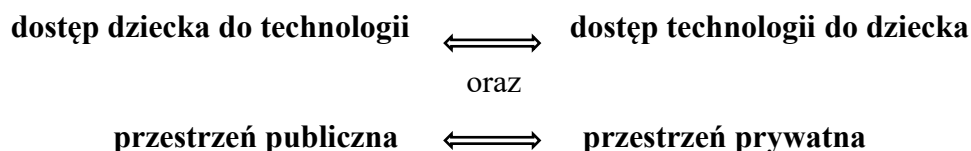


**Stanowisko Komisji Kodyfikacyjnej Prawa Rodzinnego
w sprawie cyfrowej reprezentacji dziecka**

1. Uwagi wstępne

Na ochronę cyfrowej reprezentacji dziecka składają się: skuteczne egzekwowanie praw (por. *stanowisko Komisji Kodyfikacyjnej Prawa Rodzinnego w sprawie potrzeby podjęcia działań zmierzających do wzmocnienia ochrony małoletnich w cyberprzestrzeni*) i polityk usługodawców uwzględniających zgody przedstawicieli ustawowych dziecka (i samego dziecka), kontrola śladu cyfrowego oraz reagowanie na wszelkie formy nadużyć i przemocy wobec małoletniego. W tym celu należy stworzyć bezpieczną (technologicznie i prawnie) cyberprzestrzeń, w której w równym stopniu uregulowane są dwa obszary:



Z tego też powodu należy dążyć do wprowadzenia trzech wzajemnie uzupełniających się poziomów zabezpieczeń systemowych. Niezwykle istotne jest, aby jasno określić uczestników oraz zakres ich odpowiedzialności dla każdego z tych poziomów.

- 1) **poziom pierwszy** – bank zgód (bezpieczna i spersonalizowana usługa do zarządzania i przechowywania zgód na użytkowanie technologii wymagającej/umożliwiającej umieszczenie w cyberprzestrzeni reprezentacji cyfrowej dziecka);
- 2) **poziom drugi** – bank reprezentacji cyfrowych (bezpieczna i spersonalizowana usługa do zarządzania i przechowywania reprezentacji cyfrowych, ułatwiająca prewencję technologiczną w fazie *pre-upload* oraz *post-upload*);
- 3) **poziom trzeci** – system reagowania (środki technologiczne zapewniające reakcję na nadużycia).

2. Słowniczek pojęć

1. w zakresie technologii:

- 1) **deepfake** – obraz, nagranie dźwiękowe lub ich połączenie, które zostały w całości lub częściowo wygenerowane albo zmodyfikowane przy użyciu sztucznej inteligencji w sposób sprawiający wrażenie przedstawienia rzeczywistych osób lub zdarzeń;

- 2) **friction by design** – projektowanie usług lub funkcjonalności w sposób wprowadzający celowe i proporcjonalne mechanizmy spowalniające, ostrzegające lub wymagające dodatkowego potwierdzenia przed wykonaniem czynności, w celu ograniczenia ryzyka naruszenia praw lub nadużyć;
- 3) **hasz** (*hash*) – wynik działania kryptograficznej funkcji skrótu (ang. *cryptographic hash function*) zastosowanej do reprezentacji cyfrowej, albo jej wyodrębnionych cech, umożliwiający wykrywanie identycznych albo podobnych treści bez konieczności przechowywania ich pełnej postaci;
- 4) **materiał syntetyczny wygenerowany przez sztuczną inteligencję** – obraz, nagranie dźwiękowe lub ich połączenie, a także tekst lub inna treść cyfrowa w całości lub częściowo wygenerowane albo zmodyfikowane przy użyciu sztucznej inteligencji, niezależnie od stopnia jej podobieństwa do rzeczywistych osób lub zdarzeń;
- 5) **metadane** – dane opisujące cechy materiału cyfrowego, w szczególności informacje dotyczące czasu utworzenia, urządzenia, lokalizacji, parametrów technicznych oraz historii przetwarzania, w określonych przypadkach również dane twórcy;
- 6) **reprezentacja audiowizualna** – reprezentacja cyfrowa obejmująca obraz, nagranie dźwiękowe albo ich połączenie;
- 7) **reprezentacja cyfrowa** – utwalony lub wygenerowany przy użyciu technologii cyfrowej zbiór danych odzwierciedlających cechy fizyczne, biometryczne, głosowe, behawioralne lub inne cechy umożliwiające identyfikację osób; za reprezentację cyfrową można uznać np. przetworzony lub wytworzony wizerunek dziecka, awatar, zindywidualizowane logo;
- 8) **retencja** – przechowywanie danych przez czas określony przepisami prawa powszechnie obowiązującego;
- 9) **ślad cyfrowy** – ogół danych i informacji pozostawianych przez użytkownika w związku z korzystaniem z technologii cyfrowych; obejmuje zarówno dane udostępniane świadomie, jak i te powstające automatycznie podczas korzystania z usług.

