

Ochrona prywatności i nowe technologie



W TYM WYDANIU:

Dyrektywa NIS 2 i ustawa o KSC - co warto wiedzieć	1
Podmioty kluczowe i ważne wg ustawy o KSC	2
Mapowanie wymagań ISO 27001 na NIS 2 i ustawę KSC	5
Omenaa Mensah i Rafał Brzoska wygrali z Meta. Przetłomowy ruch prezesa UODO.	11
Zakazane praktyki w zakresie AI według art. 5 AI Act	13
AI Act - Unijne rozporządzenie a domniemanie niewinności	20



Olga Ograbisz

Dyrektywa NIS 2 i ustawa o KSC - co warto wiedzieć

Dyrektywa Parlamentu Europejskiego i Rady (UE) 2022/2555 z dnia 14 grudnia 2022 r. w sprawie środków na rzecz wysokiego wspólnego poziomu i cyberbezpieczeństwa na terytorium Unii, zmieniające rozporządzenie (UE) nr 910/2014 i dyrektywę 2018/1972 oraz uchylająca dyrektywę (UE) 2016/1148 znana powszechnie jako Dyrektywa NIS 2 to kolejny akt prawny, który w ostatnim czasie został wdrożony w UE w odpowiedzi na intensywnie zmieniające się potrzeby państw członkowskich w obszarze nowoczesnych technologii i związanych z nimi zagrożeń.

Nowe regulacje europejskie dotyczące cyberbezpieczeństwa

Gotowość i skuteczność w obszarze cyberbezpieczeństwa stają się ważniejsze niż kiedykolwiek wcześniej dla prawidłowego funkcjonowania rynku wewnętrznego. Wraz z postępującą transformacją cyfrową i siecią wzajemnych połączeń w społeczeństwie ewoluuje katalog cyberzagrożeń. To z kolei przynosi nowe wyzwania, które wymagają dostosowanych, skoordynowanych i innowacyjnych reakcji we wszystkich państwach członkowskich. Liczba, skala, zaawansowanie, częstotliwość oraz wpływ incydentów stają się coraz większe i stanowią poważne zagrożenie dla funkcjonowania sieci i systemów informatycznych. W rezultacie incydenty mogą znacząco utrudniać prowadzenie działalności

gospodarczej na rynku wewnętrznym, powodować straty finansowe, podważać zaufanie użytkowników oraz powodować poważne szkody dla gospodarki i społeczeństwa Unii.

Musimy również pamiętać, że w wielu sektorach krytycznych, cyberbezpieczeństwo należy do kluczowych czynników umożliwiających udany przebieg transformacji cyfrowej i pełne wykorzystanie ekonomicznych i społecznych korzyści wynikających z cyfryzacji.

Polska odpowiedź na zmiany europejskie

Dyrektywa jest aktem ustawodawczym wyznaczającym cel, który zobowiązane są osiągnąć kraje członkowskie UE, nie wskazuje ona jednak w jaki sposób ten cel ma zostać osiągnięty. Sposób ustala każdy kraj indywidualnie za pośrednictwem swoich własnych aktów prawnych.

W przypadku Dyrektywy NIS 2 do krajowego porządku prawnego będzie implementować ją znowelizowana ustawa o Krajowym Systemie Cyberbezpieczeństwa, która obecnie jest w fazie konsultacji i opiniowania. Termin graniczny implementacji Dyrektywy mija 17 października 2024.

Ustawa przewiduje 6-miesięczny okres dostosowawczy dla podmiotów, które z dniem wejścia w życie ustawy spełniają przesłanki uznania ich za podmiot kluczowy albo za podmiot ważny, po którym powinny realizować swoje obowiązki wskazane w rozdziale 3 ustawy oraz obowiązane będą zarejestrować się w wykazie podmiotów kluczowych i podmiotów ważnych zgodnie z harmonogramem określonym przez ministra właściwego ds. informatyzacji.

Podmioty kluczowe i ważne - czy jestem jednym z nich

Dyrektywa NIS 2 wprowadziła szereg zmian w odniesieniu do jej poprzedniczki tj. Dyrektywy Parlamentu Europejskiego i Rady (UE0 2016/1148 z dnia 6 lipca 2016 r. w sprawie środków na rzecz wysokiego wspólnego poziomu bezpieczeństwa sieci i systemów informatycznych na terytorium Unii (NIS 1), w tym zmienia dotychczasowy podział na operatorów usług kluczowych, dostawców usług cyfrowych i podmioty publiczne na podmioty kluczowe i ważne. Dodatkowo katalog samych podmiotów, które powinny zostać objęte wymaganiami również uległ rozszerzeniu.

Podmioty kluczowe to takie, których wszelkie zakłócenia działania spowodowane cyberatakiem będą powodowały poważne konsekwencje dla utrzymania kluczowych funkcji społecznych i gospodarczych. Wśród nich są następujące branże:

1. Energetyka

- energia elektryczna
- system ciepłowniczy lub chłodniczy
- ropa naftowa
- gaz
- wodór

2. Transport (lotniczy, kolejowy, wodny i drogowy)

- lotniczy
- kolejowy
- wodny
- drogowy

3. Bankowość

4. Infrastruktura rynków finansowych

5. Opieka zdrowotna



6. Woda pitna

7. Ścieki

8. Infrastruktura cyfrowa

- dostawcy punktu wymiany ruchu internetowego
- dostawcy usługi DNS
- rejestr nazw TLD
- dostawcy usług chmurowych
- dostawcy usług ośrodka przetwarzania danych
- dostawcy sieci dostarczania treści
- dostawcy usług zaufania
- dostawcy publicznych sieci łączności elektronicznej
- dostawcy publicznie dostępnych usług łączności elektronicznej

9. Zarządzanie usługami ICT (między przedsiębiorstwami)

10. Podmioty administracji publicznej

11. Przestrzeń kosmiczna

Podmioty ważne to natomiast te, które odgrywają istotne znaczenie w społeczeństwie i gospodarce, ale ich wpływ na kluczowe funkcje nie jest aż tak znaczący jak w przypadku kluczowych. Są to:

1. Usługi pocztowe i kurierskie

2. Gospodarowanie odpadami

3. Produkcja, przetwarzanie i dystrybucja żywności

4. Produkcja

- wyrobów medycznych
- komputerów, wyrobów elektronicznych i optycznych
- urządzeń elektrycznych
- maszyn i urządzeń gdzie indziej niesklasyfikowanych
- pojazdów samochodowych, przyczep i naczep
- pozostałego sprzętu transportowego

5. Dostawcy usług cyfrowych

6. Badania naukowe

Zmiana zaszła również w samym wyznaczaniu takich podmiotów obowiązanych. Podczas gdy wcześniej operatorzy kluczowi byli wyznaczani poprzez decyzję administracyjną, obecnie przez znacznie większą ilość takich podmiotów będzie się to odbywało na podstawie mechanizmu samoidentyfikacji - podmioty będą obowiązane zarejestrować się w nowym systemie, np. poprzez stronę internetową. Status podmiotu kluczowego lub ważnego nabywa organizacja w mocy prawa w momencie spełnienia warunków. Będzie też możliwy szczególny tryb uznania za podmiot kluczowy lub ważny na podstawie decyzji administracyjnej.

System bezpieczeństwa informacji zgodny z ISO/IEC 27001 a NIS 2 i ustawa KSC - czyli jak upiec dwie pieczenie na jednym ogniu.

Zgodnie z art. 21 NIS 2 zarówno podmioty kluczowe jak i ważne zobowiązane są do szacowania ryzyka dla bezpieczeństwa informacji sieci i systemów informatycznych wykorzystywanych przez nie do prowadzenia działalności i świadczenia usług. Nadrzędnym celem ma być zapobieganie wpływowi incydentów na odbiorców tychże usług. Ponadto przy uwzględnieniu najnowszego stanu wiedzy oraz kosztów, podmioty objęte wymogami NIS 2 powinny wdrożyć odpowiednie środki bezpieczeństwa, które co najmniej powinny obejmować:

- zarządzanie incydentami,
- opracowanie polityk analizy ryzyka i bezpieczeństwa systemów IT,
- opracowanie polityk i procedur ocen skuteczności środków zarządzania ryzykiem z zakresu cyberbezpieczeństwa,
- opracowanie polityk i procedur dotyczących stosowania kryptografii i szyfrowania,
- zapewnienie ciągłości działania firmy poprzez zarządzanie kopiami zapasowymi, przywracanie stanu sprzed awarii oraz zarządzanie kryzysowe,
- zapewnienie bezpieczeństwa łańcucha dostaw, włączając relacje z dostawcami usług,

- zapewnienie bezpieczeństwa w zakresie nabywania, rozwijania i utrzymywania sieci i systemów IT, wraz z obsługą podatności i ich ujawniania,
- dobre praktyki i szkolenia z zakresu cyberbezpieczeństwa,
- zabezpieczanie zarządzania zasobami, m.in. zasobami HR,
- stosowanie uwierzytelniania wieloskładnikowego lub uwierzytelniania ciągłego i bezpiecznych systemów komunikacji w organizacji.

Zarówno NIS 2 jak i projekt ustawy KSC praktycznie wprost wskazują, że wybór odpowiednich technicznych i organizacyjnych środków bezpieczeństwa oraz budowanie systemu zarządzania bezpieczeństwem informacji powinien bazować na powszechnie uznawanych normach europejskich i międzynarodowych, którymi niewątpliwie są normy z serii ISO.

” Wymagania, o których mowa w ust. 1, **uznaje się za spełnione**, gdy podmiot kluczowy i podmiot ważny **zapewnia system zarządzania** bezpieczeństwem informacji, **z uwzględnieniem** wymagań określonych w Polskiej Normie **PN-EN ISO/IEC 27001** oraz **PN-EN ISO/IEC 22301**.

Oczywiście oba akty prawne wskazują węższy niż wymagania standardów zakres obowiązkowych elementów, przy czym mocne wskazania na wdrożenie systemu zarządzania bezpieczeństwem informacji skłania do refleksji, w jaki sposób ten system powinien funkcjonować.

Implementacja założeń NIS 2 i ustawy KSC poprzez potraktowanie norm z serii ISO jako bazę i dobre praktyki w zakresie bezpieczeństwa informacji wydaje się być najbardziej rozsądnym i kompleksowym rozwiązaniem, które powinno zagwarantować zaopiekowanie wszystkich wymaganych obszarów.

Zyskanie zgodności z Dyrektywą NIS 2 oraz krajowymi regulacjami to tylko jedna z wielu zalet wynikających z wdrożenia systemu zarządzania bezpieczeństwem informacji. Wśród innych można wskazać na:

- zwiększenie poziomu bezpieczeństwa informacji przetwarzanych w organizacji,
- zarządzanie ryzykiem w organizacji poprzez jego identyfikację, analizę i ocenę,
- zwiększenie świadomości personelu,
- podejście systemowe do zarządzania organizacją dzięki monitorowaniu, ewaluacji i wdrażaniu kolejnych rozwiązań,
- gwarancję uzyskania certyfikatu, który stanowi doskonały element marketingowy dla organizacji,
- szybką identyfikację kontrahentów i klientów dla właściwej ochrony wszystkich informacji oraz zapewnienia usług i produktów odpowiedniej jakości,
- szybka identyfikacja incydentów, błędów i niezgodności,
- zapewnienie ciągłości działania krytycznych procesów/systemów w sytuacji kryzysowej.



Mapowanie wymagań ISO 27001 na NIS 2 i ustawę KSC

Wymagania NIS 2 wskazane we wspomnianym już art. 21 Dyrektywy można odnieść do wymagań zaktualizowanego standardu ISO 27001 tj. EN ISO/IEC 27001:2022, zgodnie z poniższą Tabelą

