług pośrednich za pomocą zautomatyzowanego narzędzia raportowania przez ogólnych użytkowników lub przy użyciu procedury „zgłoś i usuń”,
- c. weryfikacja czasu oceny zgłoszeń przesłanych przez ogólnych użytkowników oraz ich dostępności,
- d. eskalacja — zgłaszanie nielegalnych treści przez zaufane podmioty sygnalizujące lub zaufanych partnerów,
- e. weryfikacja czasu oceny zgłoszeń przesłanych VLOP-om przez zaufane podmioty sygnalizujące lub zaufanych partnerów oraz ich dostępności,
- f. weryfikacja dostępności zgłoszonych nielegalnych treści w dniu zakończenia badania.

Kluczowe dane

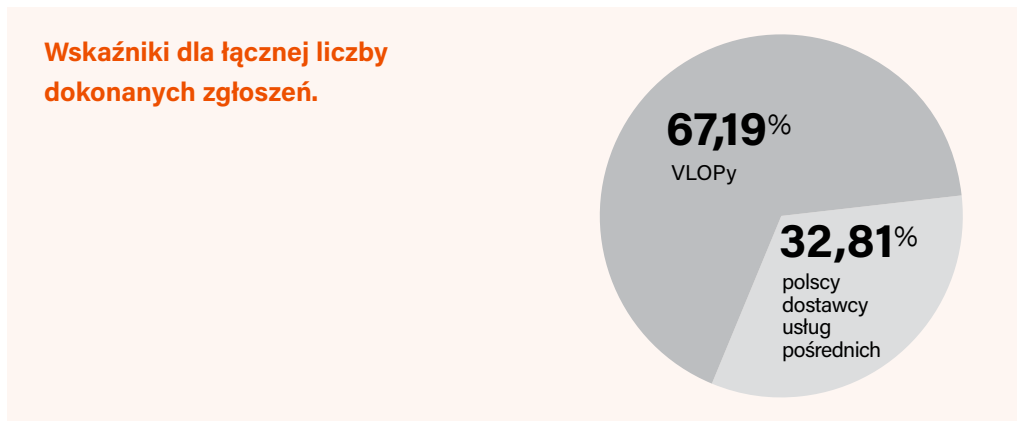
MRE koncentruje się na liczbie powiadomień (interwencji), rozumianych jako zgłoszenia wysyłane do serwisów IT. Takie podejście wynika z potrzeby oceny postępowania firm informatycznych w odpowiedzi na zgłoszenia przekazywane przez indywidualne użytkowniczki i użytkowników lub zaufane podmioty sygnalizujące i zaufanych partnerów. Liczba zgłoszeń nie jest tożsama z liczbą zgłoszonych incydentów. Dlatego całkowita liczba powiadomień (192) jest wyższa niż liczba zidentyfikowanych nielegalnych treści online (174). Ta rozbieżność wynika z faktu, że w niektórych przypadkach dotyczących polskich dostawców usług pośrednich, ta sama treść była zgłaszana dwukrotnie przez różne użytkowniczki i użytkowników z wykorzystaniem innych form raportowania treści, czyli zgłoszenia zautomatyzowanego oraz wiadomości e-mail.

- Dokonano łącznie **192** zgłoszeń, w tym: 129 zgłoszeń do VLOP-ów oraz **63** zgłoszeń do polskich dostawców usług pośrednich.
- **129** zgłoszeń do VLOP-ów zostało złożonych za pomocą zautomatyzowanych środków dostępnych użytkowniczkom i użytkownikom. Dodatkowo **60** z **129** interwencji zostało eskalowanych, co oznacza, że zostały one dodatkowo złożone przez zaufane podmioty sygnalizujące lub zaufanych partnerów.
- **63** zgłoszenia zostały złożone polskim dostawcom usług pośrednich. **45** z **63** interwencji zostało dokonanych w formie zgłoszenia zautomatyzowanego, a **18** zostało złożonych drogą e-mailową z wykorzystaniem procedury "zgłoś i usuń".
- DSA stanowi, że zgłoszenia dotyczące nielegalnych treści, przesyłane przez osoby fizyczne lub podmioty, w tym za pomocą zautomatyzowanych środków elektronicznych, powinny być traktowane w sposób „terminowy, niearbitralny i obiektywny oraz z zachowaniem należytej staranności”²⁶ Zgłoszenia przesyłane bardzo dużym platformom internetowym przez zaufane podmioty sygnalizujące powinny być traktowane priorytetowo i bez „nieuzasadnionej zwłoki”²⁷ Pojęcia „nieuzasadniona zwłoka” czy „terminowy sposób” nie określają jednak ścisłych limitów czasowych na ocenę zgłoszeń, dlatego do celów MRE zastosowaliśmy podobną metodologię do tej używanej w Unijnym kodeksie postępowania w zakresie zwalczania nielegalnego nawiązywania do nienawiści w internecie. Limity czasowe dla odpowiedzi na zgłoszenia zostały ustalone na 24 godziny, 48 godzin, mniej niż tydzień i dłużej niż tydzień.