2. w zakresie usług:

- 1) **asystent egzekwowania praw dziecka** – funkcjonalność usługi (np. mObywatel) umożliwiająca dziecku lub jego przedstawicielowi ustawowemu wykonywanie praw związanych z reprezentacją cyfrową, w szczególności zgłaszanie naruszeń, żądanie usunięcia treści oraz monitorowanie sposobu wykorzystania reprezentacji cyfrowej;
- 2) **automatyczne monitorowanie Internetu** – proces ciągłego lub okresowego wyszukiwania publicznie dostępnych materiałów cyfrowych, w tym przypadku w celu wykrycia nieuprawnionego wykorzystania reprezentacji cyfrowej;

- 3) **cyfrowy paszport** – zbiór danych dostępnych za pomocą mechanizmów identyfikacji cyfrowej, zawierający informacje niezbędne do potwierdzenia autentyczności (reprezentacji cyfrowej), zarządzania zgodami i uprawnieniami oraz wykrywania nieuprawnionego wykorzystania;
- 4) **kontrola *post-upload/pre-upload*** – zautomatyzowana analiza materiału cyfrowego wykonywana przed albo po jego opublikowaniu lub udostępnieniu;
- 5) **mechanizm potwierdzenia tożsamości dziecka i jego przedstawiciela ustawowego** – zespół środków organizacyjnych i technicznych umożliwiających potwierdzenie tożsamości dziecka oraz jego przedstawiciela ustawowego, a także istnienia między nimi relacji uprawniającej do wykonywania praw i obowiązków określonych w ustawie;
- 6) **potwierdzenie wieku** (ang. *age attestation*) – kryptograficzne (token wieku) lub urzędowe poświadczenie wieku wykonywane np. przez zaufany podmiot wyłącznie w celu zastosowania odpowiednich środków ochrony dziecka;
- 7) **weryfikacja na podstawie podpisu elektronicznego** – potwierdzenie tożsamości osoby oraz jej oświadczenia woli przy użyciu podpisu elektronicznego lub innego środka kryptograficznego zapewniającego integralność danych, identyfikację osoby składającej oświadczenie oraz możliwość weryfikacji dokonanej czynności [np. podpisy oparte na portfelach tożsamości cyfrowej (EUDI Wallet/eIDAS)];
- 8) **wykrywanie wieku** (ang. *age estimation*) – szacowanie wieku na podstawie cech biometrycznych (np. AI analizująca twarz), wykonywane wyłącznie w celu zastosowania odpowiednich środków ochrony dziecka;
- 9) **zgoda cyfrowa** – oświadczenie woli, której autentyczność i integralność mogą zostać zweryfikowane, wyrażone przy użyciu środków komunikacji elektronicznej, potwierdzające zgodę albo odmowę zgody na publikowanie, przetwarzanie, wykorzystywanie lub usunięcie reprezentacji cyfrowej dziecka.

3. w zakresie systemu ochrony dzieci:

- 1) **INHOPE CPORT** – system wymiany informacji prowadzony przez międzynarodową organizację INHOPE, umożliwiający bezpieczne przekazywanie zgłoszeń dotyczących materiałów przedstawiających seksualne wykorzystywanie dzieci oraz koordynację działań pomiędzy krajowymi punktami zgłoszeniowymi a organami ścigania;
- 2) **INTERPOL ICSE DB** (ang. *International Child Sexual Exploitation Database*) – międzynarodowa baza danych zarządzana przez Interpol, służąca identyfikowaniu małoletnich pokrzywdzonych przestępstwami o charakterze seksualnym, analizie materiałów przedstawiających seksualne wykorzystywanie dzieci oraz wspieraniu współpracy organów ścigania zrzeszonych państw;

