Ministerstwo
Cyfryzacji

Sekretarz Stanu
Dariusz Standerski

BM.WP.057.110.2026
Warszawa, 18 lipca 2026 r.

Szanowny Pan
Włodzimierz Czarzasty
Marszałek Sejmu RP

Dot. pisma z 15 czerwca br. Posła na Sejm RP Pana Waldemara Sługockiego w sprawie narastającego zagrożenia oszustwami internetowymi, wykorzystaniem technologii deepfake oraz bezpieczeństwa danych obywateli (interpelacja nr 17618)

Szanowny Panie Marszałku,

poniżej przedstawiam odpowiedzi na zadane przez Posła pytania.

Ad 1) Jakie działania podejmuje obecnie Ministerstwo Cyfryzacji w celu przeciwdziałania oszustwom internetowym wykorzystującym płatności BLIK oraz fałszywe reklamy inwestycyjne?

Ministerstwo Cyfryzacji i współpracujące z nim zespoły reagowania na incydenty komputerowe (m.in. NASK, KNF) realizują wielotorowe działania. Jednym z głównych narzędzi jest prowadzona przez CERT Polska Lista Ostrzeżeń przed niebezpiecznymi domenami, na którą w 2025 r. wpisano aż 244 341 szkodliwych domen, dzięki czemu zablokowano około 141,1 miliona prób wejść na fałszywe strony. Znaczący wkład w ten system ma CSIRT KNF, który w samym tylko 2025 r. zidentyfikował i zgłosił do blokady 41 751 niebezpiecznych domen, z czego ponad 96% dotyczyło właśnie ofert fałszywych inwestycji. Ponadto, na mocy ustawy o zwalczaniu nadużyć w komunikacji elektronicznej, uruchomiono system wymiany wzorców złośliwych SMS-ów (tzw. smishing), co zobowiązuje operatorów telekomunikacyjnych do ich automatycznego blokowania. System ten jest systematycznie rozwijany – tylko w 2025 r. na podstawie 790 opracowanych wzorców operatorzy zablokowali 1 883 610 złośliwych wiadomości SMS.

Niezależnie od wskazanych wyżej działań, istotnym wyzwaniem pozostają fałszywe reklamy publikowane w mediach społecznościowych, których skala wzrosła wraz z rozwojem narzędzi wykorzystujących sztuczną inteligencję do generowania obrazów i materiałów wideo. W reklamach tych często wykorzystywane są wizerunki osób przypominających osoby publiczne, w tym polityków oraz przedstawicieli najwyższych władz państwowych.

Ministerstwo Cyfryzacji wraz z NASK – PIB w ramach projektu „Monitoring treści o potencjale dezinformacyjnym w mediach społecznościowych i stronach internetowych oraz prace rozwojowe nad budowaniem odporności administracji publicznej oraz społeczeństwa wobec zjawiska dezinformacji poprzez działania naukowe i edukacyjne” monitoruje Polską infosferę. Każdy wykryty przypadek jest zgłaszany do platform społecznościowych celem usunięcia, jednakże do momentu wdrożenia Polska ma ograniczone możliwości wpływania na korporacje technologiczne w tym zakresie.

Ad 2) Czy Ministerstwo prowadzi analizy dotyczące liczby oszustw dokonanych z wykorzystaniem technologii deepfake? Jeśli tak, proszę o przedstawienie danych za lata 2023–2026.

Prowadzone są analizy dotyczące oszustw internetowych wykorzystujących technologie sztucznej inteligencji, w tym materiały typu deepfake, jednak precyzyjne określenie liczby takich przypadków jest istotnie utrudnione. Wynika to z charakteru zjawiska: materiały

telefon +48 22 245 54 42
adres email sekretariat.BM@cyfra.gov.pl
www.gov.pl/web/cyfryzacja

ul. Królewska 27
00-060 Warszawa

tego typu są szybko rotowane, usuwane i publikowane ponownie w kolejnych wariantach, często z niewielkimi zmianami technicznymi lub wizualnymi. Część kampanii wykorzystuje także mechanizmy maskowania treści, przekierowania do zewnętrznych stron oraz reklamy widoczne dla użytkowników (szara strefa), ale niewystępujące w takiej samej postaci w publicznych bibliotekach reklam. Z tego względu obserwowalny jest delikatny spadek liczby wykrywanych materiałów, który nie musi oznaczać proporcjonalnego spadku rzeczywistej aktywności sprawców.