Instagram potrzebował **28,5** dni, YouTube **28,5** dni, Facebook **8,26** dni, a TikTok **7** dni na podjęcie decyzji w sprawie zgłoszenia dokonanego w formie eskalacji.

26 Rozporządzenie 2022/2065, art. 16 (6).

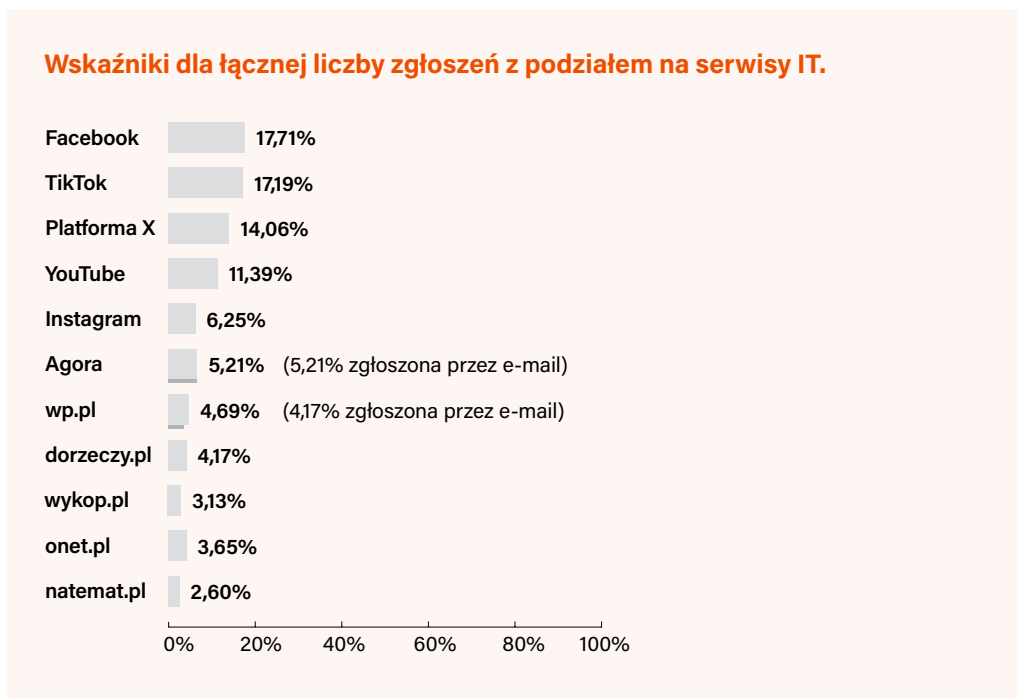
27 Ibidem, art. 22 (1).



Wykres 1.

Wśród VLOP-ów, Facebook otrzymał największą liczbę zgłoszeń (**34**), następnie **TikTok (33)**, **platforma X (27)**, **YouTube (23)** i **Instagram (12)**.

Wśród dostawców polskich usług internetowych, Agora (wyborcza.pl i gazeta.pl) otrzymała największą liczbę interwencji (**20**), następnie **wp.pl (17)**, **dorzeczy.pl (8)**, **onet.pl (7)**, **wykop.pl (6)** i **natemat.pl (5)**.



Wykres 2.

