

Wspólne oświadczenie w sprawie obrazów generowanych przez sztuczną inteligencję i ochrony prywatności 23 lutego 2026 r.

W odpowiedzi na poważne obawy związane z funkcjonowaniem systemów sztucznej inteligencji (AI), generujących realistyczne obrazy i filmy przedstawiające osoby możliwe do zidentyfikowania bez ich wiedzy i zgody, poniżsi współsygnatariusze wydają niniejsze wspólne oświadczenie.

Chociaż sztuczna inteligencja może przynieść znaczące korzyści jednostkom i społeczeństwu, ostatnie wydarzenia — w szczególności integracja narzędzi generujących obrazy i materiały wideo z wykorzystaniem AI z szeroko dostępnymi platformami mediów społecznościowych — umożliwiły tworzenie intymnych materiałów bez zgody osób, których dotyczą, treści zniesławiających oraz innych szkodliwych treści przedstawiających rzeczywiste osoby. Jesteśmy szczególnie zaniepokojeni potencjalnymi szkodami dla dzieci oraz innych szczególnie narażonych grup społecznych, w tym zjawiskami takimi jak cyberprzemoc czy wykorzystywanie.

Oczekiwanie wobec organizacji

Współsygnatariusze przypominają wszystkim organizacjom opracowującym i wykorzystującym systemy AI do generowania treści, że systemy takie muszą być opracowywane i wykorzystywane w ramach obowiązującego prawa, w tym przepisów o ochronie danych osobowych i prywatności.

Jednocześnie podkreśla się, że tworzenie intymnych materiałów bez zgody osoby, której dotyczą, może być traktowane jako przestępstwo w wielu jurysdykcjach.

Chociaż szczególne wymogi prawne różnią się w zależności od jurysdykcji, wszystkie organizacje opracowujące i wykorzystujące systemy AI do generowania treści powinny przestrzegać podstawowych zasad, w szczególności:

- **Wdrożyć solidne zabezpieczenia** mające na celu zapobieganie niewłaściwemu wykorzystaniu danych osobowych, generowaniu intymnych materiałów bez zgody i innych szkodliwych treści, w szczególności dotyczących dzieci.
- **Zapewnić znaczącą przejrzystość** w odniesieniu do możliwości systemu sztucznej inteligencji, stosowanych zabezpieczeń, dopuszczalnych zastosowań oraz konsekwencji niewłaściwego wykorzystania.
- **Zapewnić osobom fizycznym skuteczne i łatwo dostępne** mechanizmy umożliwiające żądanie usunięcia szkodliwych treści zawierających dane osobowe oraz szybkie reagowanie na takie żądania.
- **Zająć się konkretnymi zagrożeniami dla dzieci** poprzez wzmocnione zabezpieczenia oraz zapewnienie dzieciom, rodzicom, opiekunom i wychowawcom przejrzystych informacji odpowiednio dostosowanych do wieku.

Skoordynowana reakcja

Szkody wynikające z generowania intymnych, zniesławiających lub w inny sposób szkodliwych treści przedstawiających rzeczywiste osoby bez ich zgody są znaczące i wymagają pilnej uwagi ze strony organów regulacyjnych.

W celu wspierania rozwoju innowacyjnych systemów sztucznej inteligencji z poszanowaniem prywatności, współsygnatariusze niniejszego oświadczenia wyrażają swoje zaniepokojenie potencjalnymi szkodami wynikającymi z niewłaściwego wykorzystania systemów AI do generowania treści. Współsygnatariusze dążą do wymiany informacji na temat swoich podejść do rozwiązywania tych problemów, co może obejmować działania egzekucyjne, opracowywanie polityki oraz działania edukacyjne, w zakresie zgodnym z obowiązującymi przepisami prawa. Odzwierciedla to nasze zaangażowanie i skoordynowane wysiłki na rzecz przeciwdziałania globalnemu ryzyku.

Wnioski

Wzywamy organizacje do aktywnej współpracy z organami regulacyjnymi, wdrożenia solidnych zabezpieczeń już na etapie projektowania oraz zapewnienia, że postęp technologiczny będzie realizowany z poszanowaniem prywatności, godności, bezpieczeństwa oraz innych praw podstawowych, szczególnie w odniesieniu do osób należących do najbardziej wrażliwych grup społecznych na całym świecie.