Należy również podkreślić, że nie wszystkie treści wykorzystujące AI mają charakter klasycznego deepfake, czyli bezpośredniego podszywania się pod konkretną osobę. Obserwowane są również syntetyczne lub silnie przetworzone reklamy, treści wysokiego ryzyka oraz materiały wprowadzające w błąd, które mogą kierować obywateli do fałszywych stron, podejrzanych produktów lub schematów inwestycyjnych. Z prowadzonych analiz wynika, że skala zjawiska jest znacząca. Szacunki oparte na monitoringu kampanii oszukańczych wskazują, że przy utrzymaniu intensywności obserwowanej w okresach wzmożonej aktywności liczba rdzeniowych materiałów oszukańczych wykorzystujących treści syntetyczne lub manipulowane mogłaby przekraczać 160 tys. rocznie. Należy jednak podkreślić, że jest to szacunek orientacyjny. Dane za lata 2023-2026 są analizowane z uwzględnieniem zmieniającej się metodologii klasyfikacji. W poprzednich latach część takich przypadków mogła być ujmowana w szerszych kategoriach, takich jak phishing, fałszywe reklamy inwestycyjne, podszywanie się pod osoby lub instytucje, wyłudzenia danych albo reklamy wprowadzające w błąd. Z tego powodu porównywanie danych rok do roku wymaga ostrożności.

Prowadzone analizy mają charakter analityczny i służą monitorowaniu zjawiska, jednak Ministerstwo Cyfryzacji nie prowadzi odrębnych statystyk dotyczących wyłącznie oszustw z wykorzystaniem technologii deepfake. Tego rodzaju incydenty nie stanowią wyodrębnionej kategorii w oficjalnych systemach zgłoszeń i raportowania, lecz są ujmowane w szerszej kategorii „oszustw komputerowych". W związku z tym dostępne są wyłącznie dane dotyczące ogólnej liczby oszustw komputerowych, która wyniosła 97 995 incydentów w 2024 r. oraz 253 238 incydentów w 2025 r.

Ad 3) Jakie działania edukacyjne są prowadzone wobec seniorów w zakresie bezpieczeństwa cyfrowego i rozpoznawania prób wyłudzeń internetowych?

Ministerstwo Cyfryzacji prowadzi działania informacyjne i edukacyjne mające na celu podnoszenie kompetencji cyfrowych wśród seniorów, co w szerszej perspektywie wpływa na bezpieczne korzystanie w technologii oraz możliwość rozpoznawania prób wyłudzeń internetowych.

W zakresie edukacji seniorów prowadzone są także liczne działania mające na celu zwiększenie świadomości zagrożeń w sieci oraz rozwijanie umiejętności rozpoznawania prób oszustw internetowych. Organizowane są szkolenia, warsztaty i spotkania edukacyjne, m.in. na Uniwersytetach Trzeciego Wieku, w domach i klubach seniora, bibliotekach, centrach aktywności lokalnej oraz innych instytucjach i organizacjach zrzeszających osoby starsze. Działania te uzupełniają materiały edukacyjne, takie jak poradniki, broszury, publikacje i infografiki, które w przystępny sposób wyjaśniają najczęstsze metody działania cyberprzestępców oraz zasady bezpiecznego korzystania z internetu i nowych technologii.

Przykładem takich działań jest ogólnopolski program „Świadomy Senior – Bezpieczny Senior", realizowany przez Ministerstwo Sprawiedliwości we współpracy z NASK, Komendą Główną Policji, Zakładem Ubezpieczeń Społecznych, Urzędem Ochrony Konkurencji i Konsumentów, Biurem Rzecznika Praw Pacjenta oraz Krajową Radą Notarialną. W ramach programu w całej Polsce organizowane są bezpłatne spotkania

edukacyjne dla seniorów poświęcone m.in. rozpoznawaniu prób wyłudzeń, bezpiecznemu korzystaniu z internetu oraz ochronie danych osobowych.

Równolegle NASK prowadzi, we współpracy z Centralnym Biurem Zwalczania Cyberprzestępczości oraz Warszawskim Instytutem Bankowości (WIB), kampanię „#Halo! Tu cyberbezpieczny Senior", w ramach której przygotowano szeroki pakiet materiałów edukacyjnych dla osób starszych. Obejmuje on m.in. poradnik „#Halo! Tu cyberbezpieczny Senior", ulotki, plakaty oraz cykl artykułów dostępnych w Bazie Wiedzy kampanii Europejskiego Miesiąca Cyberbezpieczeństwa. Publikacje poruszają zagadnienia związane z phishingiem, spoofingiem, oszustwami telefonicznymi, socjotechniką oraz zasadami bezpiecznego korzystania z internetu i bankowości elektronicznej.

Ministerstwo Cyfryzacji koordynuje realizację rządowego Programu Rozwoju Kompetencji Cyfrowych (PRKC). Obecnie trwa aktualizacja PRKC i projekt uchwały Rady Ministrów zmieniającej uchwałę w sprawie ustanowienia programu rządowego pod nazwą „Program Rozwoju Kompetencji Cyfrowych" jest wpisany do Wykazu prac Rady Ministrów (ID274).