Wskaźniki usuwalności zgłoszeń

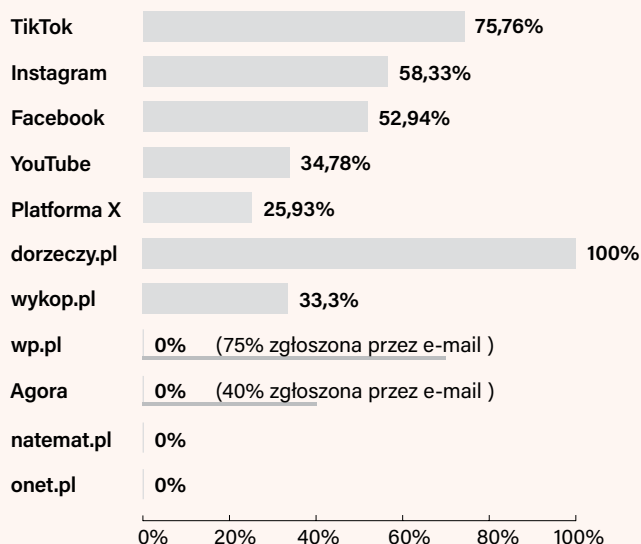
Wskaźniki łącznie usuniętych zgłoszeń.



Wykres 3.

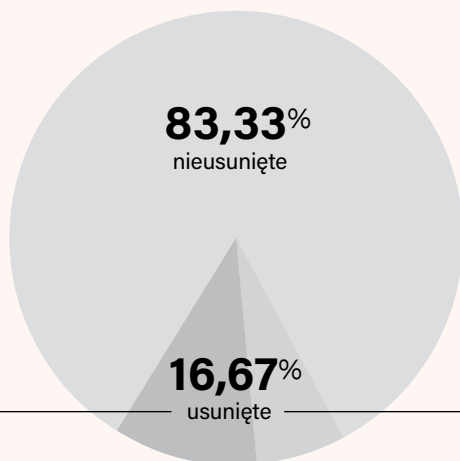
Dostępność zgłoszonych treści na dzień zamknięcia MRE.

Wskaźniki usuniętych zgłoszeń w stosunku do liczby zgłoszeń dokonanych w ramach każdego serwisu IT.



Wykres 4.

Wskaźniki usuwalności dla łącznej liczby zgłoszeń dokonanych przez indywidualnych użytkowników.



10,42%
dostawcy polskich usług pośrednich

| | |
|----------------------------|-------|
| dorzeczy.pl | 4,17% |
| wp.pl (zgłoszone e-mailem) | 3,13% |
| Agora (zgłoszone e-mailem) | 2,08% |
| wykop.pl | 1,04% |
| Agora | 0% |
| Onet.pl | 0% |
| natemat.pl | 0% |

6,25%
usunięte przez VLOPy

| | |
|-------------|-------|
| TikTok | 5,73% |
| Instagram | 0,52% |
| Facebook | 0% |
| YouTube | 0% |
| Platforma X | 0% |

Dla VLOP-ów, łączna liczba usuniętych zgłoszeń dokonanych przez indywidualnych za pomocą zgłoszenia zautomatyzowanego wyniosła 12 ze 129 interwencji, co stanowi 9,30%.

Dla dostawców polskich usług pośrednich, łączna liczba usuniętych zgłoszeń wyniosła 20 z 63 interwencji przesłanych przez zwykłych użytkowników za pomocą zgłoszenia zautomatyzowanego lub wiadomości elektronicznej, czyli 31,75%.

Wykres 5.

Wyniki dla zautomatyzowanych zgłoszeń lub wiadomości e-mail dokonanych przez indywidualnych użytkowników.

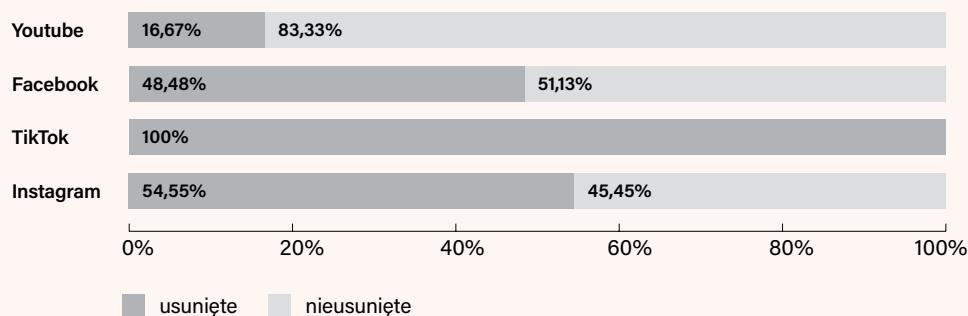
ESKALACJE

Treści nielegalne zostały zgłoszone przez zaufane podmioty sygnalizujące lub zaufanych partnerów czterem bardzo dużym platformom internetowym (Facebook, Instagram, YouTube i TikTok). Żadne zgłoszenie nie zostało przesłane do platformy X z powodu braku możliwości dokonania interwencji przez zaufanego partnera w trakcie przeprowadzania badania.