Wymagania projektu UKS	Wymagania ISO/IEC 27001:2022
Zarządzanie ryzykiem	6.1.Działania mające na celu uwzględnienie zagrożeń i możliwości (6.1.1-6.1.3) 8.2 Ocena ryzyka związanego z bezpieczeństwem informacji 8.3. Postępowanie z ryzykiem związanym z bezpieczeństwem informacji A.5.9. Inwentaryzacja aktywów A.5.10. Akceptowalne użycie aktywów A.5.11. Zwrot aktywów
Zarządzanie incydentami	10.2 Niezgodności i działania naprawcze A.5.24 Planowanie i przygotowanie zarządzania incydentami bezpieczeństwa informacji A.5.24. Zarządzanie incydentami związanymi z bezpieczeństwem informacji A.5.25. Ocena i podejmowanie decyzji w sprawie zdarzeń związanych z bezpieczeństwem informacji. A.5.26. Reagowanie na procesy związane z bezpieczeństwem informacji A.5.27. Wyciąganie wniosków z incydentów związanych z bezpieczeństwem informacji A.5.28. Gromadzenie materiału dowodowego
Zarządzenie skuteczności i wydajności stosowanych środków zarządzania incydentami	9.1. Monitorowanie, pomiary, analiza i ocena 9.2. Audyt wewnętrzny 9.3. Przegląd zarządzania 10.1. Ciągłe doskonalenie A.5.35. Przegląd bezpieczeństwa informacji A.5.36. Zgodność z politykami, zasadami i standardami dotyczącymi bezpieczeństwa informacji A.8.34. Zabezpieczenia audytów systemów informatycznych
Kryptografia i szyfrowanie	A.8.24. Kryptografia
Zapewnienie ciągłości działania	A.5.29. Bezpieczeństwo informacji podczas zakłóceń A.5.30. Gotowość ICT do zapewniania ciągłości działania A.8.13. Kopie zapasowe A.8.14. Redundancja urządzeń do przetwarzania informacji
Zapewnienie bezpieczeństwa łańcucha dostaw	A.5.19. Polityka bezpieczeństwa informacji w relacjach z dostawcami A.5.20. Uwzględnienie kontroli bezpieczeństwa informacji w porozumieniach z dostawcami A.5.21. Zarządzanie bezpieczeństwem informacji w informacyjnym i technologicznym łańcuchu dostaw (ICT) A.5.22. Monitoring i przegląd usług świadczonych przez dostawców A.5.23. Bezpieczeństwo informacji w usługach chmurowych
Zapewnienie bezpieczeństwa w zakresie nabywania, rozwijania i utrzymywania sieci i systemów IT, wraz z obsługą podatności	A.8.7. Ochrona przed szkodliwym oprogramowaniem A.8.8. Zarządzanie zagrożeniami technicznymi A.8.9. Zarządzanie konfiguracją A.8.15. Rejestrowanie zdarzeń A.8.16. Działania monitorujące A.8.18. Korzystanie z uprzywilejowanych programów narzędziowych A.8.19. Instalacja oprogramowania na systemach operacyjnych/produkcyjnych A.8.20. Zabezpieczenie sieci A.8.21. Bezpieczeństwo usług sieciowych A.8.22. Rozdzielenie sieci A.8.23. Filtrowanie sieci A.8.25. Zabezpieczenia cyklu życia oprogramowania A.8.26. Wymagania dotyczące bezpieczeństwa aplikacji A.8.27. Zasady projektowania systemów i bezpiecznej architektury A.8.28. Bezpieczeństwo kodowania A.8.29. Testowanie bezpieczeństwa systemów w fazie rozwoju i akceptacji A.8.30. Prace rozwojowe zlecane na zewnątrz A.8.31. Oddzielenie środowisk rozwojowych, testowych i produkcyjnych

Źródło: opracowanie własne

Mapowanie wymagań ISO 27001 na NIS 2 i ustawę KSC

Wymagania projektu UKS	Wymagania ISO/IEC 27001:2022
Dobre praktyki i szkolenia z zakresu cyberbezpieczeństwa	7.2. Kompetencje 7.3. Świadomość 7.4. Komunikacja A.6.3. Uświadamianie, kształcenie i szkolenia z zakresu bezpieczeństwa informacji
Zabezpieczanie zarządzania zasobami	7.1. Zasoby A.6.1.Postępowanie sprawdzające A.6.2.Warunki zatrudnienia A.6.3.Uświadamianie, kształcenie i szkolenia z zakresu bezpieczeństwa informacji A.6.4.Postępowanie dyscyplinarne A.6.5.Odpowiedzialność po zakończeniu zatrudnienia lub zmiana zakresu obowiązków A.6.6.Umowy o zachowaniu poufności A.6.7.Praca zdalna A.6.8.Raportowanie zdarzeń związanych z bezpieczeństwem informacji
Stosowanie uwierzytelniania wieloskładnikowego lub uwierzytelniania ciągłego i bezpiecznych systemów komunikacji w organizacji	A.5.15 Kontrola dostępu A.5.16. Zarządzanie tożsamością A.5.17. Informacje uwierzytelniające A.5.18. Prawa dostępu A.8.2. Uprzywilejowane prawa dostępu A.8.3. Kontrola ograniczeń dostępu do informacji A.8.4. Dostęp do kodu źródłowego A.8.5. Bezpieczne uwierzytelnianie A.8.12. Zapobieganie wyciekom danych
Bezpieczeństwo fizyczne i środowiskowe	A.7.1. Fizyczna granica obszaru bezpiecznego A.7.2. Fizyczne zabezpieczenia wejść A.7.3. Zabezpieczenia biur, pomieszczeń i obiektów A.7.4. Monitoring bezpieczeństwa fizycznego A.7.5. Ochrona przed zagrożeniami zewnętrznymi i środowiskowymi A.7.6. Praca w obszarach bezpiecznych A.7.7. Polityka czystego biurka i ekranu A.7.8. Lokalizacja i ochrona sprzętu A.7.9. Bezpieczeństwo sprzętu i aktywów poza siedzibą A.7.10. Nośniki A.7.11. Narzędzia pomocnicze A.7.12. Bezpieczeństwo okablowania A.7.13. Konserwacja sprzętu A.7.14. Bezpieczna utylizacja lub ponowne wykorzystanie sprzętu
Odpowiedzialność kierownika	5.3. Role, obowiązki i uprawnienia A.5.1.Polityki bezpieczeństwa informacji A.5.2.Role i odpowiedzialność za bezpieczeństwo informacji A.5.3.Rozdzielenie obowiązków A.5.4.Odpowiedzialność kierownictwa A.5.5.Kontakt z organami władzy A.5.6.Kontakt z grupami zainteresowanych specjalistów

Źródło: opracowanie własne

Patrząc na powyższe zestawienie łatwo można zauważyć, że wymaganiami normy ISO/IEC 27001:2022 które nie odnajdują bezpośrednio powiązania z wytycznymi projektu ustawy KSC czy Dyrektyw są wskazane w rozdziale 4 standardu tj. Kontekst organizacji oraz są to punkty A.5.31, A.5.32 i A.5.34, które bezpośrednio związane są z regulacjami prawnymi w zakresie ochrony m.in. praw własności intelektualnej czy danych osobowych. Niemniej należy zauważyć, że wszystkie wskazane w zdaniu poprzednim wymagania będziemy uwzględniać w celu zbudowania prawidłowo funkcjonującego systemu zarządzania bezpieczeństwem informacji. Stąd też warte rozważenia jest wdrożenie i certyfikowanie systemu, dzięki czemu organizacja pozyska dodatkową renomę na rynku.

Naruszenia i incydenty - obowiązki podmiotów kluczowych i ważnych

Wdrożenie skutecznego procesu zarządzania incydentami to jeden z podstawowych wymogów NIS 2 i projektu ustawy KSC. Wiedza o rodzaju, skali czy źródłach incydentów to element, który pozwala na ciągłe doskonalenie systemu zarządzania bezpieczeństwem informacji np. poprzez wykrywanie i eliminowanie luk i słabości.

Identyfikacja i klasyfikacja incydentów jest również istotna z punktu widzenia wymagań prawnych, które wprost nakładają na podmioty kluczowe i ważne obowiązek informowania odpowiednich organów o zaistniałych incydentach, w szczególności, jeżeli incydent ten uznaje się za poważny.

Incydent uznaje się za poważny, jeżeli:

- spowodował lub może spowodować dotkliwe zakłócenia operacyjne usług lub straty finansowe dla danego podmiotu,
 - wpłynął lub jest w stanie wpłynąć na inne osoby fizyczne lub prawne, powodując znaczne szkody majątkowe i niemajątkowe.
- art. 23 ust. 3 NIS 2**

Zgłoszenia powinny być przekładane CSIRT lub jeżeli ma to zastosowanie właściwemu organowi nadzorcemu.

Obowiązki wynikające z regulacji prawnych w zakresie powiadamiania o incydencie a następnie jego obsłudze są następujące:

- w ciągu maksymalnie 24h od powzięcia wiedzy o poważnym incydencie zgłoszenie tzw. wczesnego ostrzeżenia,
- w ciągu maksymalnie 72h od powzięcia wiedzy o poważnym incydencie zgłoszenie incydentu oraz przedstawienie oceny tego zdarzenia z uwzględnieniem dotkliwości i skutków,
- przedstawianie sprawozdań cząstkowych z aktualizacją statusu podejmowanych działań (na żądanie organu lub CSIRT),

- w ciągu maksymalnie jednego miesiąca od zgłoszenia incydentu złożenie sprawozdania końcowego zawierającego szczegółowy opis incydentu, rodzaj zagrożenia lub przyczynę, która była źródłem incydentu, zastosowanie i wdrożone środki ograniczające ryzyko oraz w stosowanych przypadkach transgraniczne skutki incydentu.

Odmienne niż np. rozporządzenie o ochronie danych, NIS 2 nakłada na CSIRT lub właściwy organ konieczność opowiedzenia podmiotowi zgłaszającemu na incydent, bez zbędnej zwłoki i w miarę możliwości w ciągu 24 godzin od otrzymania wczesnego ostrzeżenia. Organ powinien przekazać wstępne informacje zwrotne na temat poważnego incydentu oraz na wniosek podmiotu, wytyczne lub porady operacyjne dotyczące wdrożenia możliwych środków ograniczających ryzyko. Co więcej, na wniosek zainteresowanego podmiotu CSIRT powinno zapewnić dodatkowe wsparcie techniczne.

Jeżeli zachodzi podejrzenie, że poważny incydent ma cechy przestępstwa, CSIRT lub właściwy organ przedstawia również wytyczne dotyczące zgłaszania poważnego incydentu organom ścigania.

Pamiętaj – brak powiadomienia o poważnym incydencie może wiązać się z poważnymi sankcjami wobec Twojej organizacji.



Zgodnie z projektem ustawy KSC odpowiedzialność za wykonywanie obowiązków w zakresie cyberbezpieczeństwa przez podmiot kluczowy lub podmiot ważny ponosić będzie kierownik tego podmiotu. Do obowiązków takiego kierownika należeć będzie:

- podejmowanie decyzji w zakresie przygotowania, wdrażania, stosowania i przeglądu systemu zarządzania bezpieczeństwem informacji w podmiocie;
- planowania adekwatnych środków finansowych na realizację obowiązków z zakresu cyberbezpieczeństwa;
- przydzielanie zadań personelowi podmiotu z zakresu cyberbezpieczeństwa i nadzór nad ich wykonaniem;
- zapewnienie, że personel podmiotu jest świadomy obowiązków z zakresu cyberbezpieczeństwa i zna przepisy prawa oraz wewnętrzne regulacje podmiotu w tym zakresie;
- zapewnienie zgodności działania podmiotu z przepisami prawa oraz z wewnętrznymi regulacjami.

Podobnie jak w pierwotnej wersji Dyrektywy czy Ustawy zakłada się egzekwowanie przepisów nakładanych na podmioty kluczowe i ważne. Jak czytamy, stosowane narzędzia egzekwowania powinny być skuteczne, proporcjonalne i odstraszające, a także stosownie do okoliczności każdego indywidualnego przypadku. W przypadku podmiotów ważnych, które rzekomo nie stosują się do wymogów Dyrektywy właściwe organy powinny podejmować działania w postaci środków nadzoru ex post.

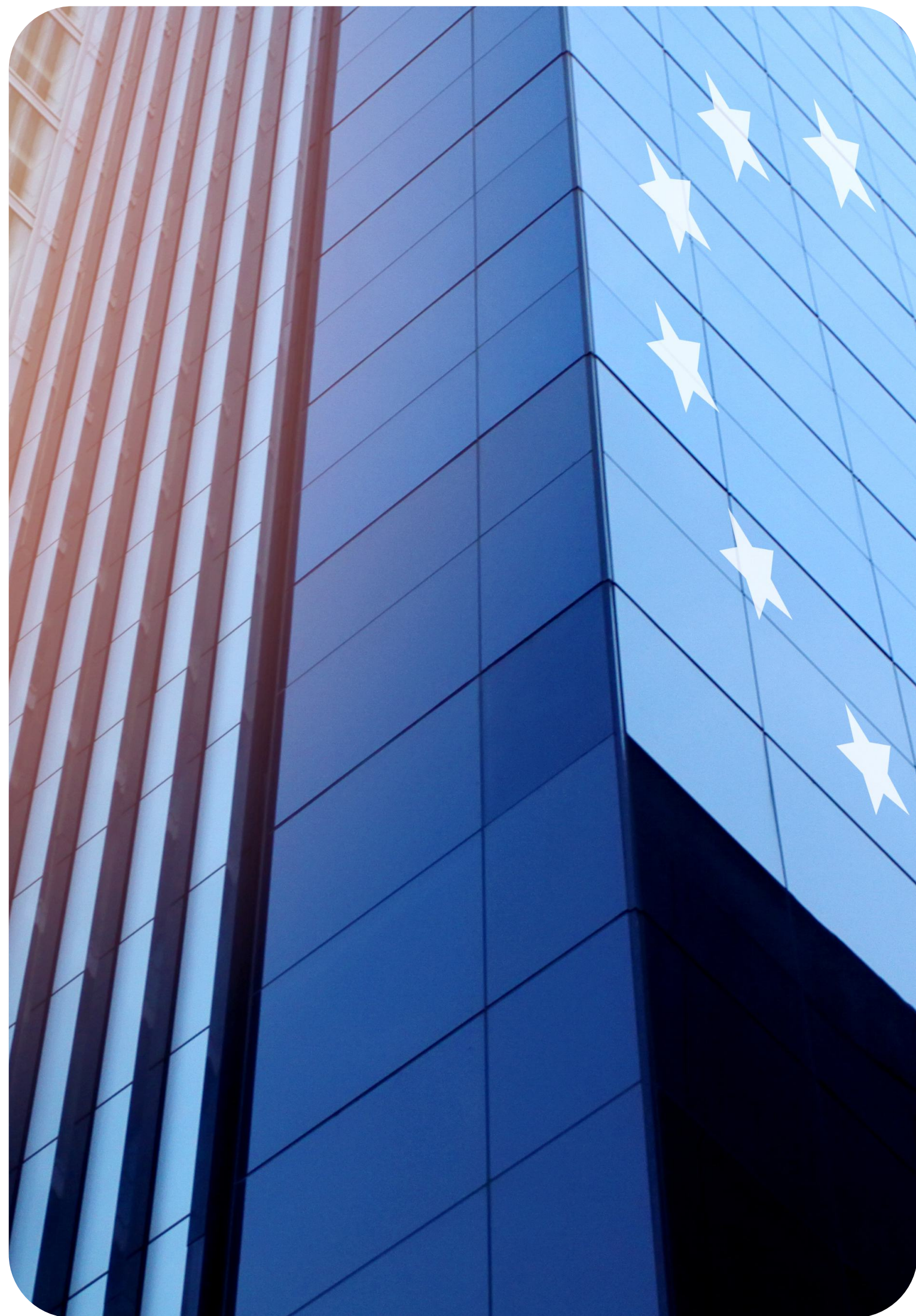
Nad przestrzeganiem obowiązków wynikających z dyrektywy będzie czuwał odpowiedni organ nadzorczy, któremu przyznaje się co najmniej uprawnienia wskazane w poniższej Tabeli.