- 3) **NCMEC Take It Down** – system prowadzony przez NCMEC (ang. *National Center for Missing & Exploited Children*), umożliwiający zgłaszanie oraz ograniczanie rozpowszechniania w Internecie intymnych materiałów przedstawiających osoby małoletnie przez wykorzystanie technologii skrótów kryptograficznych (*hashy*) oraz współpracę z dostawcami usług internetowych;
- 4) **pilny wniosek** (ang. *urgent request*) – żądanie niezwłocznego zabezpieczenia lub udostępnienia danych kierowane przez uprawniony organ lub inny podmiot w celu zapobieżenia bezpośredniemu zagrożeniu życia lub zdrowia;
- 5) **punkt zgłoszeń Dyżurnet.pl** – działający w ramach CSIRT NASK krajowy punkt zgłoszeń, zapewniający całodobowe przyjmowanie zgłoszeń dotyczących nielegalnych treści w Internecie (formularz na stronie internetowej oraz mObywatel), w szczególności materiałów przedstawiających seksualne wykorzystywanie dzieci, oraz ich analizę i przekazywanie do właściwych podmiotów;
- 6) **prawo usługodawcy do korzystania z reprezentacji cyfrowej** (*reverse subscription*) – czasowe upoważnienie do przechowywania, przetwarzania lub udostępniania reprezentacji cyfrowej udzielane usługodawcy przez dziecko oraz jego przedstawiciela ustawowego; uprawnienie przyznawane jest w określonym celu i wygasa wraz z cofnięciem zgody lub z upływem okresu, na jaki zostało udzielone, chyba że zostanie odnowione zgodnie z wolą uprawnionych osób.

3. Zagadnienia szczegółowe

Na każdym z poziomów zabezpieczeń systemowych powinno się respektować podmiotowość dziecka jako osoby w fazie dynamicznego rozwoju fizycznego, psychicznego, emocjonalnego, społecznego i poznawczego. Określenie stopnia samodzielności dziecka w podejmowaniu decyzji dotyczących jego reprezentacji cyfrowej powinno uwzględniać zróżnicowane tempo osiągania dojrzałości, przy czym można wyróżnić przykładowe¹ przedziały wiekowe:

- a) do 13 r.ż. – pełne stanowienie przez przedstawiciela ustawowego dziecka o jego reprezentacji cyfrowej i aktywności online;
- b) od 13 r.ż. do 16/17 r.ż. – współstanowienie (decyzja dziecka potwierdzona zgodą jego przedstawiciela ustawowego);
- c) od 16/17 r.ż. do 18 r.ż. – asystowanie (powiadomienie przedstawiciela ustawowego dziecka o jego aktywności, również w kontekście informowania o zagrożeniach i nadużyciach).

¹ Ustalenie przedziałów wiekowych wymaga pogłębionej analizy.

1) Poziom pierwszy

Bank zgód to bezpieczna i spersonalizowana usługa do zarządzania i przechowywania relevantnych dokumentów określających obowiązki i prawa usługodawcy, a także przedstawiciela ustawowego dziecka oraz małoletniego w kontekście zgody na publikację i przetwarzanie jego reprezentacji cyfrowej.

Usługodawca w usłudze dla pierwszego przedziału wiekowego ma obowiązek:

- a) zapewnić przedstawicielowi ustawowemu dziecka możliwość wyrażenia albo niewyrażenia zgody;
- b) zapewnić mechanizm potwierdzenia tożsamości dziecka oraz jego przedstawiciela ustawowego;
- c) przechowywać dane dotyczące zgody przez okres wymagany prawem państwa (np. na potrzeby ścigania przestępstw).

Usługodawca w usłudze dla drugiego przedziału wiekowego musi:

- a) zapewnić dziecku oraz jego przedstawicielowi ustawowemu możliwość wyrażenia albo niewyrażenia zgody;
- b) zapewnić mechanizm potwierdzenia tożsamości dziecka oraz jego przedstawiciela ustawowego;
- c) przechowywać dane dotyczące zgody przez okres wymagany prawem państwa (np. na potrzeby ścigania przestępstw).

Usługodawca w usłudze dla trzeciego przedziału wiekowego musi:

- a) zapewnić dziecku możliwość wyrażenia albo niewyrażenia zgody oraz przesłanie przedstawicielowi ustawowemu powiadomienia;
- b) zapewnić mechanizm potwierdzenia tożsamości dziecka oraz jego przedstawiciela ustawowego;
- c) przechowywać dane dotyczące zgody przez okres wymagany prawem państwa (np. na potrzeby ścigania przestępstw).

Mechanizm wyrażenia zgody powinien obejmować: potwierdzenie relacji pokrewieństwa, np. „Potwierdzam, że jestem przedstawicielem ustawowym dziecka”; potwierdzenie dbałości o dobro dziecka, np. „Potwierdzam, że użytkowanie aplikacji/publikacja nie narusza dobra dziecka”, a także potwierdzenie świadomej zgody, np. „Rozumiem ryzyka związane z udostępnianiem wizerunku małoletniego”. Dodatkowo warto dodać element dynamiczny wzmacniający świadomy wybór (tzw. *friction by design*)² taki jak pytanie: „Zdjęcie zawiera wizerunek dziecka. Czy na pewno chcesz je upublicznić?”, obowiązkowe kliknięcie: „Tak, rozumiem ryzyka”.