Działania edukacyjne zaplanowane w PRKC mają na celu podniesienie poziomu kompetencji cyfrowych w Polsce i są adresowane do całego społeczeństwa w podziale na następujące grupy docelowe:

- dzieci w wieku przedszkolnym;
- uczniowie;
- studenci;
- nauczyciele i edukatorzy;
- użytkownicy technologii cyfrowych;
- osoby stawiające pierwsze kroki w świecie cyfrowym, w tym seniorzy;
- osoby pracujące;
- przedsiębiorcy i osoby zarządzające;
- Pracownicy sektora publicznego;
- specjaliści ICT.

Kompetencje cyfrowe definiujemy jako harmonijną kompozycję wiedzy, umiejętności i postaw umożliwiających życie, uczenie się i pracę w społeczeństwie cyfrowym. Opanowanie podstawowych kompetencji cyfrowych (jednym z celów PRKC jest, aby 80% społeczeństwa miało je do 2030 r.) oznacza posiadanie co najmniej podstawowych umiejętności w każdym z 5 następujących obszarów:

1. umiejętności informacyjne i korzystania z danych;
2. umiejętności komunikacji i współpracy;
3. umiejętności tworzenia treści cyfrowych;
4. umiejętności w obszarze bezpieczeństwa;
5. umiejętności rozwiązywania problemów.

Działania edukacyjne w zakresie bezpieczeństwa cyfrowego i rozpoznawania prób wyłudzeń internetowych, mieszczą się w obszarze nr 1 i 4. Warto podkreślić, że bez opanowania umiejętności w tych obszarach nie można mówić o posiadaniu podstawowych kompetencji cyfrowych. Dlatego wszystkie interwencje mające na celu podniesienie kompetencji cyfrowych seniorów, dotyczą przedmiotowego pytania.

Do kluczowych działań PRKC, które są prowadzone wobec seniorów należą:

1) Systemowe wsparcie edukacji cyfrowej dorosłych użytkowników ICT - Kluby Rozwoju Cyfrowego (KRC). Głównym celem KRC jest podniesienie poziomu

kompetencji cyfrowych społeczeństwa przez stworzenie na poziomie lokalnym warunków wsparcia osób dorosłych, w szczególności osób starszych, w nabywaniu umiejętności cyfrowych. Oferta edukacji cyfrowej ma być dostępna dla wszystkich osób dorosłych w Polsce, a w szczególności do mieszkańców małych miejscowości i wsi. Projekt został podzielony na 3 części:

1. KRC Wsparcie ma na celu wypracowanie ram funkcjonowania KRC, tworzenie materiałów edukacyjnych, przygotowanie edukatorów oraz wsparcie merytoryczne ich działań przez cały czas realizacji projektu.
2. KRC Pilotaż, trwający obecnie w 40 gminach w Polsce, ma na celu przetestowanie modelu funkcjonowania KRC i identyfikacji najlepszych rozwiązań do wdrożenia na kolejnym etapie.
3. KRC Skalowanie ma na celu utworzenie w całej Polsce sieci ponad 2000 KRC. Będą one prowadzone przez wyłonionych w konkursie Beneficjentów (operatorów), którzy na wybranych obszarach będą, przy współudziale chętnych gmin, prowadzić określoną z góry liczbę KRC. Obecnie trwa ocena wniosków potencjalnych operatorów, natomiast Kluby z etapu KRC Skalowanie powstaną na jesieni 2026 r.

Całe przedsięwzięcie jest finansowane ze środków europejskich Programu Fundusze Europejskie dla Rozwoju Społecznego.

2) Szkolenia dla obywateli z kompetencji cyfrowych w ramach inwestycji C2.1.3 E-kompetencje. Celem działania jest podniesienie kompetencji cyfrowych umożliwiających obywatelom funkcjonowanie we współczesnym świecie. W zależności od potrzeb obywatele, w tym osoby starsze zostaną wyposażeni w takie umiejętności i kompetencje cyfrowe jak: korzystanie z e-usług, załatwianie spraw administracyjnych, obsługa komputera, komunikacja online, wykorzystanie użytecznych narzędzi cyfrowych, w tym programów ułatwiających funkcjonowanie osób niepełnosprawnych, umożliwiających nauczanie i uczenie się online, pracę zdalną, korzystanie ze środków identyfikacji elektronicznej, bezpieczne, świadome i krytyczne korzystanie z mediów i technologii cyfrowych. Łącznie do początku lipca 2026 r. planowane jest przeszkolenie około 169 00 osób w module obywatele oraz osoby wykluczone cyfrowo. Na ten cel przeznaczono około 267 mln zł ze środków Krajowego Planu Odbudowy.