Całkowita **liczba usuniętych zgłoszeń i dokonanych** przez zaufane podmioty sygnalizujące lub zaufanych partnerów wyniosła **28 z 60** interwencji. Reasumując, VLOP-y usunęły **46,67%** zgłoszeń przesłanych przez zaufanych partnerów lub podmioty sygnalizujące.

Wyniki dla zgłoszeń dokonanych przez zaufane podmioty sygnalizujące lub zaufanych partnerów.

Wskaźniki usuniętych zgłoszeń w ramach eskalacji z podziałem na VLOP-y.

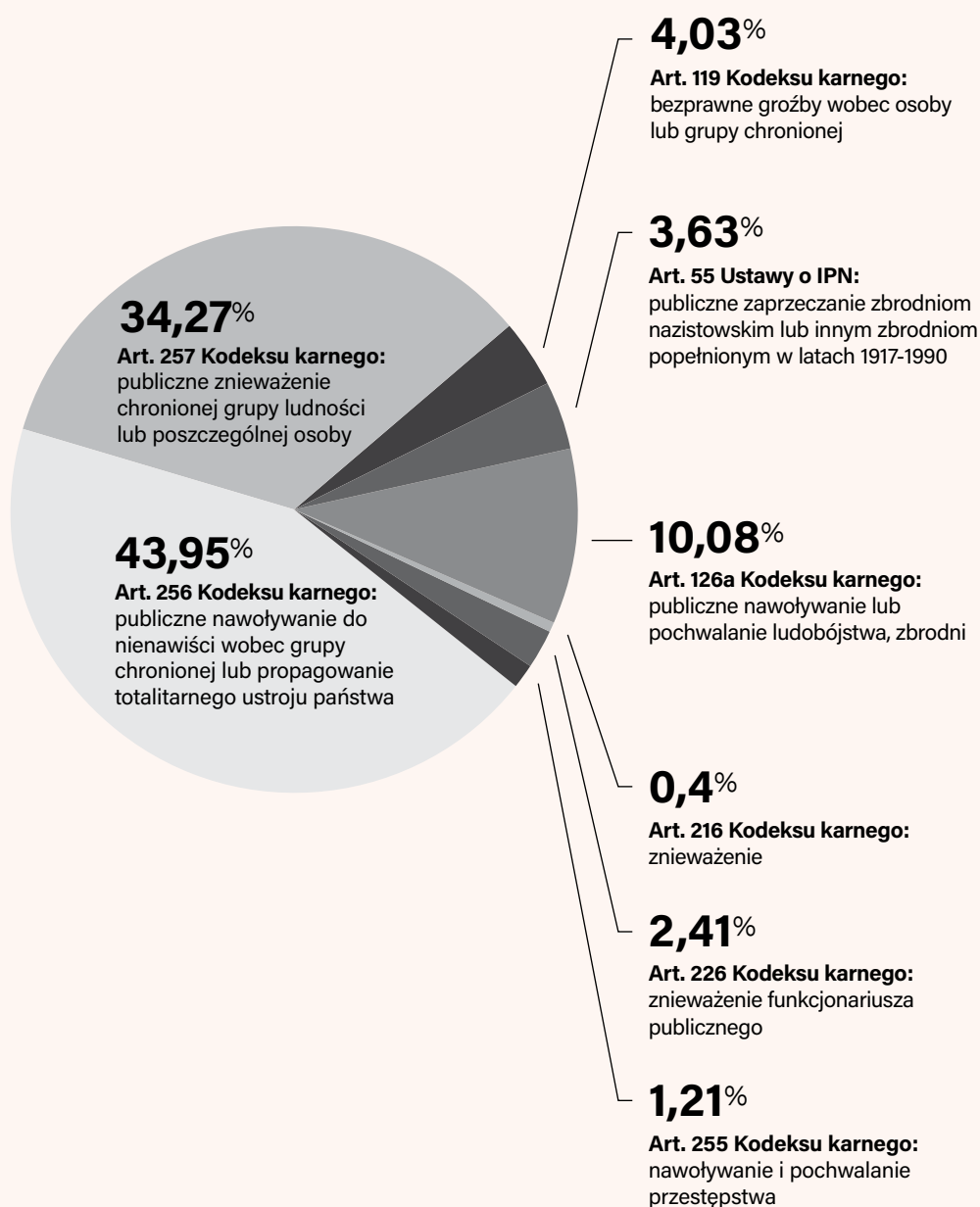


Wykres 6.

