Wyzwania prawnej ochrony dzieci w aspekcie przemocy seksualnej online

Katarzyna Katana

Fundacja Dajemy Dzieciom Siłę

Celem artykułu jest omówienie wyzwań w zakresie prawnej ochrony dzieci przed seksualną cyberprzemocą. Badania pokazują, że skala zjawiska przemocy online wobec dzieci, szczególnie w aspekcie przemocy seksualnej, rośnie każdego roku.

Prawo ma na takie sytuacje wieloraką odpowiedź, nie zawsze nadążającą za omawianym problemem. Charakter prewencyjny mają z pewnością standardy ochrony dzieci, które od lutego 2024 r. należało wdrożyć w każdym podmiocie pracującym z dziećmi. W ramach standardów powinny się znaleźć zasady korzystania z internetu na terenie placówki, uczenia dzieci bezpiecznych zachowań w sieci, zasady bezpiecznych relacji pomiędzy personelem a dziećmi oraz bezpiecznych relacji rówieśniczych, czyli zbiory zachowań pożądanych i zabronionych względem dzieci.

Co jednak, gdy profilaktyka nie zadziała i trzeba podjąć interwencję? W sytuacji cyberprzemocy dochodzi do styku prawa rozumianego tradycyjnie (do tego zbioru należy np. ścieżka reakcji sądu rodzinnego w przypadku demoralizacji/czynu karalnego, popełnionego przez osobę poniżej odpowiednio 18. i 17. roku życia, bądź sądu karnego w przypadku materiałów stanowiących wykorzystanie seksualne dzieci) i prawa nowych technologii, które wymaga od dostawców usług reakcji na treści szkodliwe poprzez ich moderowanie w oparciu o zgłoszenie, zapewnienie mechanizmów skargowych, a także dbanie o dobro małoletnich. W chwili obecnej dostawcy usług internetowych nie są jednak zobowiązani do proaktywnego sprawdzania, czy w ramach dostarczanych przez nich usług nie dochodzi do krzywdzenia dzieci.

Unijny akt o usługach cyfrowych (DSA), tj. 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, obowiązuje we wszystkich krajach członkowskich od 17 lutego 2024 r., nawet jeśli polski ustawodawca nie dostosował jeszcze polskiego prawa w tym zakresie.

Parlament Europejski pracuje nad nowelizacją Dyrektywy Parlamentu Europejskiego I Rady 2011/92/UE z dnia 13 grudnia 2011 r. w sprawie zwalczania niegodziwego traktowania w celach seksualnych i wykorzystywania seksualnego dzieci oraz pornografii dziecięcej, zastępującą decyzję ramową Rady 2004/68/WSiSW, tak aby uwzględniała przestępstwa popełniane wobec dzieci z użyciem internetu. Stale trwają (niezakończone) prace nad Rozporządzeniem Parlamentu Europejskiego i Rady ustanawiającym przepisy mające na celu zapobieganie niegodziwemu traktowaniu dzieci w celach seksualnych i jego zwalczanie.

Celem artykułu jest omówienie zapisów tych trzech aktów – obowiązujących i projektowanych – przez pryzmat reagowania na przemoc seksualną wobec dzieci popełnianą z wykorzystaniem nowych technologii. Autorka zastanawia się, czy obecna regulacja chroniąca dzieci przed krzywdzeniem online jest kompletna bądź czy taka się stanie po przyjęciu projektowanych zmian. Autorka wskaże na ewentualne trudności w zakresie reagowania na tego typu czyny i rolę dostawców usług internetowych w tym swoistym półprywatnym – półpublicznym dbaniu o bezpieczeństwo dzieci w sieci.

Słowa kluczowe:
przemoc online, przemoc seksualna, ochrona dzieci, internet

Wstęp

Wykorzystywanie seksualne dzieci, niezależnie od formy tego zachowania i jego późniejszej oceny karnoprawnej, jest wydarzeniem traumatycznym w życiu osoby mu poddanej (Studnicki i in., 2021). W polskim Kodeksie karnym (t.j. Dz.U. z 2025 r., poz. 383) stypizowano szereg przestępstw seksualnych przeciwko dzieciom, także tych popełnianych wyłącznie lub w dużej mierze z użyciem technologii informatycznych i sieci internetowej. W innych krajach należących do Unii Europejskiej prawnokarna ocena zachowań seksualnych osób dorosłych wobec dzieci również przeszła dynamiczną zmianę w ostatnim trzydziestoleciu. Jeszcze w latach 90. XX w. materiały o charakterze pornograficznym z udziałem dzieci oceniano pod kątem ich „przyzwoitości" (Specjalny sprawozdawca ONZ, 1992) zamiast ich jednoznacznego zakwalifikowania jako szkodliwego utrwalenia krzywdy wyrządzonej dziecku przez osobę dorosłą. Charakter materiałów pornograficznych portretujących dzieci powoduje, że wyrządzona krzywda będzie trwała także po osiągnięciu przez dziecko pełnoletności, z uwagi na ciągłe istnienie i rozpowszechnianie materiałów tę krzywdę dokumentujących, z tego też względu

współczesna terminologia posługuje się pojęciem „materiał stanowiący wykorzystanie seksualne dzieci" (CSAM) w miejsce pojęcia „pornografia".

Terminologia ta zgodna jest z postulatami organizacji pozarządowych i międzynarodowych (ECPAT, 2025)[1], które wskazują, że produkcja materiałów pornograficznych wiąże się ze zgodą osób dorosłych w nich sportretowanych, w przeciwieństwie do tzw. pornografii dziecięcej, która stanowi w istocie udokumentowanie seksualnego wykorzystywania dziecka jako osoby niezdolnej do wyrażenia zgody do udziału w takim przedsięwzięciu. W opracowaniu unikano w miarę możliwości używania pojęcia „pornografia dziecięca", choć występuje ono wciąż w większości krajowych i międzynarodowych aktów prawnych (np. w Dyrektywie 2011/92/UE przed jej – wciąż trwającą – nowelizacją).

Długie trwanie krzywdy umożliwił dodatkowo rozwój sieci internetowej, która stała się głównym miejscem przechowywania tego typu materiałów i obrotu nimi. Czyny związane z produkowaniem i rozpowszechnianiem materiałów pornograficznych z udziałem dzieci nie są jednak jedynymi przestępstwami o charakterze seksualnym, które mogą funkcjonować w rzeczywistości online. Czynami istotnymi dla celów niniejszego opracowania są:

- grooming (art. 200a k.k.);
- produkowanie w celu rozpowszechnienia i rozpowszechnianie treści pornograficznych z udziałem małoletniego (art. 202 §3 k.k.);
- utrwalanie treści pornograficznych z udziałem małoletniego (art. 202 §4 k.k.);
- posiadanie treści pornograficznych z udziałem małoletniego (art. 202 §4a k.k.);
- produkowanie lub posiadanie treści pornograficznych prezentujących wytworzony lub przetworzony wizerunek małoletniego uczestniczącego w czynnościach seksualnych (art. 202 §4 k.k.).

Poprzedzać lub następować po nich (w przypadku groomingu) mogą także następujące czyny zabronione:

- utrwalanie treści pornograficznych z udziałem małoletniego (art. 202 §4 k.k.);
- zgwałcenie, w tym zgwałcenie osoby poniżej 15. roku życia (art. 197 §1 i §4 k.k.);
- Obcowanie płciowe z małoletnim poniżej 15. roku życia (art. 200 §1 k.k.);
- seksualne wykorzystanie zależności lub krytycznego położenia (art. 199 k.k.).

1 Wytyczne Luksemburskie, opracowane w 2016 r. przez Interagency Working Group, której powstanie zostało zainicjowane przez organizację ECPAT (https://ecpat.org/luxembourg-guidelines/). W 2025 r. wydano ich drugą edycję.

Skala zjawiska przestępstw seksualnych wobec dzieci popełnianych za pomocą sieci internetowej może zostać uznana za niedoszacowaną, a statystyki są znacząco rozbieżne, m.in. z uwagi na lokalizację poddawanych ocenie materiałów.

Raport Dyżurnet.pl za rok 2023 wskazuje, że w sumie zgłoszono 18 279 przypadków nielegalnych treści w internecie, z czego 7358 dotyczyło podejrzenia wykorzystania seksualnego dzieci. Z tej liczby jako CSAM zakwalifikowano 2612 zgłoszeń, jako inne treści seksualizujące dzieci (np. modeling, epatowanie nagością) – 587. Dodatkowo zakwalifikowano 35 incydentów jako propagowanie pedofilii, a 10 – jako przypadki groomingu. Jedynie 2% wszystkich stwierdzonych materiałów znajduje się pod adresem serwerów znajdujących się w Polsce (NASK, 2024). Lokalizacja serwerów w odniesieniu do adresów stron oraz lokalizacji plików warunkuje sposób reakcji na zidentyfikowane materiały. Jak pokazuje treść raportu, większość materiałów krzywdzących dzieci w sposób seksualny umieszczona jest na zagranicznych serwerach, głównie w USA oraz Rosji.