3) Kampanie edukacyjno-informacyjne na rzecz rozwoju kompetencji cyfrowych. Celem działania jest zachęcanie osób dorosłych, w tym seniorów do rozwijania kompetencji cyfrowych oraz wspieranie realizacji celów PRKC. Inicjatywa składa się z kilku wielowątkowych kampanii edukacyjno-informacyjnych. Horyzontalnym tematem, pojawiającym się we wszystkich wątkach kampanii jest motywowanie społeczeństwa do poszerzania wiedzy i umiejętności cyfrowych. Nie mniej istotnym elementem kampanii jest propagowanie wiedzy w zakresie higieny cyfrowej, w celu wsparcia zdrowia psychicznego i fizycznego obywateli. Działanie jest finansowane ze środków europejskich KPO.

4) Projekt pn. Szkoła międzypokoleniowa. Projekt jest realizowany z inicjatywy Ministerstwa Cyfryzacji, we współpracy z Ministerstwem Edukacji Narodowej oraz Pełnomocnikiem Rządu do spraw Polityki Senioralnej. Operatorem projektu jest NASK – Państwowy Instytut Badawczy. Celem działań jest integracja międzypokoleniowa oraz rozwój kompetencji cyfrowych wśród osób starszych, przy wsparciu i zaangażowaniu uczniów klas VII–VIII szkół podstawowych oraz uczniów szkół ponadpodstawowych. W okresie realizacji pilotażu przedsięwzięcia (lata 2024-2025) zorganizowano ponad 5,5 tys. warsztatów edukacyjnych, zaangażowano 1554 szkoły z całej Polski. W warsztatach wzięło udział 11 260

uczniów i 10 349 seniorów. 638 z 887 szkół, które przesłały sprawozdanie, znajduje się w miejscowościach poniżej 20 tys. mieszkańców. Zajęcia skupiały się na praktycznym rozwijaniu umiejętności korzystania z Internetu (w tym z usług cyfrowych), obsługi urządzeń cyfrowych oraz rozpoznawania zagrożeń online. Młodzież pełniła rolę przewodników cyfrowych, dzieląc się wiedzą i wspierając seniorów w nauce, jednocześnie korzystając z możliwości wymiany międzypokoleniowych doświadczeń.

Obecnie trwa druga edycja projektu, w której bierze udział ponad 1000 szkół i będzie on kontynuowany w kolejnych latach.

5) Projekt szkoleniowy dla seniorów w sanatoriach pn. „eFajfy" Przykładem innego działania jest projekt szkoleniowy „eFajfy", skierowany do osób powyżej 60. roku życia przebywających w sanatoriach oraz w klubach seniorów, domach kultury i innych miejscach poświęconych seniorom (np. uniwersytetach trzeciego wieku itd.). Program obejmuje bezpłatne szkolenia teoretyczne i praktyczne, prowadzone z wykorzystaniem sprzętu cyfrowego, których celem jest zarówno wprowadzenie uczestników w podstawy korzystania z technologii, jak i rozwijanie umiejętności korzystania z e-usług publicznych oraz zwiększanie świadomości w zakresie bezpieczeństwa w internecie. W ramach zajęć omawiane są m.in. zagadnienia związane z zakładaniem kont e-mail, korzystaniem z komunikatorów, obsługą profilu zaufanego, a także identyfikowaniem prób oszustw, w tym metodą „na wnuczka". W pilotażu programu w 2025 r. przeszkolono 1 015 seniorów w 76 sanatoriach, w kolejnych latach planuje się objęcie szkoleniami ok. 8 000 seniorów rocznie.

Ad 4) Czy planowane jest uruchomienie ogólnopolskiej kampanii społecznej ostrzegającej przed oszustwami wykorzystującymi sztuczną inteligencję oraz podszywaniem się pod osoby publiczne?

Ministerstwo Cyfryzacji samodzielnie oraz we współpracy z NASK – PIB prowadzi kampanie medialne dotyczące przeciwdziałania dezinformacji, kampanie te dotyczą pośrednio kwestii oszustw i wykorzystywania sztucznej inteligencji, zachęcając do weryfikacji informacji i krytycznego myślenia.

Ad 5) Czy Ministerstwo prowadzi rozmowy z operatorami mediów społecznościowych oraz platform reklamowych w zakresie szybszego usuwania fałszywych reklam inwestycyjnych i materiałów deepfake?

Ministerstwo Cyfryzacji we współpracy z NASK – Państwowym Instytutem Badawczym prowadzi działania służące identyfikowaniu i zgłaszaniu do platform internetowych treści o charakterze dezinformacyjnym, w tym materiałów wykorzystujących technologie sztucznej inteligencji, takich jak deepfake. W 2025 r. zgłoszono do platform społecznościowych 46 511 materiałów dezinformacyjnych, jednak jedynie część z nich została niezwłocznie usunięta. Skuteczność tych działań pozostaje uzależniona od polityk moderacyjnych poszczególnych platform oraz tempa wdrażania przepisów wynikających z Aktu o usługach cyfrowych (DSA). Ministerstwo oraz NASK pozostają w stałym kontakcie z operatorami platform internetowych, przy czym obecnie prowadzone rozmowy koncentrują się przede wszystkim na zagadnieniach związanych ze stosowaniem przepisów prawa oraz ich implementacją w Polsce.