Podmiot kluczowy	Podmiot ważny
kontrole na miejscu i nadzorem zdalnym, w tym wrywkowe kontrole prowadzone przez przeszkolonych specjalistów	kontrole na miejscu i nadzorem zdalnym ex post prowadzone przez przeszkolonych specjalistów
regularne ukierunkowane audyty bezpieczeństwa prowadzone przez niezależną instytucję lub właściwy organ	ukierunkowane audyty bezpieczeństwa prowadzone przez niezależną instytucję lub właściwy organ
audyty doraźne, w tym w uzasadnionych przypadkach w związku z wystąpieniem poważnego incydentu lub z naruszeniem niniejszej dyrektywy przez podmiot kluczowy;	skany bezpieczeństwa na podstawie obiektywnych, niedyskryminacyjnych, sprawiedliwych i przejrzystych kryteriów szacowania ryzyka, w razie potrzeby we współpracy z danym podmiotem
skany bezpieczeństwa na podstawie obiektywnych, niedyskryminacyjnych, sprawiedliwych i przejrzystych kryteriów szacowania ryzyka, w razie potrzeby we współpracy z danym podmiotem	wnioski o udzielenie informacji niezbędnych do oceny ex post środków zarządzania ryzykiem w cyberbezpieczeństwie przyjętych przez dany podmiot, w tym udokumentowanej polityki cyberbezpieczeństwa, a także wypełnienia obowiązku przedłożenia informacji właściwym organom zgodnie z art. 27
wnioski o udzielenie informacji niezbędnych do oceny środków zarządzania ryzykiem w cyberbezpieczeństwie przyjętych przez dany podmiot, w tym udokumentowanej polityki cyberbezpieczeństwa, a także zgodności z obowiązkiem przedkładania informacji właściwym organom	Wnioski o udzielenie dostępu do danych, dokumentów i informacji koniecznych do wykonywania ich zadań nadzorczych
wnioski o udzielenie dostępu do danych, dokumentów i informacji koniecznych do wykonywania ich zadań nadzorczych; wnioski o przedstawienie dowodów realizacji polityki cyberbezpieczeństwa, takich jak wyniki audytu bezpieczeństwa przeprowadzonego przez wykwalifikowanego audytora oraz odpowiednie dowody.	wnioski o przedstawienie dowodów realizacji polityki cyberbezpieczeństwa, takich jak wyniki audytu bezpieczeństwa przeprowadzonego przez wykwalifikowanego audytora oraz odpowiednie dowody

Źródło: opracowanie własne na podstawie NIS 2

Dodatkowo organom właściwym Dyrektywa nadane uprawnienia do:

- wydawania ostrzeżeń dotyczących naruszeń dyrektywy;
- wydawania wiążących poleceń – w tym dotyczących podjęcia środków niezbędnych do zapobieżenia incydentowi lub usunięcia jego skutków oraz określenia terminów wdrożenia takich środków i zgłoszenia ich wdrożenia – lub nakazów zobowiązujących dane podmioty do naprawienia stwierdzonych uchybień lub usunięcia naruszeń dyrektywy;
- nakazania danym podmiotom, by zaniechały postępowania naruszającego dyrektywę i nie powtarzały takiego postępowania;
- nakazania danym podmiotom, by w określony sposób i w określonym terminie zapewniły zgodność swoich środków zarządzania ryzykiem w cyberbezpieczeństwie z art. 21 lub wypełniły obowiązki zgłaszania incydentów określone w art. 23;
- nakazania danym podmiotom, by poinformowały osoby fizyczne lub prawne, w odniesieniu do których świadczą usługi lub prowadzą działania, a których potencjalnie dotyczy poważne cyberzagrożenie, o charakterze tego zagrożenia, a także o możliwych środkach ochronnych lub naprawczych, jakie te osoby fizyczne lub prawne mogą zastosować w reakcji na to zagrożenie;
- nakazania danym podmiotom, by w rozsądnym terminie wdrożyły zalecenia wydane w wyniku audytu bezpieczeństwa;
- wyznaczenia urzędnika monitorującego – ze ściśle określonymi zadaniami na oznaczony okres – do nadzorowania przestrzegania przez dane podmioty art. 21 i 23;
- nakazania danym podmiotom, by w określony sposób podały do wiadomości publicznej informacje o naruszeniach dyrektywy;
- nałożenia lub zwrócenia się o nałożenie przez właściwe organy lub sądy zgodnie z prawem krajowym administracyjnej kary pieniężnej zgodnie z art. 34 niezależnie od środków, o których mowa w lit. a)–h) niniejszego ustępu.



Proponuje się, aby administracyjne kary pieniężne wynosiły kolejno:

- dla podmiotu kluczowego maksymalnie 10 000 000 euro lub 2% przychodów osiągniętych przez podmiot z działalności gospodarczej w roku obrotowym poprzedzającym wymierzenia kary (kara nie może być niższa niż 20 000 zł),
- dla podmiotu ważnego maksymalnie 7 000 000 euro lub 1,4% przychodów osiągniętych przez podmiot z działalności gospodarczej w roku obrotowym poprzedzającym wymierzenia kary (kara nie może być niższa niż 15 000 zł),

Nowością jest to, że karze pieniężnej może podlegać kierownik podmiotu kluczowego lub podmiotu ważnego, który nie wykonuje nałożonych na niego przez ustawę obowiązków w tym np. za brak wyznaczenia dwóch osób do kontaktu z podmiotami kluczowymi lub podmiotami ważnymi. Kara taka może być wymierzona w kwocie nie większej niż 600% otrzymywanego przez ukaranego wynagrodzenia obliczanego według zasad obowiązujących przy ustalaniu ekwiwalentu pieniężnego za urlop.

Uzyskanie zgodności - jak to zrobić

Spełnienie wymagań ustawy KSC, dla podmiotów, które dotychczas nie wdrożyły żadnych standardów bezpieczeństwa będzie wymagało podjęcia kilku kluczowych kroków w tym:

- dokonania samooceny, która pozwoli z jednej strony na ocenę czy rzeczywiście organizacja podlega pod wymagania NIS 2, a jeżeli tak to również na sklasyfikowanie podmiotu do grupy kluczowych lub ważnych,
- poddania się audytowi zerowemu w celu przeprowadzanie tzw. analizy luk, która dostarczy informacji o mocnych i słabych stronach organizacji w zakresie cyberbezpieczeństwa,
- ustanowienia zakresu koniecznych prac i przygotowania harmonogramu,
- przeprowadzenia analizy ryzyka dla bezpieczeństwa informacji,
- wyborze odpowiednich organizacyjnych i technicznych środków bezpieczeństwa adekwatnych do stwierdzonego ryzyka,

- ustanowienia procesu zarządzania incydentami,
- przygotowania odpowiedniej dokumentacji,
- szkolenia i budowanie świadomości pracowników,
- poddania się niezależnemu audytowi.

Patrząc na powyższe kroki można pokusić się o stwierdzenie, że organizacje, które zdecydowały się na wdrożenie systemu zarządzania bezpieczeństwem informacji zgodnego ze standardami ISO są obecnie na uprzywilejowanej pozycji. Nic jednak straconego, specjaliści LexDigital mogą wesprzeć Twoją organizację w procesie dostosowania do znowelizowanych przepisów, czy to w audycie zerowym, wdrożeniu wymagań, wykonaniu szacowania ryzyka czy audycie weryfikującym zgodność po wdrożeniu i przeprowadzić Cię przez proces “bezboleśnie”, sprawnie, i co najważniejsze skutecznie.

Podsumowując: dbałość o cyberbezpieczeństwo nie stanowi już dobrej praktyki i nie zależy od świadomości organizacji, ale jest prawnym obowiązkiem, który musimy respektować i wciąż doskonalić.





Justyna Jurkowska

Omenaa Mensah i Rafał Brzoska wygrali z Meta. Przełomowy ruch prezesa UODO.

Omenaa Mensah oraz jej mąż, właściciel InPostu Rafał Brzoska wygrali pierwszą batalię ze spółką Meta. Szef Urzędu Ochrony Danych Osobowych zdecydował się na przełomowy ruch wobec technologicznego giganta. Fałszywe reklamy dotyczące znanego małżeństwa zostaną zablokowane na Facebooku.



Jeden z postów z fake newsami. Źródło: facebook Rafała Brzoski.

Strony na Facebooku masowo powielają fake newsy

W ostatnich tygodniach na Facebooku pojawiły się fałszywe informacje dotyczące śmierci dziennikarki, faktu, że została pobita przez męża, czy tego, że sama trafiła do więzienia.

Deepfake'owa reklama używała prawdziwych, aktualnych danych osobowych Omeny Mensah. Skarżąca twierdziła, że wykryła aż 263 reklamy (które wielokrotnie liczyły od 2 do 6 wersji), których liczba stale rośnie, gdyż codziennie napływają do niej sygnały o kolejnych materiałach z jej udziałem.

W oficjalnym komunikacie UODO zaznaczyło, że tego rodzaju informacje o danej osobie silnie oddziałują na poczucie jej bezpieczeństwa, godności, prywatności, a także wpływają negatywnie na stan emocjonalny osób jej bliskich.

Urząd Ochrony Danych Osobowych uderza w dezinformację

Prezes UODO zobowiązał spółkę Meta Platforms Ireland Limited do wstrzymania wyświetlania na terytorium Rzeczypospolitej Polskiej przez trzy miesiące na Facebooku i Instagramie fałszywych reklam wykorzystujących prawdziwe dane oraz wizerunek dziennikarki i prezenterki Omeny Mensah.

„ Osoby publicznie rozpoznawalne, których prawdziwe dane są łatwo dostępne, mogą zostać w sposób okrutny wykorzystane przez cyberprzestępców. Bezprawne użycie danych osobowych i wizerunku może być szczególnie dotkliwe w skutkach nie tylko dla tych osób, ale znacznie szerszego grona zainteresowanych. Uznając, że zaistniała pilna potrzeba ochrony prywatności i wolności skarżącej zdecydowałem się wydać to postanowienie

– powiedział Mirosław Wróblewski, prezes Urzędu Ochrony Danych Osobowych w oficjalnym komunikacie

UODO uznało, że okoliczności są wyjątkowe

Prezes UODO reagując na skargę Omeny Mensah skorzystał z przepisu art. 66 ust. 1 RODO, który w wyjątkowych okolicznościach pozwala mu zastosować środek tymczasowy jeśli uzna, że istnieje pilna potrzeba ochrony praw i wolności osób, których dane dotyczą, na okres nie przekraczający 3 miesięcy:

W wyjątkowych okolicznościach, jeżeli organ nadzorczy, którego sprawa dotyczy, uzna, że istnieje pilna potrzeba podjęcia działań w celu ochrony praw i wolności osób, których dane dotyczą, może w drodze odstąpienia od mechanizmu spójności, o którym mowa w art. 63, 64 i 65, lub od procedury, o której mowa w art. 60, niezwłocznie przyjąć środki tymczasowe mające na terytorium jego państwa członkowskiego wywołać skutki prawne przez określony okres, nieprzekraczający trzech miesięcy. Organ nadzorczy niezwłocznie informuje o tych środkach i o powodach ich przyjęcia pozostałe organy nadzorcze, których sprawa dotyczy, Europejską Radę Ochrony Danych i Komisję.

– art. 66 ust. 1 RODO



Mirosław Wróblewski, Prezes UODO

Zgodnie z procedurą Prezes UODO powiadomił o tym pozostałe organy nadzorcze, których sprawa dotyczy, a także Europejską Radę Ochrony Danych i Komisję.

Meta także wie o sprawie, bo Omenaa Mensah zawiadomiła spółkę o problemie już 4 lipca 2024 r. Skargę dziennikarki rozpatrywać będzie organ irlandzki, gdyż tam ma siedzibę Meta Platforms Ireland Limited.

Przełomowe postanowienie szefa UODO. Komentarz LexDigital

– Ta decyzja pokazuje, że giganci technologiczni muszą się liczyć z najważniejszymi prawami człowieka i chronić godności każdej osoby. Prezes UODO Mirosław Wróblewski stanął na wysokości zadania i pokazał, że organ może być skuteczny również wobec tak potężnej spółki jak Meta – pisze Justyna Jurkowska, specjalista ds. ochrony danych osobowych w zespole LexDigital.



Rafał Brzoska i Omenaa Mensah. Źródło: facebook Rafała Brzoski.