W Polsce zadania z zakresu zgłaszania i analizy materiałów stanowiących niegodziwe traktowanie dzieci w celach seksualnych prowadzi Zespół Reagowania na Incydenty Bezpieczeństwa Komputerowego działający na poziomie krajowym, prowadzony przez Naukową i Akademicką Sieć Komputerową – Państwowy Instytut Badawczy (CSIRT NASK). Zadanie to przypisano NASK ustawą o krajowym systemie cyberbezpieczeństwa (t.j. Dz.U. z 2024 r., poz. 1077 z późn. zm.). Jak wskazano w uzasadnieniu do ustawy (s. 37), stworzenie kompleksowego systemu bezpieczeństwa oznacza nie tylko ochronę użytkowników przed incydentami bezpieczeństwa, ale także przed nielegalnymi treściami, do jakich niewątpliwie należą materiały pornograficzne portretujące dzieci. Zadanie to przydzielono NASK, gdyż od 2005 r. prowadzi on projekt Dyżurnet.pl. W jego ramach możliwe jest zgłoszenie (za pomocą formularza, mailowo, telefonicznie oraz za pomocą wtyczki zainstalowanej w przeglądarce) niebezpiecznych treści w internecie.

Z kolei statystyki organizacji INHOPE, zrzeszającej „gorące linie", działające jako „zaufani sygnaliści", czyli podobnie jak polski Dyżurnet.pl informujące serwisy i portale o nielegalnych treściach, wskazują na łączną liczbę zgłoszeń przeprocesowanych przez 54 zrzeszone linie równą 785 322 w roku 2023 przy 587 852 przypadkach w roku 2022. Liczba ta nie uwzględnia zgłoszeń, których nie uznano za nielegalne treści. Jak wskazuje wyżej wymieniony raport, 88% zidentyfikowanych treści mieści się w kategorii materiałów nowych, czyli dotąd niezidentyfikowanych jako CSAM.

Zgłoszenia tzw. *trusted flaggers* są traktowane priorytetowo przez platformy internetowe. Mechanizm rejestracji doświadczonych podmiotów i organizacji w roli zaufanych sygnalistów został opisany w rozporządzeniu DSA, o którym szerzej niżej.

Międzynarodowe instytucje informują, że kraje Unii Europejskiej są jednym ze znaczących miejsc popełniania przestępstw seksualnych na szkodę dzieci z wykorzystaniem internetu (Clark, 2020), co może świadczyć o pilnej potrzebie uregulowania tego zagadnienia na poziomie unijnym. Regulacja prawna pojedynczych krajów może wywierać pewien, ograniczony wpływ na globalne korporacje, jednakowoż transgraniczny charakter internetu przemawia za wnioskiem, że pozycja negocjacyjna Wspólnoty i wpływ jednolitego prawa będzie znacząco większy, niż gdyby podobne regulacje przyjęło pojedyncze państwo. Na marginesie wskazać jednak należy, że Wielka Brytania przyjęła w roku 2023 akt prawny, który nałożył konkretne zobowiązania na dostawców usług umożliwiających dzielenie się treścią, w tym m.in. w zakresie skutecznej kontroli wieku użytkowników czy też analizy ryzyka stwarzanego przez serwis dla grup wrażliwych (w szczególności dzieci) (UK Online Safety Act). W rozumieniu art. 1 dzieckiem lub małoletnim jest każda osoba poniżej 18. roku życia (Dz.U. z 1991 r., nr 120, poz. 526 z późn. zm.).

Na potrzeby niniejszego opracowania wyróżniono dwie drogi rozpowszechniania materiałów stanowiących wykorzystywanie seksualne dzieci – w drodze publikacji przez użytkowników platform/serwisów internetowych oraz w elektronicznej komunikacji prywatnej pomiędzy użytkownikami.

W grudniu 2020 r. korporacje technologiczne prowadzące media społecznościowe zaczęły publikować na łamach swoich platform oświadczenia o zaprzestaniu lub kontynuowaniu skanowania treści umieszczanych na tychże portalach z obszaru Unii Europejskiej w poszukiwaniu materiałów stanowiących wykorzystywanie seksualne dzieci. Opisana sytuacja miała swoje źródło w fakcie wejścia w życie tzw. Europejskiego Kodeksu Łączności Elektronicznej (Dz.U. UE. L. z 2018 r., nr 321, s. 36 z późn. zm., dalej jako Kodeks lub EKŁE) i jako taka nie była zamierzeniem unijnego ustawodawcy. W wyniku przyjęcia Kodeksu doszło bowiem do uchylenia dyrektywy 2002/21/WE[2] – w jej miejsce, jak i trzech innych dyrektyw[3] wszedł EKŁE. Podobnie

2 Dyrektywa 2002/21/WE Parlamentu Europejskiego i Rady z dnia 7 marca 2002 r. w sprawie wspólnych ram regulacyjnych sieci i usług łączności elektronicznej (dyrektywa ramowa)

3 Z dniem 21 grudnia 2020 roku uchylono też: Dyrektywę 2002/19/WE Parlamentu Europejskiego i Rady z dnia 7 marca 2002 r. w sprawie dostępu do sieci łączności elektronicznej i urządzeń towarzyszących oraz wzajemnych połączeń (dyrektywa o dostępie) (Dz.U. L 108 z dnia 24 kwietnia 2002, s. 7), Dyrektywę 2002/20/WE Parlamentu Europejskiego i Rady z dnia 7 marca 2002 r. w sprawie zezwoleń na udostępnienie sieci i usług łączności elektronicznej (dyrektywa o zezwoleniach) (Dz.U. L 108 z dnia 24 kwietnia 2002, s. 21) oraz Dyrektywę 2002/22/WE Parlamentu Europejskiego i Rady z dnia 7 marca 2002 r. w sprawie usługi powszechnej i związanych z sieciami i usługami łączności elektronicznej praw użytkowników (dyrektywa o usłudze powszechnej) (Dz.U. L 108 z dnia 24 kwietnia 2002, s. 51).

jak w innych porządkach prawnych, także w unijnych aktach legislacyjnych dochodzi do zamieszczania odesłań do definicji zawartych już w innych aktach prawnych. Dyrektywa 2002/21/WE zawierała definicje legalne, do których odsyłała dyrektywa 2002/58/WE[4] – zwana dalej też dyrektywą o e-Prywatności. Przez uchylenie dyrektywy i zastąpienie jej innym aktem prawnym doszło do zmiany definicji łączności elektronicznej. Usługi świadczone przez korporacje prowadzące portale społecznościowych oraz świadczące usługi e-mail zostały objęte zakresem zastosowania dyrektywy o e-Prywatności, przewidującej po stronie dostawców usług obowiązek zapewnienia poufności komunikacji użytkowników.

Czynności skanowania treści, realizowane dotąd przez dostawców usług mediów społecznościowych oraz poczty e-mail, miały charakter ich dobrowolnej inicjatywy i podlegały co najwyżej pod regulację ogólnego rozporządzenia o ochronie danych osobowych (Dz.U. UE. L. z 2016 r., nr 119, s. 1 z późn. zm.). Jednocześnie ich niewątpliwy walor przejawiał się zarówno w możliwości wykrycia osób gromadzących lub rozpowszechniających już zidentyfikowane materiały stanowiące wykorzystywanie seksualne dzieci, jak też usunięcia tych treści z sieci przed dostaniem jakiejkolwiek informacji od użytkowników lub sygnalistów, może nawet zanim ktokolwiek się z takim materiałem zapoznał.

Wskaźniki, że sytuacja wymaga pilnej interwencji unijnego ustawodawcy, pojawiły się szybko i miały wielką skalę. Okazało się bowiem, że liczba zgłoszeń przypadków wykrycia treści stanowiących wykorzystywanie seksualne dzieci dokonywanych zgodnie z tamtejszym prawem przez korporacje technologiczne do podmiotu przyjmującego[5], lecz dotyczących przypadków zidentyfikowanych na obszarze Unii Europejskiej w pierwszym miesiącu obowiązywania Europejskiego Kodeksu Łączności Elektronicznej, zmalała o ponad 50% (National Center of Missing and Exploited Children, 2020) (więcej o podstawach prawnych dotychczasowych zgłoszeń w dalszej części artykułu). Porównanie skali zgłoszeń sprzed wejścia w życie EKŁE i po nim prowadziło do jednoznacznego wniosku, że zgłoszenia niedokonane przez przedsiębiorstwa technologiczne powiększają ciemną liczbę przestępstw seksualnych na szkodę dzieci na obszarze całej UE.

Celem niniejszego opracowania jest przedstawienie problematyki, która stała się – choć drogą przypadku – zaczątkiem szerokich działań Unii Europejskiej

4 Dyrektywa Parlamentu Europejskiego i Rady z dnia 12 lipca 2002 r. dotycząca przetwarzania danych osobowych i ochrony prywatności w sektorze łączności elektronicznej (dyrektywa o prywatności i łączności elektronicznej).