Podstawa prawna dokonanych zgłoszeń

Wskaźnik naruszeń Kodeksu karnego w stosunku do łącznej liczby zgłoszeń.

Wykres 7.



Wnioski i rekomendacje

Badanie wykazało, że serwisy IT usunęły jedynie **50,39%** zgłoszeń nielegalnych treści motywowanych antysemityzmem. Pomimo obowiązków nałożonych na firmy IT przez unijny Akt o usługach cyfrowych oraz przepisy krajowe i unijne, a także pomimo naruszenia regulaminów serwisów, nie usunięto **49,61%** zgłoszonych naruszeń prawa. Badanie ujawniło wyższy wskaźnik usuwania nielegalnych treści, gdy zgłoszenia były składane przez zaufane podmioty sygnalizujące i zaufanych partnerów, w porównaniu do zgłoszeń przesyłanych przez użytkowników indywidualnych. Podobne wyniki można przypisać zgłoszeniom przesłanym e-mailem do polskich dostawców usług pośrednich, w przeciwieństwie do interwencji składanych za pomocą zautomatyzowanych narzędzi raportowania.

Starania o uzyskanie statusu zaufanego partnera przy VLOP-ach wskazały brak responsywności ze strony bardzo dużych platform internetowych, co stawia pod znakiem zapytania wolę współpracy z organizacjami pozarządowymi.

Bezpieczeństwo publiczne i dyskurs w przestrzeni cyfrowej są jednym z podstawowych i kluczowych wyzwań naszych czasów. Serwisy IT powinny bardziej dbać o zapewnienie bezpiecznej przestrzeni cyfrowej użytkownikom i użytkownikom usług online oraz grupom szczególnie narażonym na przestępstwa z nienawiści. W analizowanych treściach zaobserwowano przykłady dezinformacji mające na celu polaryzowanie społeczeństwa oraz tworzenie lub wzmacnianie społecznych obaw i lęków poprzez szerzenie antysemickich mitów i teorii spiskowych. W trakcie realizacji badania zaobserwowano, że zgłaszane treści mają charakter dyskryminacji interseksyjnej, w tym zawierały treści antyukraińskie.

DSA tworzy nową podstawę prawną do zwalczania nielegalnych treści w internecie. Aby skutecznie działać, podmioty zgłaszające powinny posiadać wiedzę ekspercką w określonych dziedzinach oraz zdolność do monitorowania i raportowania tych treści.

Skuteczność DSA zależy jednak przede wszystkim od uznania i wdrożenia obowiązku moderacji treści i mechanizmów działania przez serwisy IT oraz zaangażowania wszystkich podmiotów biorących udział w tworzeniu bezpieczniejszego środowiska online, w tym sprawnego przyjęcia przepisów wdrażających unijny Akt cyfrowy do polskiego porządku prawnego.

REKOMENDACJE

Szczególłą potrzebę dostrzegamy w niezwłocznym usuwaniu przez **serwisy IT** treści nielegalnych naruszających obowiązujące przepisy krajowe oraz unijne. Niezbędne wydaje się także:

- doprecyzowanie przez firmy informatyczne algorytmów mających na celu usuwanie treści nielegalnych;
- regularne szkolenie polskojęzycznych administratorów serwisów IT w celu dekodowania antysemityzmu a także znajomości prawa oraz stale rozwijającego się antysemickiego kodu językowego;
- kontynuacja przyznawania przez bardzo duże platformy internetowe statusu zaufanego partnera organizacjom eksperckim, zwłaszcza w przypadku braku powołania krajowego Koordynatora ds. usług cyfrowych.