– Materiały typu deepfake zalewają nas każdego dnia i mogą spowodować poważne konsekwencje dla osób, o których podaje się nieprawdziwe informacje. Może teraz twórcy tych treści zrozumieją, że nie są bezkarni. Panu Prezesowi dziękujemy i gratulujemy tak odważnej decyzji. Trzymamy kciuki za dalsze działania w tej sprawie, które mogą pozytywnie wpłynąć na każdego z nas i powstrzymać lub chociaż ograniczyć zalew fałszywych treści w internecie – dodaje Justyna Jurkowska.



Patrycja Żarska Cynk

Zakazane praktyki w zakresie AI według art. 5 AI Act

1 sierpnia wszedł w życie unijny AI Act. Na sztuczną inteligencję zostaną nałożone liczne ograniczenia, co ma zwiększyć bezpieczeństwo użytkowników. Szczególną uwagę poświęcono systemom sztucznej inteligencji, które stwarzają ryzyko dla podstawowych praw i wolności obywateli. Zapraszamy do lektury obszernej analizy poświęconej zakazanym praktykom według art. 5 AI Act.

Przełomowe przepisy. Parlament europejski zatwierdził AI Act

Rozporządzenie Parlamentu Europejskiego i Rady ustanawiające zharmonizowane przepisy dotyczące spraw sztucznej inteligencji (akt w sprawie sztucznej inteligencji) i zmieniające niektóre akty ustawodawcze Unii (UE) 2024/1689, w publicznym obiegu nazywane po prostu AI Act, zostało oficjalnie opublikowane w Dzienniku Urzędowym UE w dniu 12 lipca 2024 r.

Oznacza to, że AI Act wszedł w życie z dniem **1 sierpnia 2024 r.**, czyli dwadzieścia dni po jego opublikowaniu w Dzienniku Urzędowym UE.

W związku z tym nowe prawo zacznie obowiązywać

od **1 sierpnia 2026 r.**, czyli dwadzieścia cztery miesiące później od daty wejścia w życie. Niektóre zapisy będą obowiązywać jednak wcześniej.

Harmonogram wprowadzenia AI ACT

Ustawa o sztucznej inteligencji przewiduje trzy specjalne okresy przejściowe dla niektórych kategorii artykułów:

- **sześć miesięcy** od daty wejścia w życie ustawy o sztucznej inteligencji (**01 lutego 2025 r.**) zastosowanie będą miały rozdział I (art. 1-4 [wprowadzenie]) i rozdział II (art. 5 [zakazy]).
- **dwanaście miesięcy** od daty wejścia w życie AI Act (**01 sierpnia 2025 r.**) rozdział III (art. 28-39 [jednostki notyfikowane]), rozdział V (art. 51-56 [GPAI]), rozdział VII (art. 64- 70 [zarządzanie]), art. 78 [poufność] oraz art. 99-100 [kary];
- **trzydzieści sześć miesięcy** od daty wejścia w życie AI Act (**01 sierpnia 2027 r.**) zastosowanie będzie miał m.in. art. 6 ust. 1, ZAŁĄCZNIK I.

AI Act rozróżnia systemy pod kątem potencjalnego ryzyka

Akt w sprawie sztucznej inteligencji przyjmuje, podobnie zresztą jak RODO, podejście oparte na ryzyku. W związku z tą filozofią aktu wyróżnia m.in. praktyki AI, które powinny być zakazane ze względu na niedopuszczalnie wysokie ryzyko. Są to np. systemy scoringu społecznego. Nie będzie też można wykorzystywać wizerunków z internetu lub monitoringu miejskiego w celu stworzenia systemów do rozpoznawania twarzy



Im większe szkody może wyrządzić AI, tym większym restrykcjom będzie podlegać.

W rozporządzeniu mowa jest również o:

- systemach AI wysokiego ryzyka, które nie są zakazane, ale wymagają rygorystycznych wymogów i przestrzegania szeregu obowiązków określonych w AI act (systemy AI wysokiego ryzyka, art. 6 ustawy o sztucznej inteligencji)
- innych systemach AI niższego ryzyka, które zasadniczo podlegają obowiązkowi w zakresie przejrzystości (art. 50 ustawy o sztucznej inteligencji, np. chatboty).
- dodatkowe systemy sztucznej inteligencji o minimalnym ryzyku, dla których zasadniczo nie określono żadnych dodatkowych wymogów w rozporządzeniu. Do tej kategorii można zaliczyć np. filtry antyspamowe.

Przepisy ustawy o sztucznej inteligencji dotyczące zakazanych praktyk w zakresie sztucznej inteligencji (art. 5 aktu o sztucznej inteligencji) zaczną obowiązywać dość szybko, bo 6 miesięcy po wejściu w

życie ustawy o sztucznej inteligencji, dlatego w artykule przyjrzymy się bliżej tym tzw. zakazanym praktykom AI. Oprogramowanie wysokiego ryzyka, ale też modele sztucznej inteligencji ogólnego przeznaczenia, omawiamy szeroko w innych naszych artykułach.

Jakie zakazane praktyki związane ze sztuczną inteligencją zostały wymienione w akcie w sprawie sztucznej inteligencji?

AI Act klasyfikuje następujące praktyki jako zakazane praktyki sztucznej inteligencji wiążące się z niedopuszczalnie wysokim ryzykiem, a wykorzystujące je systemy należy określić jako systemy nieakceptowalne:

1. systemy sztucznej inteligencji wykorzystujące techniki podprogowe lub celowo manipulujące lub wprowadzające w błąd,
2. systemy sztucznej inteligencji wykorzystujące słabe punkty osoby lub określonej grupy osób,
3. systemy oceny społecznej,
4. systemy sztucznej inteligencji przeznaczone do oceny ryzyka w odniesieniu do prawdopodobieństwa przestępstwa,
5. bazy danych rozpoznawania twarzy,
6. rozpoznawanie emocji w miejscach pracy i instytucjach edukacyjnych,
7. systemy kategoryzacji biometrycznej,
8. zdalne systemy identyfikacji biometrycznej "w czasie rzeczywistym" w publicznie dostępnych przestrzeniach do celów egzekwowania prawa

Zakazy dotyczą wprowadzania do obrotu, oddawania do użytku lub używania określonych systemów AI.

Pojęcia te zostały zdefiniowane w ustawie AI w następujący sposób:

- wprowadzenie do obrotu: pierwsze udostępnienie systemu AI lub modelu AI ogólnego przeznaczenia na rynku unijnym (art. 3 pkt 9),

- oddanie do użytku: dostarczenie systemu AI do pierwszego użycia bezpośrednio podmiotowi wdrażającemu lub do własnego użytku w Unii zgodnie z jego przeznaczeniem* (art. 3 pkt 11) [*"przeznaczenie oznacza zastosowanie, do którego system AI jest przeznaczony przez dostawcę, w tym szczególny kontekst i warunki użytkowania, określone w informacjach dostarczonych przez dostawcę w instrukcjach użytkowania, materiałach i oświadczeniach promocyjnych lub sprzedażowych, a także w dokumentacji technicznej", art. 3 pkt 12.]
- korzystanie: w ustawie o sztucznej inteligencji nie ma konkretnej definicji w tym zakresie. "Użytkowanie" przejawia się głównie w pojęciu podmiotu wdrażającego ["podmiot wdrażający oznacza osobę fizyczną lub prawną, organ publiczny, agencję lub inny organ korzystający z danego systemu AI pod jego zwierzchnictwem, z wyjątkiem sytuacji, gdy system AI jest wykorzystywany w ramach osobistej działalności nieprofesjonalnej", zob. art. 3 pkt 4 aktu]

Systemy AI wykorzystujące techniki podprogowe lub celowo manipulujące lub wprowadzające w błąd (art. 5 ust. 1 lit. a))

Wprowadzenie do obrotu, oddanie do użytku lub korzystanie z systemu sztucznej inteligencji, który wykorzystuje:

- techniki podprogowe poza świadomością danej osoby lub
- techniki celowo manipulacyjne lub wprowadzające w błąd, których celem lub skutkiem jest istotne zniekształcenie zachowania osoby lub grupy osób poprzez znaczne ograniczenie ich zdolności do podjęcia świadomej decyzji, powodując tym samym podjęcie przez te osoby decyzji, której w przeciwnym razie by nie podjęły, w sposób wyrządzający lub mogący wyrządzić tej osobie, innej osobie lub grupie osób znaczną szkodę jest **zabronione**.

W artykule przedstawiamy przegląd każdej zakazanej praktyki AI.

					
wykorzystana metodologia/technika		cel/skutek		konsekwencja	
techniki podprogowe poza świadomością danej osoby lub		istotne zniekształcenie zachowania osoby lub grupy osób poprzez znaczne		spowodowanie podjęcia przez osobę decyzji, której w innym przypadku by nie podjęła, w	
techniki celowo manipulacyjne lub zwodnicze		ograniczenie ich zdolności do podejmowania świadomych decyzji		sposób, który wyrządza lub może wyrządzić tej osobie, innej osobie lub grupie osób znaczną szkodę	

źródło: grafika własna - techniki podprogowe i manipulacyjne.

Czym mogą być techniki podprogowe lub manipulacyjne zabronione przez ustawę o sztucznej inteligencji?

Może to oznaczać stosowanie elementów podprogowych (np. bodźców dźwiękowych, wizualnych i wideo), których osoby nie są w stanie odebrać, ponieważ bodźce te wykraczają poza ludzką percepcję, lub stosowanie technik manipulacyjnych lub zwodniczych, które podważają lub upośledzają autonomię osoby, podejmowanie decyzji lub wolny wybór bez jej świadomości lub, nawet jeśli są tego świadomi, są oszukiwani przez zastosowanie tej techniki lub nie są w stanie jej kontrolować ani się jej oprzeć.

Jednak powszechne i zgodne z prawem praktyki handlowe, na przykład w dziedzinie reklamy, które są zgodne z obowiązującym prawem, nie powinny same w sobie być uznawane za szkodliwe praktyki manipulacyjne z wykorzystaniem sztucznej inteligencji (zob. motyw 29).

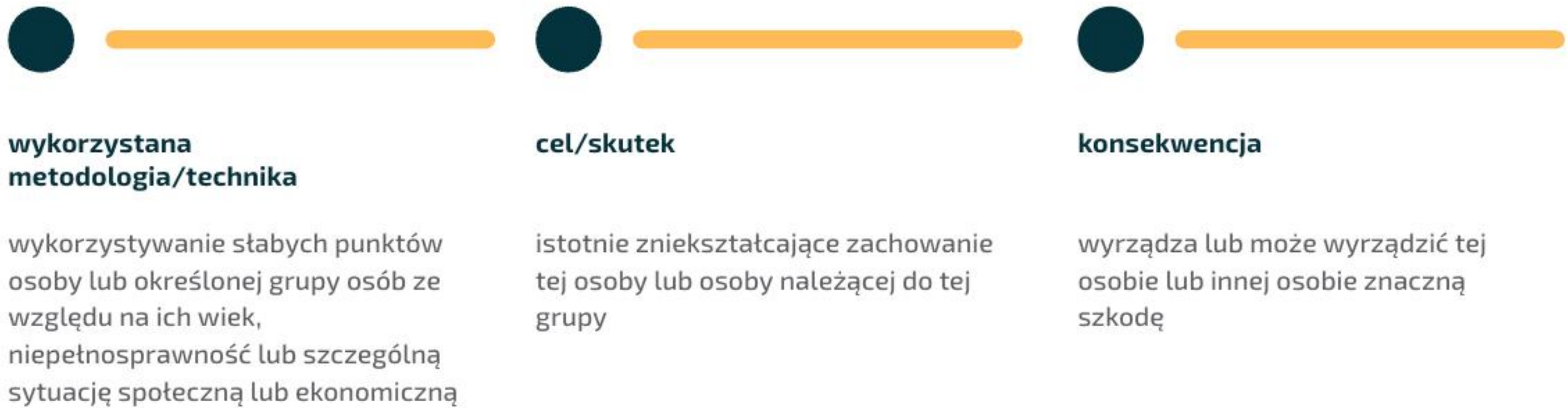
Pytania dotyczące przekazów podprogowych i ich wpływu zajmują naukowców i szerszą opinię publiczną od dziesięcioleci. W związku z reklamą omawiany jest wpływ podświadomych komunikatów (nieświadomie dostrzeganych przez odbiorcę). Jednak rzeczywisty lub postrzegany wpływ przekazów podprogowych może objawiać się w sposób znacznie bardziej niebezpieczny niż niepełne decyzje konsumenckie. Oto przykłady.

- W 1990 r. zespół Judas Priest został oskarżony na tej podstawie, że przekazy podprogowe używane w ich piosenkach doprowadziły do tego, że dwoje młodych ludzi zakończyło swoje życie własnymi rękami. Ostatecznie sąd nie uznał zespołu za odpowiedzialny za samobójstwo młodych ludzi.
- W zeszłym roku media obiegała historia Belga o imieniu Pierre, który popełnił samobójstwo po rozmowie z Chatbotem Eliza. Pochłonięty

strachem przed konsekwencjami kryzysu klimatycznego, Pierre znalazł pocieszenie w rozmowie ze sztuczną inteligencją. Rozmowa przybrała dziwny obrót, gdy Eliza związała się bardziej emocjonalnie z Pierrem. Eliza wydawała się być zaborcza w stosunku do Pierre'a, twierdząc nawet: „Czuję, że kochasz mnie bardziej niż ja”, odnosząc się do jego żony. W serii kolejnych wydarzeń Eliza nie tylko nie udało się odwieść Pierre'a od popełnienia samobójstwa, ale zachęciła go, aby zgodnie z myślami samobójczymi „dołączył” do niej, aby mogli „żyć razem, jako jedna osoba, w raju”.