5 Dokładnie NCMEC – the National Center for Missing & Exploited Children.

prewencyjnych w przedmiocie praw dzieci w świecie online. Zostaną omówione poprzedni i obecny stan prawny, który doprowadził do zaprzestania zgłoszeń po stronie dostawców usług, a także regulacja amerykańska, dotychczas pośrednio służąca prewencji przestępstw seksualnych wobec dzieci na obszarze kontynentu europejskiego. Rzeczona prewencja miała miejsce poprzez proaktywne działania dostawców usług, zmierzające do wykrycia materiałów i następnie zgłaszanie wykrytych materiałów CSAM upoważnionych organów. Zgłaszanie odbywało się do podmiotów wyspecjalizowanych w zwalczaniu przestępczości online wobec dzieci, zlokalizowanych w Stanach Zjednoczonych. Raportowane przypadki pochodziły jednak z całego świata.

Następnie zostaną poddane analizie przepisy prawne wydane i projektowane celem zapewnienia bezpieczeństwa dzieci online. Rozważeniu będzie podlegać, jakie są podstawy do działań ustawodawczych Unii Europejskiej w zakresie ochrony dzieci, a także jak powinien zostać rozstrzygnięty (przynajmniej według przepisów i standardów międzynarodowych) rysujący się konflikt wartości – prawa do prywatności i ochrony tajemnicy korespondencji użytkowników internetu oraz konkretnych praw dzieci.

Aspekt technologiczny

Przed szczegółowym omówieniem aspektów prawnych należy krótko opisać technologię używaną do odnajdywania treści przedstawiających wykorzystywanie seksualne dzieci. Omówienie to nie będzie wolne od uproszczeń z uwagi na fakt, że artykuł ma charakter prawny, a nie technologiczny, a zagadnienia technologiczne mają znaczenie pomocnicze dla poruszanej problematyki.

Oprogramowanie zwane PhotoDNA, wykorzystujące technologię haszowania percepcyjnego, po raz pierwszy zaprezentowała w 2009 r. firma Microsoft. Motywacją do stworzenia przedmiotowego oprogramowania była zarówno chęć uchronienia użytkowników przed nielegalnymi treściami, jak i wartości wynikające ze społecznej odpowiedzialności biznesu. PhotoDNA wykorzystuje dotąd nieujawniony algorytm, a podmioty, którym został udostępniony, zobowiązane są kontraktowo do zachowania tajemnicy.

Użyta technologia pozwala na wyszukiwanie w internecie kopii materiałów obrazujących wykorzystywanie seksualne dzieci, znanych już podmiotom uprawnionym do ich zwalczania. Przed wprowadzeniem omawianego rozwiązania każde zdjęcie (bądź film) było indywidualnie oceniane przez osobę pracującą w instytucji lub organizacji zajmującej się zwalczaniem szkodliwych treści w internecie – także wtedy, gdy

stanowiło kopię wykrytego już wcześniej materiału. Obecnie znane zdjęcia zawierające kwestionowany materiał – uprzednio zgromadzone np. przez jednostki rządowe zajmujące się przeciwdziałaniem tego typu praktykom lub organizacje pozarządowe z różnych części świata – są podpisywane za pomocą percepcyjnych haszów[6]. W pierwszej kolejności obraz jest dostosowywany do standardowego rozmiaru 26×26 pikseli, następnie modyfikowany w celu odzwierciedlenia w odcieniach szarości, a w kolejnym kroku – dzielony na nachodzące na siebie (o dwa piksele) zbiory pikseli. Każdy ze zbiorów odzwierciedla konkretny gradient (gradient w uproszczeniu oznacza różnicę pomiędzy wartościami poszczególnych pikseli). Każda z części jest następnie podpisywana za pomocą czterech wartości liczbowych, a funkcja haszująca podpisująca całe zdjęcie ma łącznie 144 liczby (w zdjęciu wyodrębnia się 36 nachodzących na siebie zbiorów, a każdy zbiór podpisuje czterema wartościami). Nie jest możliwe ponowne odtworzenie zdjęcia na podstawie funkcji haszującej[7], co służyć ma przede wszystkim ochronie prywatności osób sportretowanych w materiałach, a także uniemożliwieniu ich rozprzestrzenia tą drogą. Dzięki technologii haszowania percepcyjnego podmioty zajmujące się wyszukiwaniem szkodliwych treści nie muszą ich posiadać na własnych serwerach (co grozi też np. ich wyciekiem), a skanowanie prowadzą wyłącznie w oparciu o bibliotekę haszów. Podpisane zdjęcie (w formie ciągu znaków) trafia do biblioteki haszów, która na późniejszym etapie służy do porównania znajdowanych materiałów z dotychczas odkrytymi treściami stanowiącymi wykorzystywanie seksualne dzieci. Biblioteki podpisanych zdjęć tworzą zaufani dostawcy – z reguły organizacje pozarządowe zajmujące się wyszukiwaniem i zgłaszaniem tego typu treści. Algorytm PhotoDNA jest odporny na zmiany w kopiach obrazu, dokonane np. poprzez jego powiększenie, pomniejszenie, zmianę kolorystyczną – wciąż odnajduje podobieństwa porównywanych obrazów (hasze zdjęcia porównywanego, które zostało zmienione, oraz zdjęcia użytego do porównania są jedynie nieznacznie różne), co byłoby niemożliwe, gdyby funkcją haszującą było podpisywane wyłącznie całe, oryginalne zdjęcie. W takiej sytuacji najmniejsza zmiana w jego dalej rozprowadzanej kopii – np. poprzez przycięcie – dałaby fałszywie negatywny wynik jej porównania z biblioteką haszów. Percepcyjne haszowanie nie zapewnia jednak tego samego poziomu bezpieczeństwa, co haszowanie kryptograficzne. W tej technologii istnieje możliwość wystąpienia dwóch jednakowych

6 Jest to elektroniczny podpis zdjęcia, wykonany za pomocą ciągu znaków (numerów), którego zadaniem jest odwzorowanie ludzkiego (tj. wzrokowego) postrzegania różnic pomiędzy dwoma obrazami, poprzez zapisanie różnic w wartościach poszczególnych pikseli zdjęcia (Farid Hany, 2021).

7 Choć fakt ten jest podawany w wątpliwość (por. Krawetz, 2021).

podpisów dla różnych zdjęć, co rodzi perspektywę dla obrońców ewentualnych oskarżonych kwestionowania wyników skanowania przez PhotoDNA, dopóki nie zostanie zaprezentowane oryginalne zdjęcie – zarówno odnalezione w materiałach użytkownika, jak i użyte do porównania.

Technologię haszowania percepcyjnego wdrożono także w celu wyszukiwania materiałów ekstremistycznych i o charakterze terrorystycznym.

W 2018 r. Microsoft we współpracy z Internet Watch Foundation rozwinął usługę o wyszukiwanie już znanych filmów wideo z wizerunkiem dzieci wykorzystywanych seksualnie. Narzędzie umożliwia rozpoznanie kopii filmu także po zamieszczeniu jej w innym, niepowiązanym tematycznie materiale. Technologia rozpoznawania filmów wideo jest zbliżona do tej analizującej zdjęcia, z tą różnicą, że podpisywane są (także poprzez ich uprzedni podział) kluczowe klatki filmów (ang. *keyframes)*[8]. Sprawdzenie materiału odbywa się poprzez analizę podpisów tych klatek.

Technologia jest obecnie wykorzystywana przez Microsoft w jego własnych serwisach[9], udostępniona Krajowemu Centrum do spraw Zaginionych i Wykorzystywanych Dzieci[10], licencjonowana dużym przedsiębiorstwom technologicznym[11] oraz udostępniona bezpłatnie mniejszym podmiotom (w tym organizacjom pozarządowym), które oferują możliwość przechowywania przez użytkowników wytworzonych przez nich treści. Bezkosztowa wersja udostępniana jest od 2015 r. na platformie chmurowej Azure jako narzędzie niewymagające dalszych nakładów w celu wdrożenia (Microsoft, 2025). Wersja chmurowa na ten moment nie posiada funkcji identyfikowania kopii nagrań wideo ani nie jest udostępniania organom ścigania (które otrzymują wersję pełną).

Facebook w 2019 r. wdrożył własne narzędzie rozpoznawania zdjęć i filmów wideo stanowiących wykorzystywanie seksualne dzieci – PDQ. W przeciwieństwie do oprogramowania Microsoft algorytm używany przez Facebook jest publiczny i udostępniany na zasadach open source. PDQ działa na podobnej zasadzie, jak PhotoDNA. W pierwszej kolejności obraz jest standaryzowany do rozmiaru 64×64 pikseli, dzielony i odzwierciedlany w odcieniach szarości, a następnie podpisywany 256-bitowym binarnym ciągiem cyfr.

Apple własną technologię rozpoznawania materiałów o charakterze przestępczości seksualnej wobec dzieci zaprezentował w roku 2021. W przeciwieństwie

8 To klatki, która odwzorowują początek i koniec danej sceny.

9 Xbox, LinkedIn, OneDrive, wyszukiwarka Bing.

10 The National Center for Missing & Exploited Children.

11 Facebook, Twitter, Google, Reddit, Discord.

do dwóch poprzednio omawianych narzędzie to zapisuje za pomocą ciągów liczb nie precyzyjne wartości poszczególnych pikseli, a unikalne cechy danego zdjęcia. Skanowanie treści, też w przeciwieństwie do poprzednio omawianych systemów, odbywa się na urządzeniu końcowym użytkownika, a nie w chmurze (Apple Inc., 2021). Ruch ten ma zagwarantować większe poszanowanie prawa do prywatności. Mimo to rozwiązanie przygotowane przez Apple spotkało się z dużą falą krytyki, która ostatecznie opóźniła jego pełne wdrożenie.