W ramach działalności Ośrodka Analizy Dezinformacji (OAD) NASK monitorowane są również przypadki podszywania się pod osoby publiczne oraz instytucje w mediach społecznościowych. Kategoria ta obejmuje m.in. reklamy inwestycyjne wykorzystujące technologie deepfake, jak również inne materiały wykorzystujące wizerunek osób publicznych lub instytucji w sposób mogący wprowadzać odbiorców w błąd. Jednocześnie

należy zaznaczyć, że zgodnie z metodyką stosowaną przez OAD NASK kategoria ta jest szersza niż same reklamy typu deepfake i obejmuje także fałszywe konta oraz strony internetowe służące podszywaniu się pod osoby lub instytucje.

Z danych przedstawionych w raporcie „Reakcja platform społecznościowych na zgłoszenia w 2025 roku" wynika, że w 2025 r. OAD NASK zgłosił 3 573 przypadki podszywania się, stanowiące 7,7% wszystkich zgłoszonych treści. Spośród nich usunięto 2 226 materiałów, 228 poddano moderacji, natomiast 1 119 zgłoszeń zostało uznanych przez platformy za niezasadne i materiały te pozostały dostępne.

W przypadku Mety nie oferuje się dla użytkowników skutecznych, proaktywnych mechanizmów wykrywania i blokowania treści typu deepfake w krajach Unii Europejskiej. Skutkuje to między innymi pojawianiem się reklam i wpisów wykorzystujących wizerunek osób publicznie (powszechnie) znanych w celu zachęcania do udziału w fałszywych inwestycjach czy nakłaniania do kupna podejrzanych produktów. Jest to jednak szerszy problem również na innych platformach społecznościowych, jak TikTok czy YouTube (Google Ads). Ponadto w wielu przypadkach, na różnych serwisach, moderacja tego rodzaju treści jest nietransparentna i niekonsekwentna – podobne do siebie materiały i konta czasami są blokowane, a czasami nie, bez podania jasnego uzasadnienia.

Istnieją także techniczne trudności w identyfikowaniu szkodliwych reklam publikowanych na platformach. Kolejny raz za przykład posłuży serwis Mety. Informacje zawarte w udostępnionej przez firmę bibliotece reklam są aktualizowane z opóźnieniem (do 72h). Zdarza się, że użytkownicy serwisu Facebook natrafiają na szkodliwe treści, które nie są jeszcze widoczne w rejestrze, co znacząco utrudnia zgłaszanie ich do zespołu moderującego platformę. Kolejnym aspektem dotyczącym tego obszaru jest również blokowanie przez platformy, w odpowiedzi na zgłoszenie, jedynie pojedynczej reklamy, zamiast całego konta, które wielokrotnie publikuje podobne treści.

Częstą praktyką stosowaną przez aktorów dezinformacji w celu utrudnienia identyfikacji szkodliwych treści reklamowych jest publikowanie tej samej reklamy w wielu wersjach (z ang. tzw. cloaking). Główny wariant odsyła do przekazu niezwiązanego z właściwą treścią reklamy, często przybierając formę neutralnego obrazu, na przykład zachodu słońca lub elementu garderoby. Zdarza się, że w wersjach dostępnych w bibliotece reklam nie występuje oryginalny przekaz, który wcześniej był widoczny na tablicy użytkownika. Zabieg ten nie tylko utrudnia identyfikację szkodliwego materiału, ale również ogranicza możliwość przygotowania dokładnego zgłoszenia do administracji serwisu Meta. W przeszłości proceder ten był zgłaszany do punktu kontaktowego, jednak nie doprowadziło to do ograniczenia takich praktyk na platformach firmy Meta.

Trudności z monitorowaniem szkodliwych przekazów występują również w bibliotece reklam udostępnionej przez serwis Google. Nie ma tam możliwości wyszukiwania treści po słowach kluczowych. Aby uzyskać informacje o publikowanych materiałach, konieczne jest wskazanie nazwy konkretnego reklamodawcy.

Problemy te były wielokrotnie zgłaszane przez OAD NASK przedstawicielom platform społecznościowych. Przykładem są działania prowadzone w okresie kampanii prezydenckiej, podczas której identyfikowano liczne materiały wygenerowane przy wykorzystaniu sztucznej inteligencji, nieoznaczone jako treści syntetyczne. Pomimo przekazania platformom konkretnych przykładów oraz deklaracji podjęcia działań, do chwili obecnej OAD NASK nie otrzymał informacji o sposobie ich rozpatrzenia, a część wskazanych materiałów nie została odpowiednio oznaczona.

Ad 6) Jakie procedury obowiązują obecnie w przypadku zgłoszenia przez obywatela fałszywej reklamy inwestycyjnej lub próby wyłudzenia danych?