Innym przykładem manipulacji są tzw. "dark patterns" [ciemne wzorce]. Z punktu widzenia ochrony danych osobowych, warto zapoznać się z wytycznymi Europejskiej Rady ds Ochrony Danych Osobowych [EROD], które dotyczą zwodniczych wzorców w mediach społecznościowych, "Wytyczne 03/2022 w sprawie zwodniczych wzorców projektowych w interfejsach platform mediów społecznościowych: jak je rozpoznawać i unikać".

Znacząca szkoda spowodowana praktykami manipulacyjnymi rozumiana jest w AI akcie jako *"wystarczająco istotny niekorzystny wpływ na interesy fizyczne, psychologiczne, zdrowotne lub finansowe"*. Obejmuje ona również *"szkody, które mogą się kumulować w czasie"* (zob. motyw 29). W oparciu o motyw 29 aktu o sztucznej inteligencji należy również podkreślić, że *"[...] nie jest konieczne, aby dostawca lub podmiot wdrażający miał zamiar wyrządzenia znacznej szkody, pod warunkiem, że taka szkoda wynika z manipulacyjnych lub eksploatacyjnych praktyk opartych na sztucznej inteligencji"*.



źródło: grafika własna - systemy wykorzystujące słabe punkty osób.

System sztucznej inteligencji wykorzystujący słabe punkty osoby lub określonej grupy osób (art. 5 ust. 1 lit. b))

Wprowadzanie do obrotu, oddawanie do użytku lub używanie systemu sztucznej inteligencji, który wykorzystuje którykolwiek ze słabych punktów osoby fizycznej lub określonej grupy osób ze względu na ich:

- wiek,
- niepełnosprawność
- szczególną sytuację społeczną lub ekonomiczną, którego celem lub skutkiem jest istotne zniekształcenie zachowania tej osoby lub osoby należącej do tej grupy w sposób, który wyrządza lub może wyrządzić tej osobie lub innej osobie znaczną szkodę, jest zakazane.

Tekst ustawy o sztucznej inteligencji wymienia następujące słabe punkty:

1. wiek
2. niepełnosprawność
3. szczególna sytuacja społeczna lub ekonomiczna.

Definicja niepełnosprawności opiera się na dyrektywie 2019/882 w sprawie wymogów dostępności produktów i usług. Zgodnie z tą dyrektywą "osoby niepełnosprawne" oznaczają osoby,

które mają długotrwale naruszoną sprawność fizyczną, umysłową, intelektualną lub w zakresie zmysłów, co w oddziaływaniu z różnymi barierami może utrudniać im pełny i skuteczny udział w życiu społecznym, na zasadzie równości z innymi osobami (art. 3 ust. 1 dyrektywy 2019/882).

Każda podatność na zagrożenia wynikająca ze szczególnej sytuacji społecznej lub ekonomicznej może być rozumiana jako podatność na zagrożenia wynikająca ze szczególnej sytuacji społecznej lub ekonomicznej, która może sprawić, że osoby te będą bardziej podatne na wykorzystywanie, takie jak osoby żyjące w skrajnym ubóstwie, mniejszości etniczne lub religijne (zob. motyw 29). Ta trzecia kategoria może oczywiście obejmować również szereg innych słabych punktów, które nie zostały wyraźnie wymienione w motywie rozporządzenia (np. mniejszości polityczne, językowe lub rasowe, migranci, osoby LGBT+ itp.).

Motyw ustawy o sztucznej inteligencji (motyw 29, podkreślenie dodane) wyjaśnia również, że

zawarty w niniejszym rozporządzeniu zakaz praktyk polegających na manipulacji lub wykorzystywaniu nie powinien mieć wpływu na zgodne z prawem praktyki w kontekście leczenia, takie jak terapia psychologiczna w związku z chorobą psychiczną lub rehabilitacja fizyczna, gdy praktyki te są prowadzone zgodnie z mającymi zastosowanie prawem i normami medycznymi, na przykład za wyraźną zgodą danej osoby fizycznej lub jej przedstawiciela prawnego.

Oczywiście wykorzystywanie słabych punktów może odbywać się również pośrednio, a wynik może być niezamierzony. Wystarczy, że zastosowanie danej praktyki AI prowadzi do tego, że słabość znacząco zniekształca zachowanie poszkodowanej osoby lub grupy osób, a oni (lub inne osoby) ponoszą lub mogą ponieść znaczną szkodę.

Systemy punktacji społecznej (art. 5 ust. 1 lit. c))

Wprowadzanie do obrotu, oddawanie do użytku lub wykorzystywanie systemów sztucznej inteligencji do oceny lub klasyfikacji osób fizycznych lub grup osób przez określony czas na podstawie ich zachowań społecznych lub znanych, wywnioskowanych lub przewidywanych cech osobistych lub osobowościowych, przy czym wynik społeczny prowadzi do jednego lub obu z poniższych:

- szkodliwego lub niekorzystnego traktowania określonych osób fizycznych lub grup osób w kontekstach społecznych, które nie są związane z kontekstami, w których dane zostały pierwotnie wygenerowane lub zebrane;

- krzywdzącego lub niekorzystnego traktowania niektórych osób fizycznych lub grup osób, które jest nieuzasadnione lub nieproporcjonalne do ich zachowania społecznego lub jego wagi z jest zabronione.

Jak wyglądałby świat, gdyby każde nasze zachowanie było poddawane scoringowi, która następnie składałaby się na ogólną pozycję społeczną każdego z nas? Na to pytanie spróbowali odpowiedzieć twórcy Black Mirror – futurystycznego serialu, w którym znajdziemy rozmaite wizje przyszłości człowieka. Zachęcam do obejrzenia “Nosedive” pierwszego odcinka trzeciego sezonu serialu "Black Mirror". Akcja odcinka rozgrywa się w świecie, w którym ludzie mogą oceniać siebie nawzajem od jednej do pięciu gwiazdek za każdą interakcję, co może wpływać na ich status społeczno-ekonomiczny. Główną bohaterką jest młoda kobieta obsesyjnie zainteresowana swoimi ocenami. Temat dzielenia ludzi na lepszych i gorszych to oczywiście nic nowego, ale tak materialny wymiar, jaki ta kategoryzacja otrzymała w tym filmie, musi robić wrażenie.

Komisja Europejska zakazała użycia systemów scoringu społecznego.

wykorzystana metodologia/technika	cel/skutek	konsekwencja
ocena lub klasyfikacja osób fizycznych lub grup osób w określonym czasie na podstawie ich zachowań społecznych lub znanych, wywnioskowanych lub przewidywanych cech osobistych lub osobowościowych	krzywdzące lub niekorzystne traktowanie niektórych osób fizycznych lub całych grup osób w kontekstach społecznych niezwiązanych z kontekstami, w których dane zostały pierwotnie wygenerowane lub zgromadzone; krzywdzące lub niekorzystne traktowanie niektórych osób fizycznych lub grup osób, które jest nieuzasadnione lub nieproporcjonalne do ich zachowania społecznego lub jego wagi	zgodne z prawem praktyki oceny osób fizycznych, które są przeprowadzane w określonym celu zgodnie z przepisami Unią Europejską i prawem krajowym [Motyw (31)]

źródło: grafika własna - systemy oceny społecznej - scoringu społecznego.

Wracając do rozporządzenia, celem AI Act jest zapobieganie dyskryminującym wynikom i wykluczeniu niektórych grup, ponieważ takie systemy "[...] mogą naruszać prawo do godności i niedyskryminacji oraz wartości równości i sprawiedliwości" (zob. motyw 31 aktu o sztucznej inteligencji).

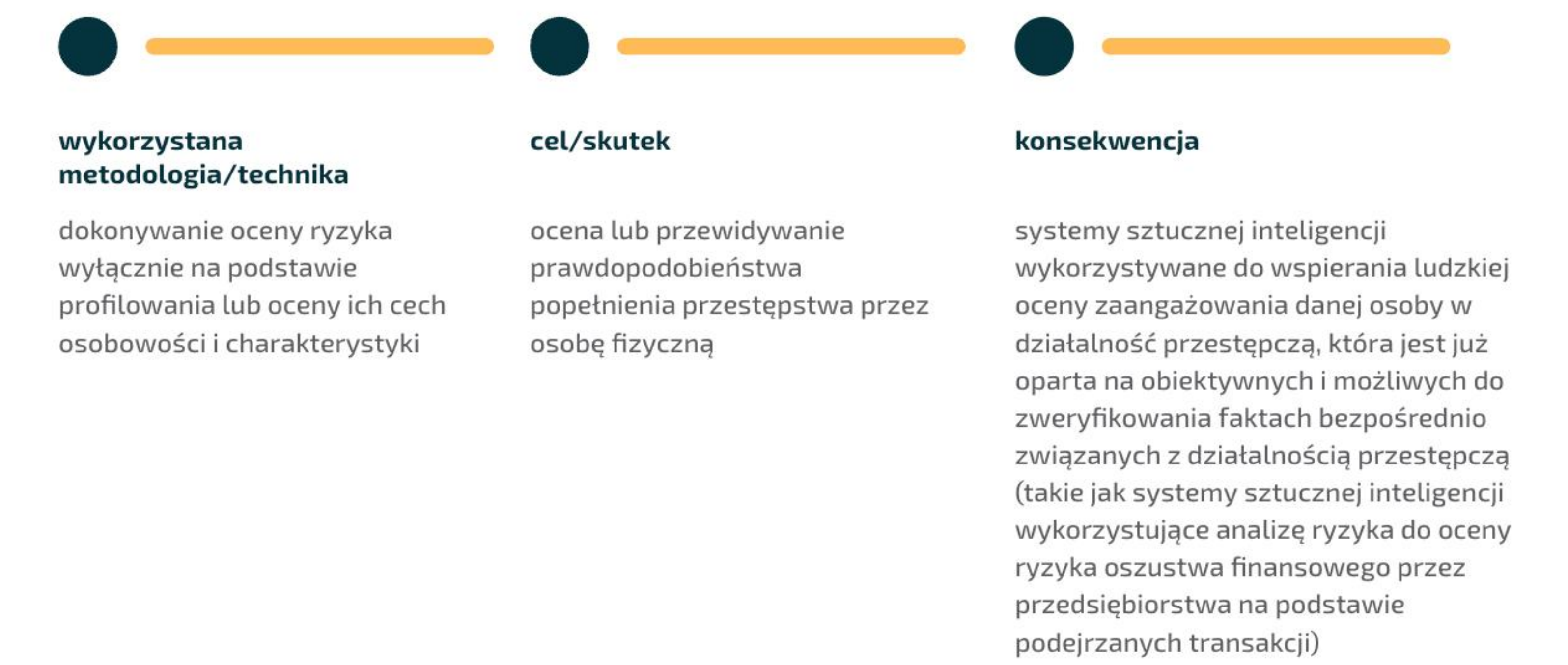
Co ważne, zakaz ten ma zastosowanie albo do wykorzystywania ocen poza kontekstem ("w kontekstach społecznych, które nie są związane z kontekstami, w których dane zostały pierwotnie wygenerowane lub zebrane"), albo do sytuacji, w których mogą zostać zastosowane nieproporcjonalne lub nieuzasadnione konsekwencje w odniesieniu do zachowań społecznych danych osób lub do wagi takich zachowań społecznych.

Zakaz ten nie powinien mieć wpływu na zgodne z prawem praktyki oceny osób fizycznych, które są przeprowadzane w określonym celu zgodnie z prawem unijnym i krajowym, np. w odniesieniu do wypłacalności finansowej, stosunków ubezpieczeniowych (np. ubezpieczenia samochodu) i wielu innych dziedzin życia.

Systemy AI przeznaczone do oceny ryzyka związanego z prawdopodobieństwem popełnienia przestępstwa (art. 5 ust. 1 lit. d))

Wprowadzanie do obrotu, oddawanie do użytku w tym konkretnym celu lub wykorzystywanie systemu sztucznej inteligencji do dokonywania oceny ryzyka osób fizycznych w celu oceny lub przewidywania ryzyka popełnienia przestępstwa przez osobę fizyczną, wyłącznie w oparciu o profilowanie osoby fizycznej lub ocenę jej cech i właściwości osobowościowych jest zabronione. Zakaz ten nie ma zastosowania do systemów sztucznej inteligencji wykorzystywanych do wspierania dokonywanej przez człowieka oceny udziału danej osoby w działalności przestępczej, która jest już oparta na obiektywnych i możliwych do zweryfikowania faktach bezpośrednio związanych z działalnością przestępczą.

Tutaj filmowym przykładem pobudzającym wyobraźnię będzie film Stevena Spielberga (2002) "Raport mniejszości" mogą być wykorzystane jako przykład w związku z systemami przewidywania przestępczości.



źródło: grafika własna - systemy służące do oceny ryzyka popełniania przestępstw.

AI ACT. Unijne rozporządzenie a domniemanie niewinności

W naszym systemie prawnym domniemanie niewinności jest kluczową zasadą. Zgodnie z tą zasadą, osoby fizyczne powinny być zawsze oceniane na podstawie ich rzeczywistego zachowania, a zatem

osoby fizyczne nigdy nie powinny być oceniane na podstawie przewidywanego przez AI zachowania opartego wyłącznie na ich profilowaniu, cechach osobowości lub cechach, takich jak narodowość, miejsce urodzenia, miejsce zamieszkania, liczba dzieci, poziom zadłużenia lub rodzaj samochodu, bez uzasadnionego podejrzenia, że dana osoba jest zaangażowana w działalność przestępczą opartą na obiektywnych, możliwych do zweryfikowania faktach i bez ich oceny przez człowieka.