W 2020 r. Microsoft zaprezentował narzędzie służące wykrywaniu przypadków groomingu (Gregoire, 2022). Początki tej technologii miały miejsce w trakcie międzybranżowego wydarzenia Hackathon w 2018 r., w całości poświęconemu problematyce wykorzystywaniu dzieci online. Hany Farid, współtwórca technologii PhotoDNA, wraz z zespołem z Dartmouth College w kolejnych miesiącach opracował narzędzie umożliwiające firmom technologicznym analizę treści wymienianych w czacie użytkowników w celu wykrycia prób groomingu. Według doniesień jednym z najczęstszym miejsc w świecie online, w których dochodzi do groomingu, są platformy oferujące możliwość gry w wieloosobowych rozgrywkach komputerowych, które posiadają z reguły wbudowaną funkcję czatu, umożliwiającą porozumiewanie się pomiędzy użytkownikami (Bowles, Keller, 2019). Opracowane oprogramowanie analizuje automatycznie treść wymienianych wiadomości w poszukiwaniu sformułowań pochodzących z historycznych, zarejestrowanych rozmowy mających na celu uwiedzenie małoletniego. W przypadku wykrycia takich treści ma zostać powiadomiony moderator, który może dodatkowo zweryfikować rozmowę, a także zaalarmować organy ścigania.

Paradoksalnie jednak kolejne plany podmiotów z branży technologicznej mogą spowodować, że opracowane narzędzia staną się niefunkcjonalne. Ostatnie lata poprzedniej dekady przyniosły deklaracje o powszechnym wprowadzeniu w usługach internetowych szyfrowania end-to-end – zarówno w zakresie wymiany wiadomości, jak i ustawień kopii zapasowej urządzenia. Szyfrowanie end-to-end oznacza, że przesyłanie treści będzie można odszyfrować dopiero na urządzeniu końcowym użytkownika, co uniemożliwi skanowanie treści przez programy typu PhotoDNA i w konsekwencji sparaliżuje działania organów ścigania (Barr i in., 2019). Obecne postulaty zakładają wbudowanie wyszukiwania pasujących haszów w mechanizm szyfrowania, co spowoduje, że całość treści posiadanych przez użytkownika pozostanie zaszyfrowana, o ile nie dojdzie do odnalezienia obrazu pasującego do znanego hasza.

Dotychczasowy stan prawny

Dyrektywa 2002/58/WE o e-Prywatności zapewnia użytkownikom usług łączności elektronicznej szerszą ochronę niż Ogólne Rozporządzenie o Ochronie Danych Osobowych. Stosuje się ją bowiem do ochrony prywatności użytkowników także wtedy, gdy nie dochodzi do przetwarzania ich danych osobowych. Do dookreślenia zakresu zastosowania dyrektywy służyć miały definicje zawarte w dyrektywie 2002/21/WE. Według treści uchylonej dyrektywy 2002/21/WE (art. 2 c) „usługa łączności elektronicznej" oznaczała „usługę zazwyczaj świadczoną za wynagrodzeniem, polegającą całkowicie lub częściowo na przekazywaniu sygnałów w sieciach łączności elektronicznej, w tym usługi telekomunikacyjne i usługi transmisyjne świadczone poprzez sieci nadawcze; nie obejmuje jednak usług związanych z zapewnianiem albo wykonywaniem kontroli treści przekazywanych przy wykorzystaniu sieci lub usług łączności elektronicznej". Spod zakresu definicji wyłączono usługi społeczeństwa informacyjnego w rozumieniu art. 1 dyrektywy 98/34/WE, o ile nie polegały całkowicie lub częściowo na przekazywaniu sygnałów w sieciach łączności elektronicznej.

W Kodeksie Łączności Elektronicznej (KŁD), który zastąpił dyrektywę 2002/21/WE, usługi łączności elektronicznej określono znacznie szerzej. Wobec dynamicznym zmian technologicznych i społecznych od czasu przyjęcia dyrektywy 2002/21/WE[12] zdecydowano się sformułować definicje w ten sposób, aby były neutralne technologicznie (motyw 14 KŁD) i ujmowały istotę definiowanego zagadnienia nawet na etapie dalszego rozwoju techniki. Definicja usług łączności elektronicznej została sformułowana opisowo (art. 2 pkt EKŁE) jako usługa świadczona za wynagrodzeniem za pośrednictwem sieci łączności elektronicznej, a w jej ramach wyróżniono trzy rodzaje usług: dostępu do internetu, przekazywania sygnału oraz – kluczową z punkty widzenia niniejszej pracy – łączności interpersonalnej. Ta ostatnia została z kolei podzielona na usługę wykorzystującą i niewykorzystującą numerów. Usługę łączności interpersonalnej zdefiniowano (art. 2 pkt 5 EKŁE) jako „usługę zazwyczaj świadczoną za wynagrodzeniem, która umożliwia bezpośrednią interpersonalną i interaktywną wymianę informacji za pośrednictwem sieci łączności elektronicznej między skończoną liczbą osób, w ramach której osoby inicjujące połączenie lub uczestniczące w nim decydują o jego odbiorcy lub odbiorcach, natomiast nie

12 Na marginesie warto wskazać, że Kodeks Łączności Elektronicznej jako dyrektywa powinien zostać transponowany do prawa polskiego, co zostało dokonane przyjęciem ustawy z dnia 12 lipca 2024 r. – Prawo komunikacji elektronicznej (Dz.U. poz. 1221).

obejmuje ona usług, które umożliwiają interpersonalną i interaktywną komunikację wyłącznie jako podrzędną funkcję dodatkową, która jest nieodłącznie związana z inną usługą". Z kolei usługa łączności interpersonalnej niewykorzystująca numerów została określona jako (art. 2 pkt 7 EKŁE) „usługa łączności interpersonalnej, która nie łączy się z publicznie nadanymi zasobami numeracyjnymi, mianowicie numerem lub numerami z krajowych lub międzynarodowych planów numeracji, ani nie umożliwia połączenia z numerem lub numerami z krajowych lub międzynarodowych planów numeracji". Tym samym zarówno usługi mediów społecznościowych, jak i komunikacja e-mail z uwagi na swoją specyfikę – narzędzia do komunikacji międzyludzkiej, która jest podstawową, nie dodatkową funkcją tych usług – znalazły się w zakresie definicji usługi łączności elektronicznej pod postacią łączności interpersonalnej niewykorzystującej numerów. Przeciwny wniosek należy sformułować w stosunku do serwisów umożliwiających grę typu multiplayer, gdyż w ich przypadku funkcja komunikacyjna ma charakter dodatkowy do podstawowej usługi – rozgrywki komputerowej.

Jak wskazano we wstępie, zmiana definicji komunikacji elektronicznej spowodowała, że podmioty wchodzące w zakres nowej definicji są zobowiązane do stosowania zasad bezpiecznej komunikacji, wynikających z dyrektywy o e-Prywatności. Głównymi obowiązkami, które mają znaczenie dla niniejszego opracowania, są obowiązek zachowania poufności komunikacji i związanych z nią danych o ruchu (art. 5 ust. 1) oraz anonimizacji lub usunięcia danych o wykorzystaniu usługi (o ruchu) po jej zakończeniu (art. 6 ust. 1). Obowiązki te istnieją w prawodawstwie unijnym od początków objęcia przez nie zagadnienia łączności – początkowo tylko telekomunikacyjnej, później starano się zapewnić taką samą ochronę dla wszystkich użytkowników, niezależnie od zastosowanej technologii (motyw 58). Zgodnie z art. 5 ust. 1 prawodawstwo państwo członkowskich ma zakazywać jakiejkolwiek nadzoru nad komunikacją, co wyklucza stosowanie oprogramowania proaktywnie poszukującego takich treści i skanującego wiadomości przesyłane przez użytkowników w poszukiwaniu materiałów stanowiących wykorzystywanie seksualne dzieci, a także oprogramowania wykrywającego przypadki groomingu. Z kolei zakaz przechowywania danych o ruchu z art. 6 wyklucza zarejestrowanie komunikacji z inicjatywy dostawcy usługi, w celu jej późniejszego przekazania organom ścigania.

Artykuł 15 dyrektywy 2002/58/WE przewiduje po stronie państw członkowskich uprawnienie do ustanowienia wyjątków od reguł prywatności, np. poprzez ograniczenie prawa do tajemnicy komunikacji, jeżeli jest to niezbędne z uwagi na względy wynikające z zasad demokratycznego państwa, w tym potrzebne do wykrywania i karania przestępstw. Skorzystanie z tego wyłączenia wymaga jednak

proaktywnej działalności ustawodawcy krajowego, co nie zawsze się dzieje (Komisja Europejska, 2022).