² <https://designingfriction.com/>

PRZYKŁAD: usługa, z której korzysta użytkownik (np. mObywatel, zakładka „bank zgód”), przechowuje zgodę wyrażoną w związku z przystąpieniem do korzystania z aplikacji A. Usługodawca oferujący aplikację A, po zweryfikowaniu tożsamości użytkownika (dziecka lub jego przedstawiciela ustawowego), jego praw i obowiązków, przekazuje do „banku zgód” informację o zgodzie.

2) Poziom drugi

Bank reprezentacji cyfrowych to bezpieczna i spersonalizowana usługa do zarządzania i przechowywania reprezentacji cyfrowych, ułatwiająca prewencję technologiczną w fazie *pre-upload* oraz *post-upload*.

Reprezentacja cyfrowa dostępna w przestrzeni publicznej musi obejmować następujące rozwiązania:

1) zarządzany przez przedstawiciela ustawowego dziecka bank reprezentacji cyfrowych lub hashy wizerunku, a także głosu dziecka:

a) istotny z punktu widzenia ustalania autentyczności treści audiowizualnych w związku z podejrzeniem nadużyć i przestępstw (przedstawiciela ustawowego dziecka, AI, osoby trzeciej);

b) niezbędne jest określenie dla tego rozwiązania technologicznego ryzyk związanych np. z ewentualnymi cyberatakami, czy wyciekami danych;

2) automatyczne wykrywanie wizerunku oraz głosu dziecka – w zależności od wieku i kontekstu: pełna blokada *uploade'u*; *upload* ograniczony czasowo z możliwością przedłużenia (np. powiązanie reprezentacji cyfrowej dziecka z jego kontem społecznościowym odnawiane co 30 dni – zastosowanie wobec platform prawa do korzystania z cyfrowej reprezentacji; *upload* na stałe po uzyskaniu zgód – pkt 1);

3) bezpieczeństwo reprezentacji cyfrowej dziecka przez stosowanie takich rozwiązań jak:

a) blurowanie, zastąpienie awatarem, modulacja głosu;

b) blokada pobierania lub powiadomienie o próbie pobrania;

c) pełna dostępność uzyskiwana tylko w wyniku autoryzacji podpisem cyfrowym;

4) automatyczne usuwanie metadanych – obowiązek retencji usuniętych metadanych na potrzeby tzw. *urgent request*;

5) dostęp technologii do dziecka powinien obejmować zakazy:

a) profilowania dziecka w celach marketingowych;

b) podejmowania zautomatyzowanych decyzji wobec dziecka;

c) uzależniania od zachowań patologicznych, szkodliwych (np. przemoc, hazard, pornografia).

Reprezentacja cyfrowa dostępna w przestrzeni prywatnej³ musi obejmować następujące rozwiązania:

- 1) zarządzany przez dziecko bank reprezentacji cyfrowych lub hashy wizerunku, a także głosu dziecka:
 - a) istotny z punktu widzenia ustalania autentyczności materiału audiowizualnego w związku z podejrzeniem nadużyć i przestępstw (przedstawiciela ustawowego dziecka, AI, osoby trzeciej);
 - b) niezbędne jest określenie dla tego rozwiązania technologicznego ryzyk związanych np. z ewentualnymi cyberatakami, czy wyciekami danych;
- 2) wykrywanie wizerunku dziecka na zdjęciu (*age detection*) uzależnione od wieku – w zależności od wieku i kontekstu: pełna blokada *uploade'u*; *upload* ograniczony czasowo z możliwością przedłużenia (np. konta społecznościowe odnawiane co 30 dni – zastosowanie wobec platform *reverse subscription*; po uzyskaniu wskazanych wyżej zgód – *upload* na stałe treści opatrzonej znakiem wodnym);
- 3) bezpieczeństwo reprezentacji cyfrowej dziecka to pełna widoczność uzyskiwana tylko w wyniku autoryzacji podpisem cyfrowym; usuwanie przez użytkownika plików z komunikatora w dowolnym momencie;
- 4) automatyczne usuwanie metadanych – obowiązek retencji usuniętych metadanych na potrzeby tzw. *urgent request*.