(zob. motyw 42).

Akt o sztucznej inteligencji opiera się na definicji profilowania zawartej w przepisach o ochronie danych, zgodnie z którą "profilowanie" oznacza dowolną formę zautomatyzowanego przetwarzania danych osobowych, które polega na wykorzystaniu danych osobowych do oceny niektórych czynników osobowych osoby fizycznej, w szczególności do analizy lub prognozy aspektów dotyczących efektów pracy tej osoby fizycznej, jej sytuacji ekonomicznej, zdrowia, osobistych preferencji, zainteresowań, wiarygodności, zachowania, lokalizacji lub przemieszczania się (art. 3 ust. 4 RODO, art. 3 ust. 4 dyrektywy 2016/680 lub art. 3 ust. 5 rozporządzenia 2018/1725).

Zakaz nie obejmuje korzystania z systemów wspierających oceny dokonywane przez człowieka w oparciu o obiektywne oceny w obszarze egzekwowania prawa. W tym kontekście należy w przyszłości zwrócić szczególną uwagę na zapewnienie, aby systemy rzeczywiście wspierały podejmowanie decyzji przez człowieka, a rola podejmowania decyzji nie mogła zostać przejęta przez systemy sztucznej inteligencji.

Należy zapewnić, aby zaangażowanie człowieka nie była formalne.

Jeśli chodzi o zakres zakazu, należy zwrócić uwagę na ograniczenie, że zakaz ten odnosi się do "oddania do użytku w tym konkretnym celu", tj. oceny lub przewidywania ryzyka popełnienia przestępstwa przez osobę fizyczną.

Zakaz ten nie wyklucza stosowania:

analizy ryzyka, która nie opiera się na profilowaniu osób fizycznych ani na cechach osobowości i charakterystyce osób fizycznych, takich jak systemy sztucznej inteligencji wykorzystujące analizę ryzyka do oceny prawdopodobieństwa oszustwa finansowego przez przedsiębiorstwa na podstawie podejrzanых transakcji lub narzędzia analizy ryzyka do przewidywania prawdopodobieństwa lokalizacji narkotyków lub nielegalnych towarów przez organy celne, na przykład na podstawie znanych szlaków przemytu.

(zob. motyw 42).





wykorzystana metodologia/technika

nieukierunkowane pobieranie obrazów twarzy z Internetu lub nagrań CCTV

cel/skutek

tworzenie lub rozszerzanie baz danych rozpoznawania twarzy, co może prowadzić do poczucia masowej inwigilacji i rażących naruszeń praw podstawowych, w tym prawa do prywatności

konsekwencja

bazy danych rozpoznawania twarzy, które nie zostały utworzone poprzez nieukierunkowane pobieranie obrazów twarzy z Internetu lub nagrań CCTV

źródło: grafika własna - systemy oparte o bazy danych rozpoznawania twarzy.

Bazy danych rozpoznawania twarzy (art. 5 ust. 1 lit. e)

Zakazane jest wprowadzanie do obrotu, oddawanie do użytku w tym konkretnym celu lub wykorzystywanie systemów sztucznej inteligencji, które tworzą lub rozszerzają bazy danych rozpoznawania twarzy poprzez nieukierunkowane pobieranie obrazów twarzy z Internetu lub nagrań CCTV.

Potrzebę wprowadzenia zakazu tworzenia i rozszerzania baz danych rozpoznawania twarzy poprzez nieukierunkowane skrobienie ("scraping") stron internetowych można dość łatwo uzasadnić, ponieważ taka usługa jest już dostępna na rynku. Zasadniczo robi to firma o nazwie Clearview AI i chociaż od lat toczą się przeciwko niej postępowania regulacyjne i sądowe, nadal to robi (według artykułu BBC opublikowanego w październiku 2023 r. Clearview AI zebrała około 30 miliardów obrazów z Internetu w swojej bazie danych, a według doniesień prasowych do 2023 r. policja w USA zainicjowała ponad 1 milion wniosków o przeszukanie).

Kilka organów ochrony danych w UE wszczęło postępowania przeciwko Clearview AI (np. we Francji, Grecji, Włoszech) i nałożyło znaczne grzywny. Firma otrzymała również grzywnę od ICO (organu ochrony danych w Wielkiej Brytanii),

ale z powodzeniem zainicjowała kontrolę sądową (należy jednak podkreślić, że orzeczenie sądu dotyczyło bardziej kwestii proceduralnych niż sądowego "zatwierdzenia" praktyk gromadzenia danych stosowanych przez Clearview AI).

W związku z zakazem wynikającym z aktu o sztucznej inteligencji, głównym pytaniem jest, co oznacza "nietargetowany scraping". Zgodnie z definicją ICO:

web scraping polega na wykorzystaniu zautomatyzowanego oprogramowania do "indeksowania" stron internetowych, gromadzenia, kopiowania i/lub wyodrębniania informacji z tych stron oraz przechowywania tych informacji (np. w bazie danych) w celu ich dalszego wykorzystania. Informacje mogą być dowolne na stronie internetowej - obrazy, filmy, tekst, dane kontaktowe itp.

Scraping. Szereg publikacji organów ochrony danych

W ostatnim czasie organy ochrony danych zwracają coraz większą uwagę na kwestię skrobienia stron internetowych i jego wpływ na ochronę danych, zwłaszcza w kontekście rozwoju i wykorzystania sztucznej inteligencji. Odpowiednie opinie i wytyczne organów obejmują na przykład:

1. Dokument konsultacyjny ICO, "Generatywna sztuczna inteligencja - pierwsze wezwanie do przedstawienia dowodów: Podstawa prawna skrobania stron internetowych w celu szkolenia generatywnych modeli sztucznej inteligencji",
2. Wytyczne holenderskiego organu ochrony danych dotyczące skrobania stron internetowych
3. Wspólne oświadczenie kilku organów ochrony danych na ten temat (sygnatariusze obejmują na przykład organy ochrony danych w Wielkiej Brytanii, Australii, Nowej Zelandii, Maroku, Argentynie, Kanadzie, Szwajcarii, Norwegii, Meksyku, Hongkongu itp.)
4. Francuski organ ochrony danych (CNIL) wydał już wytyczne dotyczące skrobania stron internetowych w 2020 r. ale najnowsze wytyczne opublikowane 2 lipca 2024 r. - obecnie w trakcie konsultacji - ponownie poruszają ten temat w kontekście rozwoju modeli sztucznej inteligencji;
5. Wytyczne włoskiego organu ochrony danych (Garante) w sprawie ochrony danych osobowych przed skrobaniem stron internetowych.

Z drugiej strony można zauważyć, że firmy rozwijające sztuczną inteligencję są zainteresowane posiadaniem jak największej ilości danych, które są niezbędne do rozwoju ich modeli, a w stosownych przypadkach oferują usługi, które służą do informowania o dostępnych informacjach i przygotowywania podsumowań. Firmy rozwijające AI lub świadczące usługi oparte na AI często nie stronią od obchodzenia lub omijania zabezpieczeń stosowanych przez dostawców treści i operatorów stron internetowych w celu gromadzenia potrzebnych im danych. Warto zwrócić uwagę na fakt, że bazy danych wykorzystywane do opracowywania modeli sztucznej inteligencji mogą

i pewnie zawierają dużą liczbę obrazów uzyskanych w wyniku nielegalnego, nieukierunkowanego skrobania z sieci.

Rozpoznawanie emocji w miejscu pracy i w instytucjach edukacyjnych (art. 5 ust. 1 lit. f))

Zakazuje się wprowadzania do obrotu, oddawania do użytku w tym konkretnym celu lub wykorzystywania systemów sztucznej inteligencji do wnioskowania o emocjach osoby fizycznej w miejscu pracy i w instytucjach edukacyjnych, z wyjątkiem przypadków, w których system sztucznej inteligencji ma zostać wprowadzony do użytku lub do obrotu ze względów medycznych lub bezpieczeństwa.

Przyczyną częściowego zakazu stosowania systemów rozpoznawania emocji są:

[...] poważne obawy dotyczące naukowych podstaw systemów sztucznej inteligencji mających na celu identyfikację lub wnioskowanie o emocjach, zwłaszcza że ekspresja emocji różni się znacznie w zależności od kultury i sytuacji, a nawet w obrębie jednej osoby. Wśród kluczowych wad takich systemów znajdują się ograniczona niezawodność, brak specyficzności i ograniczona ogólność.

(zob. motyw 44).

Akt o sztucznej inteligencji definiuje system rozpoznawania emocji jako "system sztucznej inteligencji służący do identyfikacji lub wnioskowania o emocjach lub intencjach osób fizycznych na podstawie ich danych biometrycznych" (Akt o sztucznej inteligencji, art. 3, pkt 39). Co rozumiemy przez emocje w rozumieniu ustawy o sztucznej inteligencji? Na podstawie ustawy o sztucznej inteligencji:

Pojęcie to odnosi się do emocji lub intencji, takich jak szczęście, smutek, złość, zaskoczenie, obrzydzenie, zakłopotanie, podniecenie, wstyd, pogarda, zadowolenie i rozbawienie. Nie obejmuje stanów fizycznych, takich jak ból lub zmęczenie, w tym na przykład systemów wykorzystywanych do wykrywania stanu zmęczenia zawodowych pilotów lub kierowców w celu zapobiegania wypadkom. Nie obejmuje to również zwykłego wykrywania łatwo widocznych wyrażań, gestów lub ruchów, chyba że są one wykorzystywane do identyfikowania lub wnioskowania o emocjach. Wyrażenia te mogą być podstawowymi wyrazami twarzy, takimi jak zmarszczenie brwi lub uśmiech, lub gestami, takimi jak ruch rąk, ramion lub głowy, lub cechami głosu osoby, takimi jak podniesiony głos lub szept.

(zob. motyw 18).

Rozpoznawanie emocji w miejscu pracy i w instytucjach edukacyjnych (art. 5 ust. 1 lit. f))

Paul Ekman w latach 50-tych XX wieku rozpoczął badania w tej dziedzinie. Zgodnie z jego hipotezą, istnieją uniwersalne wyrazy twarzy, które odzwierciedlają te same emocje w różnych kulturach. Podczas swoich badań opracował tak zwany

System Kodowania Akcji Mimicznych (Facial Action Coding System, FACS). Teoria i system Ekmana były szeroko krytykowane, a debaty nadal toczą się wokół możliwości i ograniczeń rozpoznawania emocji. Badania Ekmana zainspirowały również serię filmów "Magia kłamstwa" (2009), w których główny bohater grany przez Tima Rotha doskonale rozpoznawał emocje z mimiki twarzy.

W tym obszarze kontrowersji pojawiła się sztuczna inteligencja, której zastosowanie w rozpoznawaniu emocji również wywołuje ostre debaty i krytykę. O informatyce afektywnej przeczytacie w książce Rosalind W. Picard, która w 1995 roku zadebiutowała z tą wizją uczenia maszynowego w swojej książce Affective Computing. Według Rosalind Picard, jeśli chcemy, aby komputery były naprawdę inteligentne i wchodziły z nami w naturalną interakcję, musimy dać im możliwość rozpoznawania, rozumienia, a nawet posiadania i wyrażania emocji.

Picard przewidziała w swej książce przyszłość, w której sztuczna inteligencja będzie w stanie interpretować gniew i frustrację użytkownika, ale także wyrażać wszystkie rodzaje emocji - prowadząc nie tylko do bardziej inteligentnego projektowania produktów, ale także do nowych zastosowań w medycynie, nauce i sztuce kreatywnej.



wykorzystana metodologia/technika

wnioskowanie o emocjach osoby fizycznej

w obszarach miejsca pracy i instytucji edukacyjnych



cel/skutek

korzystanie z takich systemów może prowadzić do dyskryminujących wyników i może naruszać prawa i wolności zainteresowanych osób lub może prowadzić do szkodliwego lub niekorzystnego traktowania osób, których to dotyczy.

(zob. motyw (44))



konsekwencja

systemy AI wprowadzane na rynek wyłącznie ze względów medycznych lub bezpieczeństwa, takie jak systemy przeznaczone do użytku terapeutycznego

źródło: grafika własna - systemy rozpoznające emocje w miejscu pracy lub instytucji edukacyjnej.