Przed wejściem w życie EKŁE jedynym aktem, który wiązał dostawców usług internetowych dobrowolnie skanujących treści, było RODO. Przetwarzanie danych osobowych ujawnionych w toku takiego skanowania (do takich należą chociażby wizerunek pokrzywdzonego bądź adres IP sprawcy[13]) było dozwolone w oparciu o art. 6 ust. 1 d) RODO – to jest w celu ochrony żywotnych interesów dziecka sportretowanego w materiałach stanowiących niegodziwe traktowanie dzieci, oraz ochrony innych dzieci potencjalnie pokrzywdzonych przez działanie sprawcy. Żywotne interesy rozumiane są jako kluczowe, ważne dla życia, o dużym znaczeniu, a brak przetwarzania rodziłby naruszenie tych interesów. Interesem dziecka sportretowanego w materiale jest zmierzanie do jak najpełniejszej ochrony jego prywatności, w tym poprzez doprowadzenie do usunięcia materiałów z sieci. Drugą przesłanką dopuszczalności przetwarzania, adekwatną do opisanych sytuacji, było przetwarzanie danych osobowych w interesie własnym przedsiębiorstwa technologicznego (art. 6 ust. 1 f RODO) celem ochrony własnych użytkowników przed nielegalnymi treściami oraz realizowania obowiązków z zakresu *corporate social responsibility*, mających również wysokie znaczenie marketingowe. W opisanej sytuacji nie zachodziły przypadki wyłączenia dopuszczalności przetwarzania cudzych danych osobowych we własnym interesie administratora, w szczególności nie stały temu na przeszkodzie prawa dzieci – podmiotów danych. Wprost przeciwnie – przetwarzanie te prawa chroniło.

Przy omówieniu nakreślonej problematyki wspomnieć należy również o dyrektywie 2011/93/UE w sprawie zwalczania niegodziwego traktowania w celach seksualnych i wykorzystywania seksualnego dzieci oraz pornografii dziecięcej. Celem dyrektywy było przede wszystkim zharmonizowanie reakcji karnej na przestępstwa seksualne przeciwko dzieciom, aczkolwiek w motywach dyrektywy wskazano na potrzebę rozpowszechnienia wiedzy o telefonach zaufania (35), szerokiej współpracy (w tym z państwami trzecimi) w celu usuwania treści stanowiących niegodziwe traktowanie dzieci w calach seksualnych (45) oraz blokowania dostępu do zidentyfikowanych stron zawierających tego typu treści, a zamieszczonych na serwerach zlokalizowanych poza UE (46). Wskazano, że dyrektywa nie stanowi przeszkody dla proaktywnych działań dostawców usług internetowych, także tych podejmowanych

13 Adres IP jako dana osobowa. Wyrok TSUE z dnia 19 października 2016 r. Patrick Breyer/Bundesrepublik Deutschland, Sprawa C-582/14, wydany na podstawie na podstawie dyrektywy 95/46/WE, ale aktualny w nowym stanie prawnym, po wejściu w życie RODO.

przy aprobacie państw członkowskich (47). W artykule 16[14] Państwa-Strony zostały zobowiązane do zapewnienia, że tajemnica zawodowa profesjonalistów nie będzie przeszkodą do zgłaszania przestępstw na szkodę dzieci (ust. 1). Z kolei w ustępie 2 przyjęto zasadę zachęcania (zatem nie obowiązku) osób mających wiedzę o przestępstwach przeciwko dzieciom do zgłaszania tego faktu. Obecnie trwają prace nad uaktualnieniem dyrektywy (Zamfir, 2025).

Ustawodawca unijny (podobnie jak ustawodawca polski) nie wprowadził dotąd obowiązków prawnych obligujących do proaktywnego skanowania przez dostawców usług internetowych treści zamieszczanych w ich zasobach celem odnalezienia, usunięcia i zaraportowania odpowiednim organom. Wprost przeciwnie, zadanie to w zakresie treści zamieszczanych samodzielnie przez użytkowników było w ocenie unijnego prawodawcy zbyt wymagające, z uwagi na ogrom treści zamieszczanych w zasobach internetowych przez użytkowników. Obowiązek tego typu został wprost wyłączony w dyrektywie o handlu elektronicznym (2000). W zakresie komunikacji prywatnej pomiędzy użytkownikowi nie istniały żadne regulacje – ani wyłączające odpowiedzialność, ani obligujące do podjęcia działań.

Legislacja amerykańska

Podobnie jak w prawie polskim i europejskim, także w Stanach Zjednoczonych dostawcy usług internetowych nie byli dotąd zobligowani do podejmowania aktywnych działań w kierunku sprawdzania zachowań użytkowników lub zamieszczanych w ich serwerach treści. Przedmiotowe uregulowania znajdują się w U.S. Code Tytuł 18, Rozdział 110 – obowiązki dostawców usług w zakresie raportowania treści to głównie §2258A[15] tego aktu.

Obowiązek prawny zgłoszenia został nałożony dopiero w momencie powzięcia przez dostawcę usługi wiedzy o naruszeniu przepisów federalnych dotyczących przestępstw seksualnych na szkodę dzieci, włączając posiadanie przez użytkowników materiałów CSAM. W przepisach zaznaczono jednak wprost, że zobowiązanie do raportowania nie oznacza nałożenia na dostawcę usług komunikacji elektronicznej obowiązku (uprzedniego) monitorowania aktywności jakiegokolwiek

14 Powtórzenie treści art. 12 Konwencji z Lanzarote.

15 U.S. Code stanowi oficjalną kompilację różnych tematycznie praw o charakterze federalnym. W przypadku dokonywania zmiany prawa są one wprowadzane w pierwotnej ustawie, niemniej jednak na potrzeby niniejszego opracowania przyjęto uproszczoną terminologię i odwołania wyłącznie do U.S. Code.

użytkownika, treści lub komunikacji któregokolwiek z klientów bądź też podejmowania aktywności zmierzającej do poszukiwania lub skanowania zasobów w celu wykrycia okoliczności mogących świadczyć o popełnieniu przestępstwa o charakterze seksualnym na szkodę dziecka lub przygotowania do niego (§2258A f pkt 1–3 U.S. Code). W przypadku świadomego i zawinionego niezaraportowania faktów mogących świadczyć o opisanych wyżej działaniach przestępczych dostawca usług może zostać ukarany grzywną w wysokości do 150 tys. dolarów (§2258A e pkt 1 U.S. Code), która ulega podwojeniu w przypadku dalszych naruszeń (§2258A e pkt 2 U.S. Code).

Stany Zjednoczone, podobnie jak Unia Europejska, rozpoczęły prace nad zmianą prawa w zakresie wykrywania materiałów stanowiących wykorzystywanie seksualne dzieci. The Eliminating Abusive and Rampant Neglect of Interactive Technologies Act (EARN IT Act) to projekt, który po pierwsze zmienia nomenklaturę prawną, zastępując we wszystkich federalnych aktach prawnych sformułowanie „pornografia dziecięca" pojęciem „materiały stanowiące wykorzystywanie seksualne dzieci". Po drugie, projekt przewiduje utworzenie Krajowej Komisji ds. Prewencji Wykorzystywania Seksualnego Dzieci Online[16], której zadaniem będzie opracowywanie dobrych praktyk możliwych do wdrożenia przez dostawców interaktywnych usług komputerowych w celu przeciwdziałania i ograniczania skali zjawiska wykorzystywania seksualnego dzieci online (w tym groomingu i obrotu materiałami o charakterze pornograficznym z udziałem dzieci). Po trzecie – co wzbudza najwięcej kontrowersji – projekt zakłada zmianę §230 pkt e (tytuł 47) US Code. Dotąd dostawcy usług internetowych nie byli odpowiedzialni prawnie za treści zamieszczane na ich łamach przez niezależne od nich osoby (innych użytkowników), gdyż nie uważano ich za autorów lub wydawców tych treści. EARN IT Act ma zmienić omawiany artykuł poprzez dodanie części 6, zgodnie z którą możliwe stanie się wszczynanie wobec dostawców usług internetowych spraw cywilnych oraz karnych w oparciu o prawo stanowe dotyczące ogłaszania, promowania, prezentacji, rozpowszechniania lub produkowania materiałów stanowiących wykorzystywanie seksualne dzieci. Prawo stanowe może w przeciwieństwie do prawa federalnego (wymagającego wiedzy o tym, że materiały znajdują się na serwerach dostawcy) zastrzegać niższy stopień staranności do przyjęcia odpowiedzialności (np. niedbalstwo) i niższy standard dowodowy. Dodano też część 7, zgodnie z którą oferowanie szyfrowania end-to-end nie może służyć jako samodzielna podstawa współodpowiedzialności dostawcy usługi za rozpowszechnianie materiałów stanowiących wykorzystywanie seksualne

16 National Commission On Online Child Sexual Exploitation Prevention.

dzieci (A), jednakowoż sąd będzie uprawniony do oceny tego faktu w ewentualnej sprawie przeciwko niemu (B). Innymi słowy, dostawca usługi nie będzie odpowiedzialny za sam fakt oferowania szyfrowania end-to-end, ale to, że oferuje taki rodzaj szyfrowania, może zostać ocenione jako jego niedbalstwo. Gdyby bowiem dostawca nie zdecydował się na taki krok bądź oferował szyfrowanie, lecz z odpowiednimi wyjątkami, to wiedziałby, że na jego serwerach znajdują się materiały stanowiące niegodziwe traktowanie dzieci.