PRZYKŁAD: usługa, z której korzysta użytkownik (np. mObywatel, zakładka „bank reprezentacji cyfrowych”), przechowuje wizerunek dziecka na zaszyfrowanych subkontach, do których operator usługi (państwo) nie ma dostępu. Subkonta wydzielone są dla każdej grupy wiekowej oraz – wewnątrz nich – dla przestrzeni publicznej i przestrzeni prywatnej. Usługodawca oferujący aplikację A, po zweryfikowaniu tożsamości użytkownika (dziecka lub jego przedstawiciela ustawowego), jego praw i obowiązków, przekazuje do „banku reprezentacji” informację o działaniach własnych oraz osób trzecich podejmowanych wobec reprezentacji cyfrowej na podstawie ustalonego katalogu.

3) Poziom trzeci

System reagowania to środki technologiczne zapewniające reakcję na nadużycia.

1. **Środki obywatelskie** obejmują takie rozwiązania jak:

1.1. Asystent egzekwowania praw dziecka do reprezentacji cyfrowej

³ W kontekście respektowania podmiotowości dziecka oraz jego prawa do prywatności, warto rozważyć ustalenie wieku zgody na prowadzenie nienadzorowanej korespondencji.

Po osiągnięciu 13 (lub 16/17) r.ż. dziecko ma prawo:

- a) zweryfikować wszystkie swoje reprezentacje cyfrowe opublikowane przez przedstawiciela ustawowego dziecka;
- b) zażądać ich usunięcia tzw. jednym kliknięciem;
- c) zablokować dalsze publikacje.

1.2. Cyfrowy „paszport dziecka”

Każde dziecko (oraz osoba pełnoletnia) posiada cyfrowy paszport połączony z:

- a) bankiem zgód oraz bankiem reprezentacji cyfrowych;
- b) systemem zgłaszania nadużyć i przestępstw.

1.3. Automatyczne monitorowanie Internetu

Dziecko oraz jego przedstawiciel ustawowy może:

- a) otrzymywać alerty, gdy reprezentacja cyfrowa dziecka pojawi się w cyberprzestrzeni w kontekście nadużyć;
- b) wykrywać nieautoryzowane użycia (np. *deepfake*, memy);
- c) zgłaszać naruszenia.

2. Środki organów ścigania i wymiaru sprawiedliwości obejmują:

2.1. bezpieczny kanał zgłoszeniowy i procedurę szybkiego zabezpieczenia dowodów – skoordynowany system wymiany danych obejmujący zgłaszanie naruszeń do organów ścigania oraz wyznaczonych – według katalogu – jednostek z zakresu cyberbezpieczeństwa krajowego (np. WCPR 112, CERT, Dyżurnet.pl, Niebieska Linia 116 123) z automatycznym powiadomieniem drugiej strony skutkujący ochroną pokrzywdzonego przed instrumentalizacją i wtórną wiktyimizacją w toku postępowania;

2.2. bazę wizerunków i hashy, która obejmuje:

- a) system agregujący dane i informacje związane z przestępstwami;
- b) szybki i bezpieczny obieg plików w ramach procesu karnego (w uzasadnionych przypadkach także innych postępowań sądowych i przedsądowych);
- c) uzasadnioną dobrem dziecka współpracę z podmiotami zagranicznymi, w tym administratorami ICSE DB, NCMEC Take It Down, INHOPE CPORT itp.

3. Laboratorium AI – system oceny autentyczności i pochodzenia treści⁴.

4. Wnioski

- 1) Cyfrowa reprezentacja dziecka powinna służyć przede wszystkim ochronie jego dobra, godności, prywatności i bezpieczeństwa, a nie wyłącznie realizacji uprawnień osób pełnoletnich.

⁴ <https://c2pa.org/>

- 2) Udostępnianie wizerunku, danych i informacji dotyczących dziecka wymaga uwzględnienia jego wieku, dojrzałości oraz – w miarę możliwości – jego zdania.
- 3) Wykonywanie władzy rodzicielskiej w środowisku cyfrowym nie może prowadzić do nadmiernej ingerencji w prywatność dziecka ani do tworzenia trwałego śladu cyfrowego, który może naruszać jego interesy w przyszłości.
- 4) Konieczne jest wyraźne określenie granic reprezentacji dziecka przez jego przedstawicieli ustawowych, a także reguł działania instytucji publicznych oraz podmiotów prowadzących platformy internetowe.
- 5) Regulacje prawne powinny zapewniać dziecku skuteczne środki żądania usunięcia treści, ograniczenia ich rozpowszechniania oraz ochrony przed dalszym naruszaniem jego praw.
- 6) KKPR opowiada się za stworzeniem spójnego modelu ochrony dziecka w środowisku cyfrowym, opartego na zasadzie dobra dziecka, jego podmiotowości oraz stopniowo zwiększającej się autonomii.

Opracowanie:

P. Oberszt