W przypadku zgłoszenia przez obywatela fałszywej reklamy inwestycyjnej lub próby wyłudzenia danych (phishingu), obowiązują obecnie procedury obejmujące zarówno mechanizmy funkcjonujące po stronie platform internetowych, jak i działania właściwych instytucji publicznych.

Zgłoszenie do platformy internetowej w oparciu o obowiązki wynikające z unijnego rozporządzenia 2022/2065 Akt o usługach cyfrowych

Zgodnie z rozporządzeniem Parlamentu Europejskiego i Rady (UE) 2022/2065 (akt o usługach cyfrowych – DSA), dostawcy usług pośrednich, w tym platform internetowych, są zobowiązani do:

- zapewnienia użytkownikom łatwo dostępnych mechanizmów zgłaszania nielegalnych treści,
- niezwłocznego rozpatrzenia zgłoszenia w ramach procedury: „mechanizmy zgłaszania i działania",
- poinformowania użytkownika o podjętej decyzji oraz jej uzasadnieniu.

Zgłoszenie incydentu cyberbezpieczeństwa

Obywatel może zgłosić próbę wyłudzenia danych do:

- CSIRT NASK (np. za pośrednictwem portalu https://incydent.cert.pl lub poprzez przekazanie podejrzanych wiadomości SMS na numer 8080),

Zgłoszenia te są analizowane pod kątem identyfikacji zagrożeń, a w uzasadnionych przypadkach podejmowane są działania mające na celu ograniczenie ich rozprzestrzeniania.

Zgłoszenie organom ścigania

W przypadku poniesienia szkody majątkowej lub podejrzenia przestępstwa, obywatel może złożyć zawiadomienie do Policji lub prokuratury.

Zgłoszenie do właściwych organów nadzorczych

W zależności od charakteru zdarzenia możliwe jest również:

- zgłoszenie do Komisji Nadzoru Finansowego (KNF) – w przypadku fałszywych ofert inwestycyjnych,
- zgłoszenie do Prezesa UOKiK – w przypadku praktyk naruszających zbiorowe interesy konsumentów.

Ad 7) Czy Ministerstwo rozważa wprowadzenie obowiązku oznaczania treści generowanych przez sztuczną inteligencję, które mogą wpływać na bezpieczeństwo obywateli?

Zgodnie z przepisami rozporządzenia w sprawie sztucznej inteligencji ('AI Act') przewidziano obowiązek informowania odbiorców o tym, że dana treść została wygenerowana lub zmodyfikowana z wykorzystaniem systemów AI, w szczególności w przypadkach treści syntetycznych, takich jak obraz, audio, wideo czy tekst, a także deepfake'ów. Celem tych regulacji jest zwiększenie przejrzystości oraz ograniczenie ryzyka wprowadzania obywateli w błąd, zwłaszcza w sytuacjach, w których treści te mogą oddziaływać na opinię publiczną, bezpieczeństwo informacyjne lub zaufanie do instytucji publicznych.

Ministerstwo prowadzi obecnie prace związane z umożliwieniem stosowania rozporządzenia w Polsce poprzez projekt ustawy o systemach sztucznej inteligencji. Projekt ustawy znajduje się obecnie u Prezydenta. W tym kontekście zastosowanie znajdą

przepisy art. 50 rozporządzenia, które przewidują obowiązek oznaczania treści generowanych lub zmanipulowanych przy użyciu sztucznej inteligencji. Zgodnie z art. 50 ust. 2 dostawcy systemów AI generujących treści będą zobowiązani do zapewnienia, aby wygenerowane materiały były oznakowane w sposób umożliwiający ich wykrycie jako sztucznie wygenerowanych lub zmanipulowanych. Natomiast art. 50 ust. 4 AI Act nakłada obowiązek ujawniania przez podmiot stosujący system sztucznej inteligencji, który generuje obrazy, treści audio lub wideo stanowiące treści deepfake lub który manipuluje takimi treściami, że treści te zostały sztucznie wygenerowane lub zmanipulowane. Przepisy te będą stosowane bezpośrednio we wszystkich państwach Unii Europejskiej od 2 sierpnia 2026, z wyjątkiem art. 50 ust. 2, który wejdzie w życie 2 grudnia 2026.

Ad 8) Ile incydentów związanych z naruszeniem bezpieczeństwa danych obywateli odnotowano w latach 2023–2025 w podmiotach publicznych oraz spółkach Skarbu Państwa?

CSIRT NASK nie prowadzi statystyk dotyczących wyłącznie incydentów związanych z naruszeniem bezpieczeństwa danych obywateli ani odrębnych zestawień obejmujących wyłącznie spółki Skarbu Państwa.

Ad 9) Czy planowane są zmiany legislacyjne zwiększające odpowiedzialność platform internetowych za publikowanie reklam o charakterze oszukańczym?