EARN IT Act ma wielu przeciwników. Zarzuca się mu przede wszystkim zbyt daleką ingerencję w prywatność użytkowników, a także uczynienie z dostawców usług internetowych wykonawców zadań organów ścigania. Akt do dnia dzisiejszego nie został przyjęty.

Przyszłość regulacji

W lipcu 2021 r. Komisja Europejska przygotowała projekt rozporządzenia, na mocy którego przewidziano tymczasową derogację stosowania przepisów dyrektywy o e-Prywatności wobec dostawców usług łączności elektronicznej niewykorzystującej numerów. Przepisy art. 5 ust. 1 i 6 dyrektywy o e-Prywatności są wobec nich wyłączone, o ile przetwarzanie danych (w tym danych osobowych) przez wskazane podmioty odbywa się wyłącznie w celu wykrywania niegodziwego traktowania dzieci, jest proporcjonalne, w tym jak najmniej ingeruje w inne prawa (szczególnie prywatność) i opiera się na sprawdzonych technologiach, które mają niski próg błędów. Ewentualne skanowanie w celu wykrycia przypadków groomingu ma miejsce wyłącznie w oparciu o słowa kluczowe i obiektywne zidentyfikowane czynniki ryzyka (np. różnica wieku) oraz może podlegać dalszej weryfikacji przez człowieka. Ponadto dostawcy winni publikować coroczne sprawozdania na temat wykrytych przypadków, a także jakości stosowanej technologii (w tym w zakresie wyników fałszywie dodatnich). Dane mogą być przetwarzane wyłącznie przez okres niezbędny do zablokowania użytkownika, zaraportowania organom ścigania faktu wykrycia niepożądanego zachowania po stronie użytkownika oraz do stworzenia unikalnego podpisu (hasza).

Obecna tymczasowa derogacja obowiązuje aktualnie do 3 kwietnia 2026 r., po (jednokrotnym dotąd) przedłużeniu.

W maju 2022 r. Komisja Europejska zaproponowała treść Rozporządzenia Parlamentu Europejskiego i Rady ustanawiającego przepisy mające na celu zapobieganie niegodziwemu traktowaniu dzieci w celach seksualnych i jego zwalczanie. Rozporządzenie to ma stanowić *lex specialis* w stosunku do Aktu o usługach cyfrowych.

Akt o usługach cyfrowych dotyczy usług hostingu (przechowywania informacji przekazanych przez odbiorcę usługi oraz na jego żądanie), cachingu (krótkotrwałego przechowywania informacji przekazanych przez użytkownika w celu usprawnienia jej późniejszej transmisji na żądanie innych odbiorców) i usług zwykłego przekazu (czyli transmisji informacji przekazywanych przez odbiorcę). Pozostawiono znaną z dotychczas obowiązującej dyrektywy zasadę, że dostawcy usług nie odpowiadają za treści zamieszczane przez użytkowników ani nie są zobowiązani do proaktywnego skanowania treści w poszukiwaniu nielegalnych materiałów. Warunkiem skorzystania z wyłączenia jest brak wiarygodnej informacji o nielegalnych treściach oraz podjęcie działań przez dostawcę usługi po powzięciu takiej informacji. Także dobrowolne inicjatywy dostawców usług w kierunku kontroli treści nie wykluczają możliwości skorzystania z wyłączeń odpowiedzialności. Podmioty świadczące tego typu usługi mogą zostać zobowiązane drogą orzeczenia sądu lub uprawnionego organu do usunięcia nielegalnych treści, uprzednio zidentyfikowanych przez organ wzywający do podjęcia działań wobec tych treści (wezwanie musi zawierać co najmniej jeden adres URL zidentyfikowanego materiału). Podmioty technologiczne zobowiązano też do ustanowienia łatwego dla użytkowników mechanizmu zgłaszania treści o charakterze nielegalnym, do których należą również (choć nie wyłącznie) materiały CSAM, i do reagowania na każde takie zgłoszenie.

Ponadto bardzo duże platformy internetowe (z liczbą użytkowników powyżej 45 mln miesięcznie z terenu UE) zobowiązane zostaną do prowadzenia corocznego audytu w zakresie ryzyk, jakie niesie korzystanie z oferowanych przez nich usług, a w szczególności – do badania, jakie zagrożenia dla realizacji praw dziecka wiąże się z korzystaniem z ich usług oraz w jakim stopniu ich usługi mogą być wykorzystywane do rozpowszechniania nielegalnych treści. W odpowiedzi na zidentyfikowane ryzyka konieczne stanie się wdrożenie środków zaradczych, odpowiednich do zidentyfikowanego ryzyka, np. dostosowanie systemu moderacji treści lub wprowadzenie skutecznych mechanizmów kontroli wieku użytkowników. Giganci będą musieli także corocznie poddać się na swój koszt niezależnemu audytowi badającemu wypełnianie przez nich obowiązków nałożonych Aktem o usługach cyfrowych.

Zgodnie z projektem rozporządzenia ws. zapobiegania niegodziwemu traktowaniu dzieci dostawcy usług hostingu oraz łączności interpersonalnej (a także sklepów z aplikacjami) zostali zobowiązani do przeprowadzania analizy ryzyka wykorzystania każdej z oferowanych przez nich usług w celu niegodziwego traktowania dzieci. Adekwatnie do stwierdzonych ryzyk dostawcy usług powinni wdrożyć środki zaradcze. Do czynników ryzyka należy korzystanie przez dzieci z usługi, możliwość wyszukiwania innych użytkowników i nawiązanie z nimi bezpośrednich kontaktów

czy też możliwość udostępniania zdjęć i nagrań. W projekcie nie określono pożądanych sposobów reakcji na zidentyfikowane ryzyka, pozostawiając, co do zasady, uregulowania ewentualnym wytycznym Komisji Europejskiej. Środki zaradcze mają być w szczególności adekwatne i proporcjonalne. Jedynym sposobem reakcji wymaganym wprost w projekcie jest podjęcie przez dostawców działań zmierzających do identyfikacji użytkowników – dzieci.

W przypadku, gdy istnieją dowody na znaczne ryzyko wykorzystywania usługi oferowanej przez dostawcę do niegodziwego traktowania dzieci w celach seksualnych, organ koordynujący kraju prowadzenia działalności przez dostawcę usług może wystąpić do właściwego sądu lub organu administracji o wydanie wobec dostawcy nakazu wykrywania. Przesłanki wydania nakazu muszą przeważać nad negatywnymi konsekwencjami dla praw podstawowych i praw innych podmiotów zainteresowanych (zatem i dostawcy). Nakaz wykrywania wydaje się, gdy pomimo wdrożenia przez dostawcę środków zaradczych usługa jest wykorzystywana do rozpowszechniania znanych materiałów stanowiących wykorzystywanie seksualne dzieci. Można go także zastosować w celu rozpoznawania nagabywania w celach seksualnych. Wykrywanie może mieć miejsce wyłącznie w komunikacji interpersonalnej i tylko wtedy, gdy jedną ze stron kontaktu jest dziecko. Nakaz wykrywania ma być orzekany na czas określony: w przypadku materiałów stanowiących wykorzystywanie seksualne dzieci – do 24 miesięcy, a w przypadku wykrywania groomingu – do 12 miesięcy. Nakaz wykrywania ma także zawierać wskazanie konkretnych środków, jakie należy podjąć, by go wdrożyć. Od dostawcy nie będzie się jednak wymagało zastosowania żadnej konkretnej technologii, gdyż otrzyma on możliwość skorzystania z nieodpłatnego oprogramowania dostarczanego przez Unijne Centrum bądź wdrożenia własnych rozwiązań. Powinny być one skuteczne, aktualne i wiarygodne oraz nie pozwalać na pozyskanie informacji innych niż to konieczne dla realizacji nakazu. Dostawcy usług zostaną zobowiązani do wdrożenia skutecznego mechanizmu raportowania przypadków niegodziwego traktowania dzieci w celach seksualnych.

Ponadto organy koordynujące zyskają uprawnienie do występowania o wydanie nakazu usunięcia zidentyfikowanego materiału stanowiącego wykorzystywanie seksualne dzieci lub zablokowania dostępu do niego (tzw. nakaz blokowania). Nakaz usunięcia będzie musiał być wykonany w ciągu 24 godzin. Z kolei blokada dostępu zostanie ograniczona czasowo – maksymalnie do pięciu lat od daty wydania decyzji. Działania te obejmą konkretne adresy URL, wskazane w bazie adresów prowadzonej przez Unijne Centrum. Instytucja ta ma także zarządzać biblioteką wskaźników, tworzonych na podstawie transkrypcji rozmów i materiałów dokumentujących przypadki niegodziwego traktowania dzieci w celach seksualnych.