W naszej ocenie polski ustawodawca powinien w szczególności zadbać o:

- powołanie krajowego Koordynatora do spraw usług cyfrowych, który będzie miał do dyspozycji odpowiednie zasoby zarówno osobowe jak i finansowe;
- doprecyzowanie na gruncie polskiego porządku prawnego roli i wytycznych sprawozdawczych dla zaufanych podmiotów sygnalizujących oraz udzielenie im rzeczywistego wsparcia finansowego w realizacji podejmowanych działań;
- powołanie Rady konsultacyjnej przy Koordynatorze do spraw. usług cyfrowych celem współpracy przede wszystkim strony rządowej z organizacjami posiadającymi status zaufanego podmiotu sygnalizującego w kontekście realizacji założeń i przepisów DSA.

SPIS RYCIN

- Wykres 1. Wskaźniki dla łącznej liczby dokonanych zgłoszeń.
- Wykres 2. Wskaźniki dla łącznej liczby zgłoszeń z podziałem na serwisy IT.
- Wykres 3. Wskaźniki łącznie usuniętych zgłoszeń.
- Wykres 4. Wskaźniki usuniętych zgłoszeń w stosunku do liczby zgłoszeń dokonanych w ramach każdego serwisu IT.
- Wykres 5. Wskaźniki usuwalności dla łącznej liczby zgłoszeń dokonanych przez indywidualnych użytkowników.
- Wykres 6. Wskaźniki usuniętych zgłoszeń w ramach eskalacji z podziałem na VLOP-y.
- Wykres 7. Wskaźnik naruszeń Kodeksu karnego w stosunku do łącznej liczby zgłoszeń.

BIBLIOGRAFIA

Akty normatywne i źródła prawa

Decyzja Ramowa 2008/913/WSiSW z dnia 28 listopada 2008 r. w sprawie zwalczania pewnych form i przejawów rasizmu i ksenofobii za pomocą środków prawnokarnych, Dz.U. L 328/55 z 6.12.2008, <https://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:32008F091>.

Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/2065 z dnia 19 października 2022 r. w sprawie jednolitego rynku usług cyfrowych oraz zmiany dyrektywy 2000/31/WE (akt o usługach cyfrowych), Dz. U. UE L 277/1 z 27.10.2022, <https://eur-lex.europa.eu/legal-content/PL/TXT/PDF/?uri=CELEX:32022R2065>.

Ustawa z dnia 18 grudnia 1998 r. o Instytucie Pamięci Narodowej – Komisji Ścigania Zbrodni przeciwko Narodowi Polskiemu, Dz. U. z 2023 r. poz. 102, <https://dziennikustaw.gov.pl/D2023000010201.pdf>

Ustawa z dnia 6 czerwca 1997 r. - Kodeks karny, Dz. U. 1997, nr 88, poz. 553, <https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU19970880553>.

Ustawa z dnia 18 lipca 2002 r. o świadczeniu usług drogą elektroniczną, Dz. U. 2002, nr 144, poz. 1204, <https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20021441204>.

Inne źródła

Civic Resilience Initiative, Detector Media i Instytut Kościuszki, "Odporność na dezinformację w krajach trójkąta lubelskiego", 2022, https://ik.org.pl/wp-content/uploads/2023/08/Resilience-to-Disinformation0612_4.pdf.

Council of Europe, Recommendation No. R (97) 20 of the Committee of Ministers to member states on "hate speech", adopted by the Committee of Ministers on 30 October 1997, <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680505d5b>.

Council of Europe, Recommendation CM/Rec (2022) 16 of the Committee of Ministers to Member States on combating hate speech, adopted by the Committee of Ministers on 20 May 2022, <https://search.coe.int/cm?i=0900001680a67955>.

Institute for Strategic Dialogue, B'nai B'rith International, "Online Antisemitism: A Toolkit for Civil Society" 2022, <https://www.isdglobal.org/isd-publications/online-antisemitism-a-toolkit-for-civil-society/>

Institute for Strategic Dialogue, "Narratives of Hate: Post-7 October Antisemitism and Anti-Muslim Hate on Social Media", 2024, <https://www.isdglobal.org/isd-publications/narratives-of-hate-post-7-october-antisemitism-and-anti-muslim-hate-on-social-media/>.

Międzynarodowy Sojusz na rzecz Pamięci o Holokauście, "Robocza definicja antysemityzmu", <https://holocaustremembrance.com/resources/robocza-definicja-antysemityzmu-przyjeta-przez-miedzynarodowy-sojusz-na-rzecz-pamieci-o-holokauscie>

Ministerstwo Kultury i Dziedzictwa Narodowego, "Oświadczenie w sprawie. definicji IHRA", <https://www.gov.pl/web/archiwum-kulturaiseport/oswiadczenie-w-sprawie-definicji-ihra>.