Prowadzone są prace legislacyjne w ramach dwóch nowelizacji ustawy o świadczeniu usług drogą elektroniczną (projekty o numerach w wykazie prac legislacyjnych Rady Ministrów: UC140 i UC141). Są to prace zmierzające do zapewnienia pełnego stosowania unijnego rozporządzenia 2022/2065 Akt o usługach cyfrowych. Przewidziane rozwiązania prawne mają na celu zwiększenie odpowiedzialności platform internetowych oraz skuteczniejszą walkę z nielegalnymi treściami, w tym z oszustwami. Oba projekty zostały przyjęte przez Radę Ministrów 2 czerwca 2026 r. (więcej na ten temat na stronie gov.pl) i zostały 9 czerwca br. przekazane do Sejmu RP. 17.06 projekty skierowano do I czytania w komisjach sejmowych. Projektom tym nadano numery druków sejmowych: druk nr 2694 (projekt UC140) oraz druk nr 2693 (projekt UC141).

Wdrożenie Aktu o usługach cyfrowych (DSA) wprowadza w zakresie powiązanym z pytaniem następujące rozwiązania:

- procedura szybkiego blokowania treści nielegalnych: nowelizacja wprowadza krajową procedurę wydawania nakazów ograniczenia dostępu do nielegalnych treści,
- katalog czynów zabronionych, które mogą być podstawą do wydania takiego nakazu, obejmuje m.in. oszustwa internetowe i phishing, a także doprowadzenie do niekorzystnego rozporządzenia mieniem przez wprowadzenie w błąd (art. 286 § 1 i 2 kodeksu karnego) oraz oszustwo komputerowe (art. 287 § 1 kodeksu karnego),
- platformy internetowe będą zobowiązane do blokowania treści wskazanych w decyzjach uprawnionych organów. Organy te (np. Policja, prokuratura) będą wydawały decyzje w krótkich terminach – standardowo 7 dni, a w pilnych sprawach prowadzonych przez służby – nawet w 2 dni.

Wprowadzenie tych mechanizmów ma na celu zwiększenie skuteczności przeciwdziałania nielegalnym treściom w internecie.

Ad 10) Jak Ministerstwo ocenia aktualny poziom przygotowania administracji publicznej do przeciwdziałania cyberzagrożeniom wykorzystującym sztuczną inteligencję?

Ministerstwo Cyfryzacji ocenia, że poziom przygotowania administracji publicznej do przeciwdziałania cyberzagrożeniom wykorzystującym sztuczną inteligencję jest

systematycznie wzmacniany, jednak ze względu na dynamiczny rozwój technologii AI oraz stale ewoluujące metody działania cyberprzestępców wymaga ciągłego doskonalenia kompetencji, narzędzi oraz procedur.

Istotnym elementem budowania odporności administracji są działania edukacyjne i szkoleniowe. W ramach programu SecureV konsekwentnie realizowane są szkolenia dla osób pełniących kluczowe funkcje w państwie oraz kadry zarządzającej administracją publiczną w zakresie zagrożeń socjotechnicznych, cyberhigieny oraz metod przeciwdziałania oszustwom. Do końca 2025 r. programem objęto 16 729 osób.

W 2025 r. zorganizowano 32 szkolenia, w których uczestniczyło blisko 43 tys. osób. Działania te obejmowały m.in.:

- Szkolenia z cyberhigieny: Realizowane we współpracy z NASK-PIB (m.in. cykle „Cyberzagrożenia - bądź na bieżąco!"), w których w 2025 r. przeszkolono ponad 18 tysięcy osób z podstawowych zasad bezpieczeństwa.
- Zajęcia we współpracy z rynkiem prywatnym (PWCyber): Przeprowadzono 24 szkolenia online przy współpracy z partnerami technologicznymi Programu PWCyber, ukierunkowane na podniesienie praktycznych umiejętności operacyjnych i radzenia sobie w sytuacjach kryzysowych.

Wiedzę kadr administracji i podmiotów KSC wzmacnia się także poprzez zaawansowane warsztaty stacjonarne. Należy tu wymienić m.in. specjalistyczne warsztaty z zakresu Cyber Threat Intelligence (CTI) zorganizowane z jednym z partnerów PWCyber dla kilkunastu kluczowych instytucji (w tym analityków z CSIRT GOV, CSIRT KNF, CBZC czy Ministerstwa Infrastruktury).

W zakresie przeciwdziałania Dezinformacji trwa budowa mechanizmów przeciwdziałania Dezinformacji (w tym generowanej przez AI) o skali działań i możliwości weryfikacyjnych świadczy ponad 46 tysięcy zgłoszeń dezinformacji do platform w roku 2025. Stopień przygotowania na tego typu zagrożenia zawsze może być wyższy, reakcja szybsza i skuteczniejsza. Ocena zagrożeń przyszłych i niepewnych, zależnych od wielu czynników jest trudna do sparametryzowania, czy też pomiaru, stale mierzymy się z nowymi zagrożeniami i reagujemy na nowe działania i metody stosowane przez dezinformujących i poprawiamy swoje zdolności.