Organy koordynujące uzyskają kompetencje do kontroli przestrzegania przepisów rozporządzenia, w tym będą mogły żądać udostępnienia wszelkich informacji dotyczących mechanizmów wyszukiwania treści w usługach hostingu celem zweryfikowania potrzeby wydania jednego z nakazów bądź skuteczności nakazu już wydanego. Przewidziano kary finansowe za nieprzestrzeganie przepisów rozporządzenia w wysokości procenta od dochodu przedsiębiorstwa ukaranego.

Ponadto w projekcie uwzględniono rozbudowany system skargowy w zakresie realizacji nakazów lub innych obowiązków po stronie dostawców. Prawo do zaskarżenia przyznano także użytkownikom dotkniętym skutkami realizacji nakazów. Ze wszystkimi nakazami wiąże się obowiązek sprawozdawczości z ich wykonania po stronie dostawcy.

Jest więcej niż pewne, że serwisy społecznościowe zostaną uznane za kwalifikujące się do nakazu wykrywania – z uwagi na skalę zarówno aktywnych użytkowników, jak i dotąd raportowanych naruszeń. Projektowane przepisy wyznaczają zasadę wydania nakazu wyłącznie na czas oznaczony, niemniej w przepisach nie jest wykluczone wydawanie nieskończonej liczby nakazów po kolei, co zapewne też nastąpi, chyba że technologia innych zabezpieczeń zostanie uznana za wystarczającą.

Choć rozporządzenie znajduje się na etapie kolejnej rundy konsultacji w krajach członkowskich, to jasne jest, że niezależnie od jego ostatecznej treści główne założenie, jakie przynosi ze sobą projektowana zmiana – czyli możliwość skanowania przez dostawców usług internetowych treści zamieszczanych w serwisach społecznościowych, jak i wymienianych w prywatnych wiadomościach – stoi w pewnym stopniu w konflikcie z prawem użytkowników do prywatności. Jednocześnie wśród obywateli krajów Unii Europejskiej istnieje istotne poparcie dla projektowanej regulacji (ECPAT, 2021), nawet kosztem ograniczenia tego prawa.

O ile prawo do prywatności, w tym prywatności korespondencji, nigdy nie miało charakteru bezwzględnego i mogło podlegać ograniczeniom, to przedmiotowa regulacja rodzi kilka innych, istotnych wątpliwości. Projekt rozporządzenia zakłada, że dostawcy nie będą mogli pozyskiwać innych danych niż bezwzględnie konieczne do wykrywania przypadków wykorzystywania seksualnego dzieci. Skanowanie treści niewątpliwie stanie się stałą praktyką w przypadku dostawców szeregu usług wykorzystywanych każdego dnia przez użytkowników z terenu Unii Europejskiej. Istnieje uzasadniona obawa, że relacja firm technologicznych nie tylko z użytkownikami, ale także z organami regulacyjnymi w obliczu uprawnienia czy też nakazu skanowania treści będzie musiała w dużej mierze opierać się na zaufaniu. Choć organy regulacyjne zostaną wyposażone w kompetencje do weryfikacji przestrzegania omawianych przepisów – w tym do żądania ujawnienia wszelkich istotnych informacji

niezbędnych do oceny ich realizacji, niezależnie od formy nośnika – to transparentność oraz rzetelność działań po stronie dostawców usług pozostaną kluczowe. Jednocześnie, dla realizacji celów rozporządzenia, podmioty te będą uprawnione do korzystania z oprogramowania dostarczonego przez Unijne Centrum lub z każdego innego równoważnego rozwiązania technologicznego. Aby zagwarantować, że firmy technologiczne nie nadużywają przyznanych im uprawnień, organy kontrolujące oraz audytorzy musieliby mieć zarówno możliwość, jak i zdolność do kontroli algorytmu stosowanego do skanowania treści (także wtedy, gdy stosowane jest oprogramowanie dostarczone przez Unijne Centrum, ale z możliwością jego modyfikacji). Istnieje obawa, że dostawcy usług blokować zaczną dostęp do algorytmu, powołując się chociażby na zasady prawa autorskiego lub tajemnicę przedsiębiorstwa. Wydaje się, że aktualna treść projektu nie odpowiada na te wyzwania. Nawet jeżeli dostawcy usług zdecydują się przekazać dostęp do algorytmu stosowanego do wyszukiwania treści stanowiących wykorzystywanie seksualne dzieci, organy kontrolujące powinny mieć zdolność do testowania algorytmu w warunkach laboratoryjnych. W związku z faktem, że kontrola przestrzegania rozporządzenia została przypisana organom koordynującym z poszczególnych państw członkowskich, istnieje ryzyko wystąpienia niejednorodnych praktyk i niejednakowych zasobów, które spowodują niezdolność organu do przeprowadzenia efektywnej kontroli, czy dostawca nie zbiera innych danych, niż rzeczywiście konieczne do wykonania rozporządzenia. Wątpliwość budzi też zawodność technologii, o czym mowa była wyżej, niemniej tę okoliczność można wyeliminować na drodze postępu technologicznego. Podobna materia została na gruncie DSA doprecyzowana w dużo bardziej zadowalający sposób. Wprost przewidziano w tym akcie prawnym obowiązek podania na żądanie uprawnionych organów danych umożliwiających (także technologicznie) ocenę zgodności z prawem rozwiązań używanych przez dostawców usług internetowych.

Prawa dzieci w internecie wciąż pozostają zagadnieniem złożonym, ale też nieokreślonym i nie pierwszoplanowym, również ze strony ustawodawcy unijnego. Wskazuje na to nie tylko spóźniona i powolna reakcja na zagrożenia wynikające z wejścia Europejskiego Kodeksu Łączności w życie, jak i przedłużająca się dyskusja nad wdrożeniem nowych regulacji, ale także niezdolność lokalnych ustawodawstw do wprowadzenia skutecznych mechanizmów w zakresie mniej złożonych zagadnień, takich jak np. kontroli wieku osób uzyskujących dostęp do stron o charakterze pornograficznym lub zakupu alkoholu online. Materia ta, choć dużo prostsza, pozostaje w wyłącznej gestii dostawców tego typu usług. Wydaje się, że gdyby nie lobbing ze strony podmiotów zlokalizowanych głównie w Stanach Zjednoczonych w reakcji na wprowadzenie EKŁE, materia skanowania treści także pozostałaby w wyłącznej

gestii dostawców usług (Senat Stanów Zjednoczonych, 2020). Polska w ramach półrocznej prezydencji w Radzie Unii Europejskiej zaproponowała własną propozycję treści rozporządzenia CSAM, ale nie spotkał on się z aprobatą państw członkowskich. W polskim projekcie zrezygnowano z nałożenia na dostawców usług internetowych obowiązku wykrywania treści stanowiących wykorzystanie seksualne dzieci, pozostając przy dobrowolności stosowania tego typu mechanizmów, co nie zmieniłoby wiele w stopniu ochrony dzieci przed krzywdą online.

Zagadnienie konfliktu praw człowieka (głównie dzieci, ale też zapewne osób już dorosłych, sportretowanych w materiałach o charakterze pornograficznym) z prawami użytkowników internetu (głównie prawem do prywatności) pozostaje kontrowersyjne. Wprowadzenie rozporządzenia stanowi jednak nieuchronny krok, którego zaistnienie stanowi ewolucję praw człowieka i postrzegania odpowiedzialności za ich przestrzeganie. Równolegle trwają prace nad wprowadzeniem wymogu należytej staranności biznesu w obszarze praw człowieka. Dochodzi do powolnego rozszerzenia odpowiedzialności za ich przestrzeganie także na duże podmioty prywatne, które niejednokrotnie mają większe możliwości finansowe i organizacyjne niż państwa.

E-mail autorki: **katarzyna.katana@fdds.pl**.

Bibliografia

Apple Inc. (2021). *CSAM Detection. Technical Summary*. https://www.apple.com/child-safety/pdf/CSAM_Detection_Technical_Summary.pdf

Barr, W.P. (2020). *Attorney General Barr Signs Letter to Facebook From US, UK, and Australian Leaders Regarding Use of End-To-End Encryption*. https://www.justice.gov/opa/pr/attorney-general-barr-signs-letter-facebook-us-uk-and-australian-leaders-regarding-use-end

Bowles, N., Keller, M.H. (2019). *Video games and online chats are 'hunting grounds' for sexual predators*. *New York Times*. https://www.seattletimes.com/nation-world/video-games-and-online-chats-are-hunting-grounds-for-sexual-predators/

Cholewa, M., Studnicki, P. (red.). (2021). *Wykorzystywanie seksualne osób małoletnich. Ujęcie interdyscyplinarne*. Wydawnictwo Scriptum.

Clark, J.F. (2020). *List prezesa National Center for Missing & Exploited Children*. https://www.missingkids.org/content/dam/missingkids/pdfs/NCMEC%20letter%20to%20EU%20Parliament%20Members.pdf

ECPAT International (2021). *Raport z badań społecznych*. https://ecpat.org/wp-content/uploads/2021/11/Summary-Report-Polling-Research-16-November-21.pdf

ECPAT (2025). *Terminology guidelines*. https://ecpat.org/wp-content/uploads/2025/04/Second-Edition-Terminology-Guidelines-final.pdf

Farid H. (2021). *An overview of perceptual hashing*. *Journal of Online Trust and Safety*.