Komisja Europejska, "Nadzór nad wyznaczonymi bardzo dużymi platformami internetowymi i wyszukiwarkami w ramach aktu o usługach cyfrowych", <https://digital-strategy.ec.europa.eu/pl/policies/list-designated-vlops-and-vloses>.

Komisja Europejska, "Unijny akt o usługach cyfrowych. Bezpieczeństwo w zakresie zwalczania nielegalnego nawoływania do nienawiści w internecie", https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_pl.

Komisja Europejska, "Unijny kodeks postępowania w zakresie zwalczania nielegalnego nawoływania do nienawiści w internecie", https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-counteracting-illegal-hate-speech-online_en?prefLang=pl&etran=pl

Żydowskie Stowarzyszenie Czulent, "Antysemickie incydenty w 2023 roku. Badanie doświadczenia społeczności żydowskiej", 2024, <https://czulent.pl/antysemityczne-incydenty-w-roku-2023-badanie-doswiadczenia-spolecznosci-zydowskiej/>.

WIDZISZ → REAGUJ → ZGŁOŚ ANTYSEMITYZM

Dlaczego trzeba zgłaszać incydenty antysemickie?

Traktujemy poważnie każde doświadczenie, które postrzegasz jako antysemickie. Twoja perspektywa jako ofiary lub świadka jest w naszej pracy najważniejsza. Stosujemy podejście zorientowane na ofiary i koncentrujemy się na ich potrzebach.

Nasza rola jest unikalna: przyjmujemy zgłoszenia incydentów, które nie zawsze stanowią przestępstwa w świetle prawa. Wszystkie dane, które dostarczysz, są traktowane jako poufne. Ty podejmujesz decyzję, co do ich dalszego wykorzystania.

Chcemy poinformować cię o możliwościach uzyskania pomocy: doradztwo antydyskryminacyjne, psychospołeczne i prawne.

Jak zgłaszać incydenty o charakterze antysemickim?

za pośrednictwem strony
www.zglosantysemityzm.pl



przez e-mail:
info@zglos-antysemityzm.pl

lub telefonicznie pod numerem
[12 400 00 08](tel:12400008)
(godziny pracy na stronie www.zglosantysemityzm.pl)

Żydowskie Stowarzyszenie Czulent

Żydowskie Stowarzyszenie Czulent to niezależna organizacja non-profit, działająca zarówno na poziomie krajowym, jak i międzynarodowym, koncentrująca się głównie na działaniach rzeczniczych. Nasza platforma zrzesza profesjonalistów ze społeczności żydowskiej w Polsce oraz za granicą.

Czulent podejmuje kompleksowe inicjatywy mające na celu przeciwdziałanie antysemityzmowi, opracowując analizy i raporty dotyczące tego zjawiska w krajach Grupy Wyszehradzkiej oraz prowadząc działania z zakresu litygacji strategicznej. Stworzona przez Stowarzyszenie platforma internetowa zglosantysemityzm.pl umożliwia raportowanie incydentów i przestępstw o charakterze antysemickim oraz oferuje wsparcie prawne dla osób nimi pokrzywdzonych.

Działania Czulentu obejmują aspekty polityczne, społeczne i prawne, realizowane poprzez wdrażanie innowacyjnych rozwiązań edukacyjnych oraz budowanie koalicji na rzecz otwartości, przeciwdziałania antysemityzmowi, rasizmowi i dyskryminacji. Współpracujemy z instytucjami, administracją publiczną oraz organizacjami dialogu, aby wpływać na zmianę postaw społecznych oraz kształtowanie polskiego ustawodawstwa w zakresie tolerancji i zwalczania rasizmu. Wśród naszych partnerów znajdują się m.in. Biuro Instytucji Demokratycznych i Praw Człowieka OBWE (ODIHR), American Jewish Committee Central Europe oraz National Democratic Institute.

W ramach koalicji międzynarodowych — European Network on Monitoring Antisemitism (ENMA), Coalition to Counter Online Antisemitism (CCOA), European Network for Countering Antisemitism through Education (ENCATE) oraz European Network Against Racism (ENAR) — gromadzimy i promujemy dobre praktyki oraz rekomendujemy rozwiązania na poziomie europejskim.



stowarzyszenie
żydowskie
czulent

 zglosantysemityzm.pl