Jednocześnie rozwój sztucznej inteligencji wpływa na cały krajobraz cyberbezpieczeństwa. Technologie AI są wykorzystywane zarówno przez cyberprzestępców, jak i przez podmioty odpowiedzialne za ochronę użytkowników i reagowanie na incydenty. Z perspektywy obrońców nie oznacza to jednak pojawienia się całkowicie nowych kategorii zagrożeń, lecz przede wszystkim zwiększenie skali, szybkości i stopnia automatyzacji już znanych metod ataków, takich jak phishing, fałszywe reklamy inwestycyjne czy wyłudzanie danych.

Odpowiedzią na te wyzwania jest rozwój zdolności do wykrywania, analizowania i neutralizowania zagrożeń oraz stałe wzmacnianie współpracy pomiędzy administracją publiczną, CSIRT-ami poziomu krajowego, sektorem prywatnym i partnerami międzynarodowymi. Istotnym elementem tego procesu jest również wdrażanie dyrektywy NIS2 oraz rozwój sektorowych CSIRT-ów, które zwiększają możliwości wymiany informacji i koordynacji działań.

We wzmacnianiu odporności na cyberzagrożenia kluczowy jest również rozwój narzędzi wspierających administratorów i właścicieli infrastruktury.

Przykładem jest serwis moje.cert.pl, w którym udostępniane są bezpłatne narzędzia wspierające monitorowanie bezpieczeństwa domen i usług internetowych. Pozwalają one administratorom i właścicielom domen identyfikować potencjalne zagrożenia, monitorować wycieki danych uwierzytelniających, otrzymywać powiadomienia o

podatnościach i błędnych konfiguracjach, czy spojrzeć na infrastrukturę z zewnątrz, z perspektywy potencjalnego atakującego. Z platformy korzysta już ponad 20 tys. użytkowników. Przeskanowaliśmy dla nich ponad 4 mln domen, subdomen i adresów IP, identyfikując niemal 800 tys. podatności i błędnych konfiguracji, co pozwala usuwać potencjalne słabości jeszcze zanim zostaną wykorzystane przez cyberprzestępców.

Równocześnie specjaliści ds. cyberbezpieczeństwa wykorzystują rozwiązania oparte na sztucznej inteligencji do wspierania swojej pracy. CERT Polska prowadzi badania nad zastosowaniem nowoczesnych metod, w tym autonomicznego fuzzingu, które pomagają szybsze wykrywanie podatności oraz usprawniać analizę zagrożeń (działania te opisane są również w publicznie dostępnym artykule, celem wymiany wiedzy i inspiracji wśród specjalistów z branży: https://cert.pl/posts/2026/05/autonomiczny-fuzzing/).

W kontekście przeciwdziałania oszustwom prowadzonym za pośrednictwem platform internetowych istotne znaczenie będzie miało będzie miało wdrożenie mechanizmu zaufanych podmiotów sygnalizujących (Trusted Flaggers), przewidzianego w rozporządzeniu Digital Services Act (DSA). Obecnie polskie instytucje nie mają możliwości korzystania z tego narzędzia, co sprawia, że zgłoszenia dotyczące fałszywych reklam inwestycyjnych, kampanii phishingowych czy innych oszustw nie są traktowane priorytetowo przez duże platformy internetowe. Ogranicza to skuteczność szybkiego usuwania szkodliwych treści i utrudnia efektywną ochronę użytkowników przed cyberoszustwami. Pilne uruchomienie tego mechanizmu w Polsce jest niezbędne dla zwiększenia skuteczności walki z oszustwami internetowymi.

Mając na uwadze dynamiczny rozwój technologii sztucznej inteligencji oraz stale zmieniający się charakter cyberzagrożeń, Ministerstwo Cyfryzacji konsekwentnie rozwija działania w zakresie edukacji, budowy kompetencji, wzmacniania narzędzi analitycznych oraz współpracy krajowej i międzynarodowej. Poziom przygotowania administracji publicznej do przeciwdziałania cyberzagrożeniom jest systematycznie podnoszony, jednak działania w tym obszarze mają charakter ciągły i są na bieżąco dostosowywane do nowych wyzwań wynikających z rozwoju technologii oraz metod działania cyberprzestępców.

Z wyrazami szacunku
Dariusz Standerski
Sekretarz Stanu
/dokument podpisany elektronicznie/

Do wiadomości:
Kancelaria Prezesa Rady Ministrów - Departament Spraw Parlamentarnych

telefon +48 22 245 54 42
adres email sekretariat.BM@cyfra.gov.pl
www.gov.pl/web/cyfryzacja

ul. Królewska 27
00-060 Warszawa