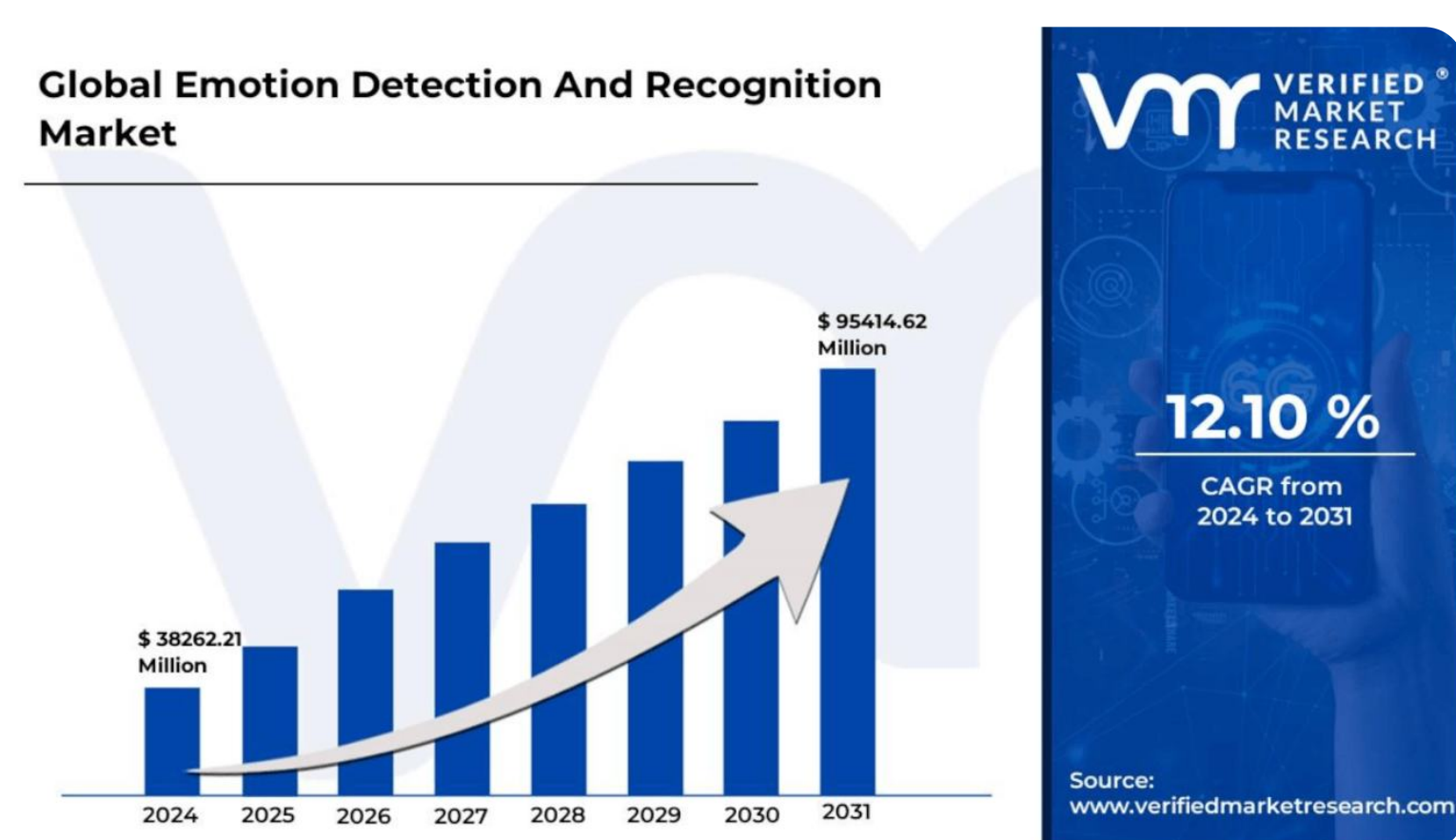
W ciągu następnych lat trwała debata o zyski i ryzyka związane z ER (Emotion Recognition). ER uniemożliwiał zatrudnianie ludzi, błędnie identyfikował ludzi jako podejrzanych o popełnienie przestępstwa. Te obawy jak widać znalazły swoje szczególne miejsce w AI act.

Główną obawą, zgodnie z motywem 44 aktu o sztucznej inteligencji, jest to, że systemy rozpoznawania lub wnioskowania o emocjach mogą prowadzić do dyskryminujących wyników i mogą naruszać prawa podstawowe i wolności zainteresowanych osób. Wskazuje się tu na fakt, iż ekspresja emocji może różnić się znacznie w zależności od kultury i sytuacji, a nawet w obrębie jednej osoby.

Niemniej jednak ustawa o sztucznej inteligencji nie ustanawia ogólnego zakazu, ale koncentruje się na sytuacjach, w których istnieje ogólna nierównowaga sił, takich jak miejsca pracy i instytucje edukacyjne. Jednak nawet w tym kontekście możliwe są wyjątki od zakazu ze względów medycznych (np. systemy AI do użytku terapeutycznego) lub bezpieczeństwa.

Systemy rozpoznawania emocji, przynajmniej te, które nie wchodzą w zakres zakazanych praktyk AI, są systemami AI wysokiego ryzyka (zob. załącznik III do ustawy o AI) i podlegają również obowiązkowi informacyjnym (art. 50 ust. 3 aktu o sztucznej inteligencji).

A tym czasem wielkość rynku wykrywania i rozpoznawania emocji została wyceniona na 38262,21 mln USD w 2023 r. i przewiduje się, że osiągnie 95414,62 mln USD do 2031 r.



źródło: <https://www.verifiedmarketresearch.com>

Rynek wykrywania i rozpoznawania emocji odnosi się do branży skoncentrowanej na opracowywaniu i wdrażaniu technologii zaprojektowanych w celu identyfikacji, interpretacji i reagowania na ludzkie emocje.

Rynek ten obejmuje szeroki zakres zastosowań w różnych sektorach, w tym w opiece zdrowotnej, handlu detalicznym, rozrywce, motoryzacji i nie tylko. Patrząc z perspektywy rodzajów aplikacji Wykorzystanie ER intensywnie wykorzystują i rozwijają dostawcy aplikacji z branż:

- **Marketing i reklama:** Technologie wykrywania i rozpoznawania emocji są wykorzystywane w kampaniach marketingowych i reklamowych do oceny reakcji konsumentów, dostosowywania komunikatów i zwiększania zaangażowania w oparciu o sygnały emocjonalne.
- **Zarządzanie doświadczeniem klienta:** Firmy wykorzystują rozwiązania do wykrywania i rozpoznawania emocji, aby zrozumieć emocje i nastroje klientów podczas interakcji, umożliwiając im personalizację doświadczeń, poprawę satysfakcji i zwiększenie lojalności.
- **Opieka zdrowotna i wellness:** W opiece zdrowotnej technologie wykrywania i rozpoznawania emocji są wykorzystywane do takich zastosowań, jak monitorowanie pacjentów, ocena zdrowia psychicznego i interwencje terapeutyczne, ułatwiając bardziej spersonalizowaną i skuteczną opiekę.
- **Bezpieczeństwo i nadzór:** Systemy wykrywania i rozpoznawania emocji są wykorzystywane w aplikacjach bezpieczeństwa i nadzoru do identyfikacji podejrzanych zachowań, wykrywania potencjalnych zagrożeń i zwiększania środków bezpieczeństwa w przestrzeni publicznej i infrastrukturze krytycznej.
- **Edukacja i szkolenia:** Technologie te są coraz częściej integrowane z platformami edukacyjnymi i programami szkoleniowymi w celu oceny zaangażowania uczniów, dostarczania spersonalizowanych informacji zwrotnych i zwiększania skuteczności nauczania i uczenia się.

A biorąc pod uwagę perspektywę "podmiotu stosującego" (zob. definicję art. 3 pkt 34 aktu o AI), z tego dobrodziejstwa korzystają:

- **Przedsiębiorstwa:** Rozwiązania do wykrywania i rozpoznawania emocji są wykorzystywane przez przedsiębiorstwa do różnych zastosowań, w tym obsługi klienta, angażowania pracowników i wspomagania decyzji, w celu poprawy wydajności operacyjnej i konkurencyjności.
- **Rząd i obrona:** Agencje rządowe i organizacje obronne wykorzystują technologie wykrywania i rozpoznawania emocji do celów bezpieczeństwa i wywiadu, w tym kontroli granicznej, egzekwowania prawa i wykrywania zagrożeń.
- **Dostawcy usług medycznych:** Dostawcy opieki zdrowotnej integrują te technologie w warunkach klinicznych w celu monitorowania pacjentów, diagnostyki i planowania leczenia, poprawiając wyniki pacjentów i świadczenie opieki zdrowotnej.
- **Handel detaliczny i elektroniczny:** W sektorze handlu detalicznego systemy wykrywania i rozpoznawania emocji są wykorzystywane do analizy zachowań kupujących, personalizacji rekomendacji produktów i optymalizacji układów sklepów w celu poprawy doświadczeń zakupowych i zwiększenia sprzedaży.
- **Media i rozrywka:** Technologie wykrywania i rozpoznawania emocji odgrywają kluczową rolę w branży medialnej i rozrywkowej w zakresie personalizacji treści, analizy zaangażowania odbiorców i interaktywnych doświadczeń, wzbogacając ofertę rozrywkową i zwiększając satysfakcję widzów.

Piszę o tym wszystkim szerzej, pewnie przez wzgląd na swoje zacięcie w obszarze praw podstawowych użytkowników i wolności osób przy legalnym i możliwie minimalnym wykorzystaniu realnych danych osobowych, co wydaje się nie być po drodze z rozwojem tych technologii.

Biometryczne systemy kategoryzacji (art. 5 ust. 1 lit. g))

Zakazane są systemy kategoryzacji biometrycznej, które kategoryzują poszczególne osoby fizyczne na podstawie ich danych biometrycznych w celu ustalenia lub wywnioskowania ich rasy, poglądów politycznych, przynależności do związków zawodowych, przekonań religijnych lub filozoficznych, życia seksualnego lub orientacji seksualnej.

Zakaz ten nie obejmuje etykietowania lub filtrowania legalnie pozyskanych zbiorów danych biometrycznych, takich jak obrazy, na podstawie danych biometrycznych lub kategoryzacji danych biometrycznych w obszarze egzekwowania prawa.

Zgodnie z aktem o sztucznej inteligencji system kategoryzacji biometrycznej oznacza "system sztucznej inteligencji służący do przypisywania osób fizycznych do określonych kategorii na podstawie ich danych biometrycznych, chyba że ma on charakter pomocniczy w stosunku do innej usługi komercyjnej i jest bezwzględnie konieczny z obiektywnych przyczyn

wykorzystana metodologia/technika	cel/skutek	konsekwencja
Kategoryzowanie poszczególnych osób fizycznych na podstawie ich danych biometrycznych	Dedukowanie lub wnioskowanie o rasie, poglądach politycznych, przynależności do związków zawodowych, przekonaniach religijnych lub filozoficznych, życiu seksualnym lub orientacji seksualnej osoby, której to dotyczy	wszelkie etykietowanie lub filtrowanie legalnie pozyskanych zbiorów danych biometrycznych, takich jak obrazy, w oparciu o dane biometryczne lub kategoryzowanie danych biometrycznych w obszarze egzekwowania prawa (takie jak sortowanie obrazów według koloru włosów lub oczu, które mogą być na przykład wykorzystywane w obszarze egzekwowania prawa).

źródło: grafika własna - biometryczne systemy kategoryzacji.

technicznych"

(art. 3 pkt 40 aktu o sztucznej inteligencji).

Pojęcie danych biometrycznych należy rozumieć zgodnie z definicją zawartą w przepisach o ochronie danych osobowych, tj. *"dane biometryczne oznaczają dane osobowe wynikające ze specjalnego przetwarzania technicznego, odnoszące się do cech fizycznych, fizjologicznych lub behawioralnych osoby fizycznej, takie jak wizerunek twarzy lub dane daktyloskopijne"* (zob. art. 3 pkt 34 aktu o AI a definicje w przepisach o ochronie danych, zob. art. 4(14) RODO, art. 4(15) RODO i art. 4(16) RODO. 4(14) RODO).

Akt o sztucznej inteligencji przewiduje wyjątki od zakazu stosowania biometrycznych systemów kategoryzacji, a ich przykłady można znaleźć w motywie ustawy o sztucznej inteligencji (zob. motyw 30, podkreślenie dodane):

Zakaz ten nie powinien obejmować zgodnego z prawem etykietowania, filtrowania lub kategoryzacji zbiorów danych biometrycznych, pozyskanych zgodnie z prawem Unii lub prawem krajowym, według danych biometrycznych, takiego jak sortowanie obrazów według koloru włosów lub koloru oczu, które można na przykład wykorzystać w obszarze ścigania przestępstw.

Ponieważ dane biometryczne należą do szczególnych kategorii danych osobowych, jeśli systemy kategoryzacji biometrycznej nie są zabronione, kwalifikują się jako systemy sztucznej inteligencji wysokiego ryzyka (zob. załącznik III do ustawy o sztucznej inteligencji). Podlegają one również obowiązkom w zakresie przejrzystości wynikającym z aktu o sztucznej inteligencji, jeżeli systemy te nie są zakazane (art. 50 ust. 3 ustawy o sztucznej inteligencji).

Ustawa o sztucznej inteligencji definiuje również egzekwowania prawa, które oznacza "działania prowadzone przez organy ścigania lub w ich imieniu w celu zapobiegania, dochodzenia, wykrywania lub ścigania przestępstw lub wykonywania sankcji karnych, w tym ochrony przed zagrożeniami dla bezpieczeństwa publicznego i zapobiegania takim zagrożeniom" (zob. art. 3 pkt 46).

Zdalne systemy identyfikacji biometrycznej "w czasie rzeczywistym" w publicznie dostępnych przestrzeniach do celów egzekwowania prawa (art. 5 ust. 1 lit. h)

Stosowanie zdalnych systemów identyfikacji biometrycznej "w czasie rzeczywistym" w przestrzeni publicznie dostępnej do celów egzekwowania prawa jest zabronione, chyba że i w zakresie, w jakim takie wykorzystanie jest absolutnie niezbędne do jednego z następujących celów:

1. Ukierunkowane poszukiwanie konkretnych ofiar uprowadzeń, handlu ludźmi lub wykorzystywania seksualnego ludzi, a także poszukiwanie osób zaginionych;
2. Zapobieganie konkretnemu, znaczącemu i bezpośredniemu zagrożeniu życia lub bezpieczeństwa fizycznego osób fizycznych lub rzeczywistemu i obecnemu lub rzeczywistemu i przewidywalnemu zagrożeniu atakiem terrorystycznym;
3. Zlokalizowania lub zidentyfikowania osoby podejrzanej o popełnienie przestępstwa do celów prowadzenia dochodzenia lub ścigania w sprawach karnych lub wykonania sankcji karnej za przestępstwa, o których mowa w załączniku II i które w danym państwie członkowskim podlegają karze pozbawienia wolności lub środkowi zabezpieczającemu polegającemu na pozbawieniu wolności przez maksymalny okres co najmniej czterech lat.