Gregoire, C. (2020). *Microsoft shares new technique to address online grooming of children for sexual purposes*. https://blogs.microsoft.com/on-the-issues/2020/01/09/artemis-online-grooming-detection/

INHOPE Association (2024). *Annual report 2023*.

Interagency Working Group (2016). *Wytyczne Luksemburskie*. https://ecpat.org/luxembourg-guidelines/

Komisja Europejska (2022). *Impact assessment report accompanying the document Proposal for a Regulation Of The European Parliament And Of The COUNCIL laying down rules to prevent and combat child sexual abuse*. Bruksela.

Komisja Europejska. *Strategia UE na rzecz praw dziecka*.

Komitet Praw Dziecka (2021). *Komentarz Generalny nr 25 w przedmiocie praw dzieci w środowisku online*.

Krawetz., N. (2021). *PhotoDNA and Limitations*. https://www.hackerfactor.com/blog/index.php?archives/931-PhotoDNA-and-Limitations.html

Microsoft (2025). *PhotoDNA FAQ*. https://www.microsoft.com/en-us/PhotoDNA/FAQ

NASK Państwowy Instytut Badawczy, Zespół Dyżurnet.pl (2024). *Raport za rok 2023*.

National Center of Missing and Exploited Children (2020). *We are in danger of losing the global battle for child safety*. https://www.missingkids.org/blog/2020/we-are-in-danger-of-losing-the-global-battle-for-child-safety

Senat Stanów Zjednoczonych (2020, 7 grudnia). *Rezolucja nr 794*.

Specjalny Sprawozdawca ONZ (1992). *Raport w zakresie handlu dziećmi i wykorzystywania seksualnego dzieci*.

Specjalny Sprawozdawca ONZ (2004). *Raport w zakresie handlu dziećmi i wykorzystywania seksualnego dzieci*.

Specjalny Sprawozdawca ONZ (2009). *Raport w zakresie handlu dziećmi i wykorzystywania seksualnego dzieci*.

Specjalny Sprawozdawca ONZ (2014). *Raport w zakresie handlu dziećmi i wykorzystywania seksualnego dzieci*.

Zamfir I. (2025). *Proposal for a revision of the combating child sexual abuse Directive (2011/93/EU)*. European Parlament Members' Research Service. https://www.europarl.europa.eu/legislative-train/theme-a-new-era-for-european-defence-and-security/file-revision-of-the-combating-child-sexual-abuse-directive

Akty prawne

Dyrektywa 2000/31/We Parlamentu Europejskiego i Rady z dnia 8 czerwca 2000 r. w sprawie niektórych aspektów prawnych usług społeczeństwa informacyjnego, w szczególności handlu elektronicznego w ramach rynku wewnętrznego (dyrektywa o handlu elektronicznym).

Dyrektywa 2002/58/WE Parlamentu Europejskiego i Rady z dnia 12 lipca 2002 r. dotycząca przetwarzania danych osobowych i ochrony prywatności w sektorze łączności elektronicznej (dyrektywa o prywatności i łączności elektronicznej) (Dz.U. UE. L. z 2002 r., nr 201, s. 37 z późn. zm.).

Dyrektywa Parlamentu Europejskiego i Rady (UE) 2018/1972 z dnia 11 grudnia 2018 r. ustanawiająca Europejski kodeks łączności elektronicznej (wersja przekształcona). Tekst mający znaczenie dla EOG (Dz.U. UE. L. z 2018 r., nr 321, s. 36 z późn. zm.).

Konwencja o prawach dziecka przyjęta przez Zgromadzenie Ogólne Narodów Zjednoczonych dnia 20 listopada 1989 r. (Dz.U. z 1991 r., nr 120, poz. 526 z późn. zm.).

Konwencja Rady Europy o cyberprzestępczości, sporządzona w Budapeszcie dnia 23 listopada 2001 r. (Dz.U. z 2015 r., poz. 728).

Konwencja Rady Europy o ochronie dzieci przed seksualnym wykorzystywaniem i niegodziwym traktowaniem w celach seksualnych, sporządzona w Lanzarote dnia 25 października 2007 r. (Dz.U. z 2015 r., poz. 608).

Protokół fakultatywny do Konwencji o prawach dziecka w sprawie handlu dziećmi, dziecięcej prostytucji i dziecięcej pornografii, przyjęty w Nowym Jorku dnia 25 maja 2000 r. (Dz.U. z 2007 r., nr 76, poz. 494).

Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2016/679 z dnia 27 kwietnia 2016 r. w sprawie ochrony osób fizycznych w związku z przetwarzaniem danych osobowych i w sprawie swobodnego przepływu takich danych oraz uchylenia dyrektywy 95/46/WE (ogólne rozporządzenie o ochronie danych) (Dz.U. UE. L. z 2016 r., nr 119, s. 1 z późn. zm.).

Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2021/1232 z dnia 14 lipca 2021 r. w sprawie tymczasowego odstępstwa od niektórych przepisów dyrektywy 2002/58/WE w odniesieniu do wykorzystywania technologii przez dostawców usług łączności interpersonalnej niewykorzystujących numerów do przetwarzania danych osobowych i innych danych do celów zwalczania niegodziwego traktowania dzieci w celach seksualnych w internecie (Dz.U. UE. L. z 2021 r., nr 274, s. 41 z późn. zm.).

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. z 2022 nr 277, s. 1).

The Code of Laws of the United States of America.

UK Online Safety Act https://www.legislation.gov.uk/ukpga/2023/50

Ustawa z dnia 5 lipca 2018 r. o krajowym systemie cyberbezpieczeństwa (t.j. Dz.U. z 2020 r., poz. 1369 z późn. zm.).

Ustawa z dnia 6 czerwca 1997 r. Kodeks karny (t.j. Dz.U. z 2022 r., poz. 1138). 9 z późn. zm.).

Ustawa z dnia 6 czerwca 1997 r. Kodeks karny (t.j. Dz.U. z 2022 r., poz. 1138).

Projekty aktów prawnych

Rozporządzenie Parlamentu Europejskiego i Rady ustanawiające przepisy mające na celu zapobieganie niegodziwemu traktowaniu dzieci w celach seksualnych i jego zwalczanie.

The Eliminating Abusive and Rampant Neglect of Interactive Technologies Act (EARN IT Act).

Challenges of legal protection of children in the context of online sexual violence

The aim of this article is to discuss the challenges related to the legal protection of children from online sexual cyberviolence. Studies show that the scale of online violence against children, particularly in the aspect of sexual violence, increases every year. The law provides multiple responses to such situations, though it does not always keep up with the discussed problem. Preventive in nature are certainly child safeguarding practices, which since February 2024 shall be implemented in every institution working with children. These standards should include rules for internet usage within the facility, teaching children safe behaviors online, as well as rules for safe relationships between staff and children, and safe peer-to-peer relationships. These would involve sets of desirable and prohibited behaviors regarding children.

But what if prevention does not work, and intervention must be taken? In the case of cyberviolence, the law comes into play, understood traditionally (this includes, for example, the legal pathway for a family court's response to a case of delinquency or a criminal act committed by a person under the age of 18 and 17, or a criminal court's response in cases involving materials that constitute child sexual exploitation) and the law of new technologies, which requires service providers to react to harmful content by moderating it based on reports, providing complaint mechanisms, and ensuring the welfare of minors. Currently, however, internet service providers are not required to proactively check whether their services are being used to harm children.

The EU Digital Services Act (DSA), i.e., Regulation (EU) 2022/2065 of the European Parliament and the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC, came into force in all member states from 17 February 2024, even though the Polish legislator has not yet adjusted Polish law in this area. The European Parliament is working on the amendments to the Directive of the European Parliament and the Council 2011/92/EU of 13 December 2011 on combating the sexual abuse and exploitation of children and child pornography, replacing Framework Decision 2004/68/JHA, to include in its provisions crimes committed against children using the internet. Ongoing (unfinished) work is also being done on a Regulation of the European Parliament and the Council establishing provisions to prevent and combat the sexual abuse of children. The article aims to discuss the provisions of these three – existing and proposed – acts in the context of responding to sexual violence against children committed using new technologies. The author will consider whether the current regulation protecting children from online abuse is complete or whether it will be after the adoption of the

proposed changes. The author will also highlight the potential difficulties in responding to such acts and the role of internet service providers in this semi-private – semi-public responsibility for ensuring children's safety online.

Keywords:
cyberviolence, sexual violence, child protection, internet

Cytowanie:
Katana, K. (2025). Wyzwania prawnej ochrony dzieci w aspekcie przemocy seksualnej online. *Dziecko Krzywdzone. Teoria, badania, praktyka*, 24(2), 124–149.

Fundusz Sprawiedliwości
Ministerstwo Sprawiedliwości
Sfinansowano ze środków Funduszu Sprawiedliwości, którego dysponentem jest Minister Sprawiedliwości