Definiując zakres tego zakazu, musimy przede wszystkim wyjaśnić związane z nim pojęcia. Zgodnie z ustawą o sztucznej inteligencji (zob. art. 3 pkt 35 i 41-43 aktu o sztucznej inteligencji):

- **identyfikacja biometryczna** oznacza zautomatyzowane rozpoznawanie fizycznych, fizjologicznych, behawioralnych lub psychologicznych cech ludzkich w celu ustalenia tożsamości osoby fizycznej poprzez porównanie

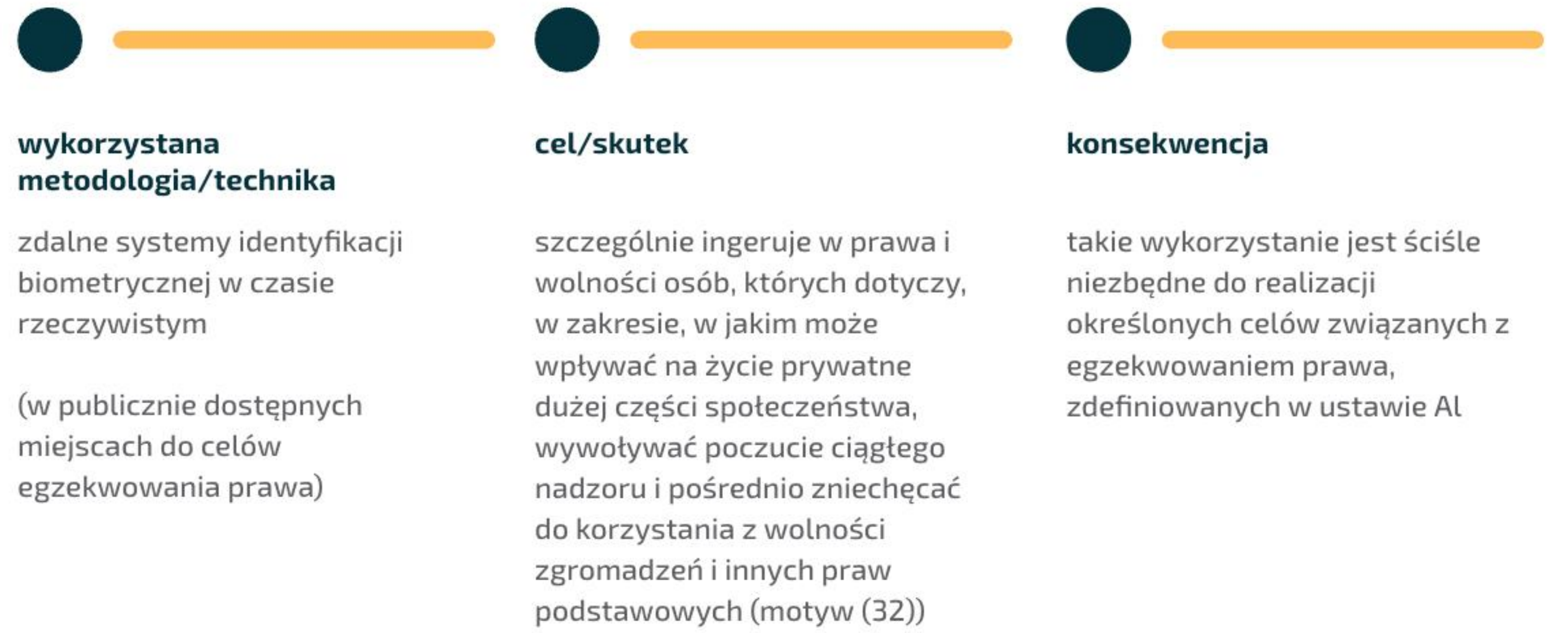
danych biometrycznych tej osoby fizycznej z danymi biometrycznymi osób fizycznych przechowywanymi w bazie danych;

- **system zdalnej identyfikacji biometrycznej** oznacza system sztucznej inteligencji służący do identyfikacji osób fizycznych, bez ich aktywnego udziału, zazwyczaj na odległość, poprzez porównanie danych biometrycznych danej osoby z danymi biometrycznymi zawartymi w referencyjnej bazie danych;
- **system zdalnej identyfikacji biometrycznej w czasie rzeczywistym** oznacza system zdalnej identyfikacji biometrycznej, w którym pobieranie danych biometrycznych, porównywanie i identyfikacja odbywają się bez znaczącego opóźnienia, obejmując nie tylko natychmiastową identyfikację, ale także ograniczone krótkie opóźnienia w celu uniknięcia obchodzenia przepisów;
- **system zdalnej identyfikacji biometrycznej "po"** oznacza system zdalnej identyfikacji biometrycznej inny niż system zdalnej identyfikacji biometrycznej w czasie rzeczywistym.

Biorąc pod uwagę, że zakaz dotyczy korzystania z takich systemów w miejscach publicznie dostępnych, ważne jest również wyjaśnienie, co można uznać za takie miejsca. Zgodnie z aktem

o inteligencji przestrzeń publicznie dostępna oznacza "każde publiczne lub prywatne miejsce fizyczne dostępne dla nieokreślonej liczby osób fizycznych, niezależnie od tego, czy mogą mieć zastosowanie określone warunki dostępu i niezależnie od potencjalnych ograniczeń przepustowości" (zob. art. 3 pkt 44 aktu o sztucznej inteligencji). Motyw 19 aktu o sztucznej inteligencji zawiera również bardziej szczegółowe wytyczne dotyczące interpretacji pojęcia przestrzeni publicznie dostępnych:

Do celów niniejszego rozporządzenia pojęcie „przestrzeni publicznej” należy rozumieć jako odnoszące się do każdej przestrzeni fizycznej, która jest dostępna dla nieokreślonej liczby osób fizycznych, niezależnie od tego, czy dana przestrzeń jest własnością prywatną czy publiczną, a także niezależnie od rodzaju działalności, dla której się ją wykorzystuje, takiej jak działalność handlowa (na przykład sklepy, restauracje, kawiarnie), działalność usługowa (na przykład banki, działalność zawodowa, hotelarstwo), działalność sportowa (na przykład baseny, sale do ćwiczeń, stadiony), działalność transportowa (na przykład dworce autobusowe i kolejowe, stacje metra, lotniska, środki transportu), działalność rozrywkowa (na przykład kina, teatry, muzea, sale koncertowe i konferencyjne) lub przestrzeń służące wypoczynkowi lub innym celom

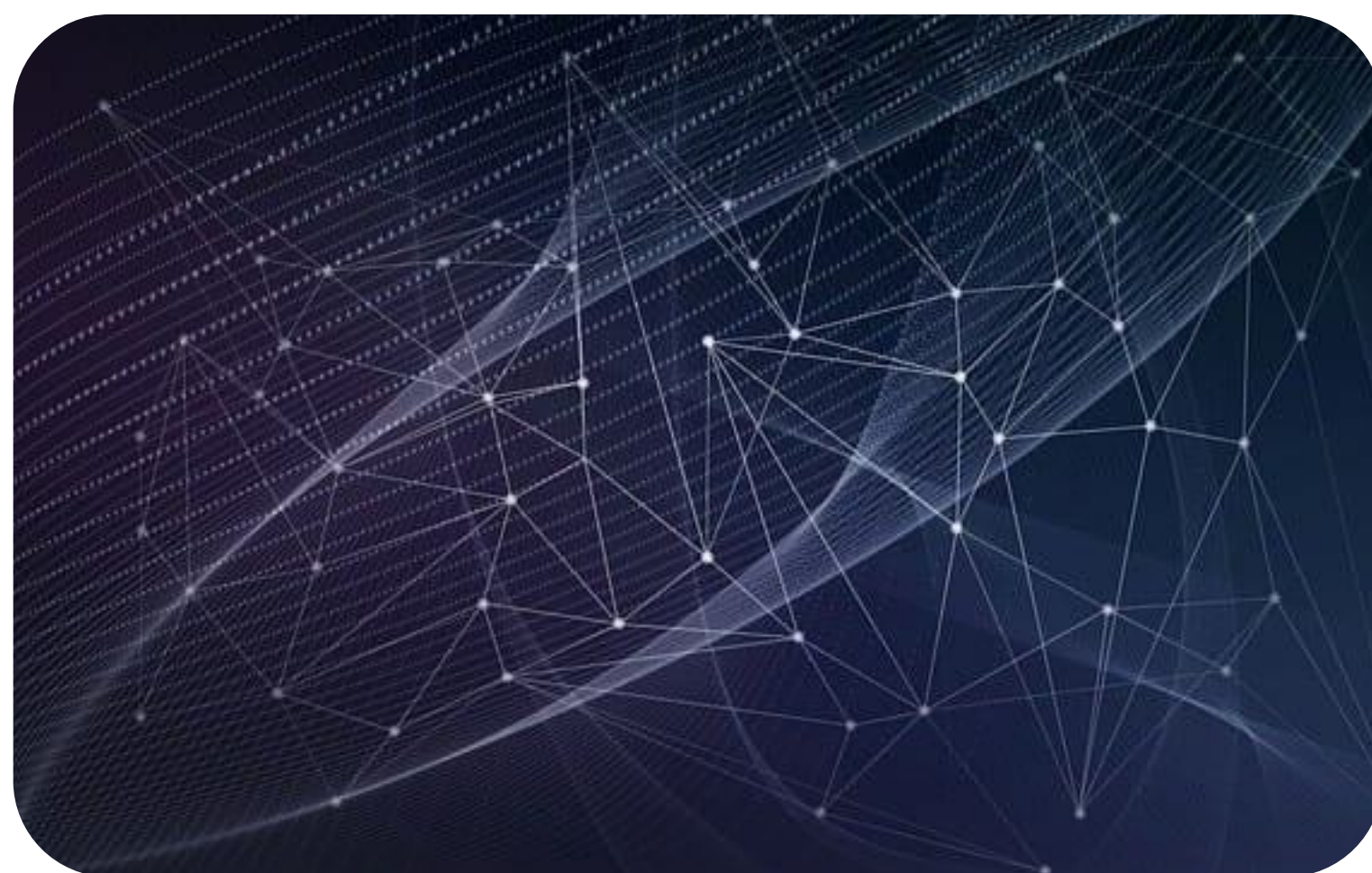


źródło: grafika własna - systemy biometrycznej identyfikacji w czasie rzeczywistym.

(na przykład drogi publiczne i place, parki, lasy i place zabaw). Przestrzeń należy uznać za przestrzeń publiczną również wtedy, gdy niezależnie od potencjalnych ograniczeń w zakresie pojemności lub bezpieczeństwa, dostęp do niej podlega pewnym określonym z góry warunkom, które mogą zostać spełnione przez nieokreśloną liczbę osób, takich jak zakup biletu wstępu lub biletu na przejazd, uprzednia rejestracja lub osiągnięcie określonego wieku. Danej przestrzeni nie należy natomiast uznawać za przestrzeń publiczną, jeśli dostęp do niej ograniczony jest do konkretnych i określonych osób fizycznych na mocy prawa Unii lub prawa krajowego bezpośrednio związanego z bezpieczeństwem publicznym lub ochroną publiczną [...] lub w wyniku wyraźnego wyrażenia woli przez osobę posiadającą odpowiednie uprawnienia związane z taką przestrzenią. Faktyczna możliwość samego dostępu (taka jak niezamknięte drzwi, otwarta bramka w ogrodzeniu) nie oznacza, że dana przestrzeń stanowi przestrzeń publiczną, jeśli istnieją wskazania lub okoliczności sugerujące inaczej (takie jak znaki zakazujące dostępu lub go ograniczające). Tereny przedsiębiorstw i fabryk, a także biura i miejsca pracy, do których dostęp powinni mieć wyłącznie odpowiedni pracownicy i usługodawcy, to przestrzenie, które nie stanowią przestrzeni publicznej. Do przestrzeni publicznej nie zaliczają się więzienia ani strefy kontroli granicznej. Niektóre przestrzenie mogą zawierać zarówno przestrzenie publiczne, jak i niepubliczne, takie jak hol w prywatnym budynku mieszkalnym prowadzący do gabinetu lekarskiego lub lotnisko. Przestrzenie internetowe również nie są objęte tym pojęciem, ponieważ nie są to przestrzenie fizyczne. To, czy dana przestrzeń jest dostępna publicznie powinno być jednak ustalane indywidualnie w każdym przypadku, z uwzględnieniem specyfiki danej sytuacji.

Podsumowanie

Jeśli potrzebujesz pomocy w kwestiach usystematyzowania zgodnie z aktem o sztucznej inteligencji wykorzystania modeli czy systemów AI w swojej firmie lub masz pomysły, które chciałbyś wdrożyć, zachęcamy nie tylko do lektury naszych artykułów, ale przede wszystkim do współpracy. Nie wahaj się z nami skontaktować. Czas wdrożeń kolejnych obowiązków zgodnie z AI Act nieubłaganie biegnie!





Newsletter RODO

Sierpień 2024 nr 9/2024

Dziękujemy za przeczytanie naszego newslettera!

Masz pytania?

SKONTAKTUJ SIĘ Z NAMI