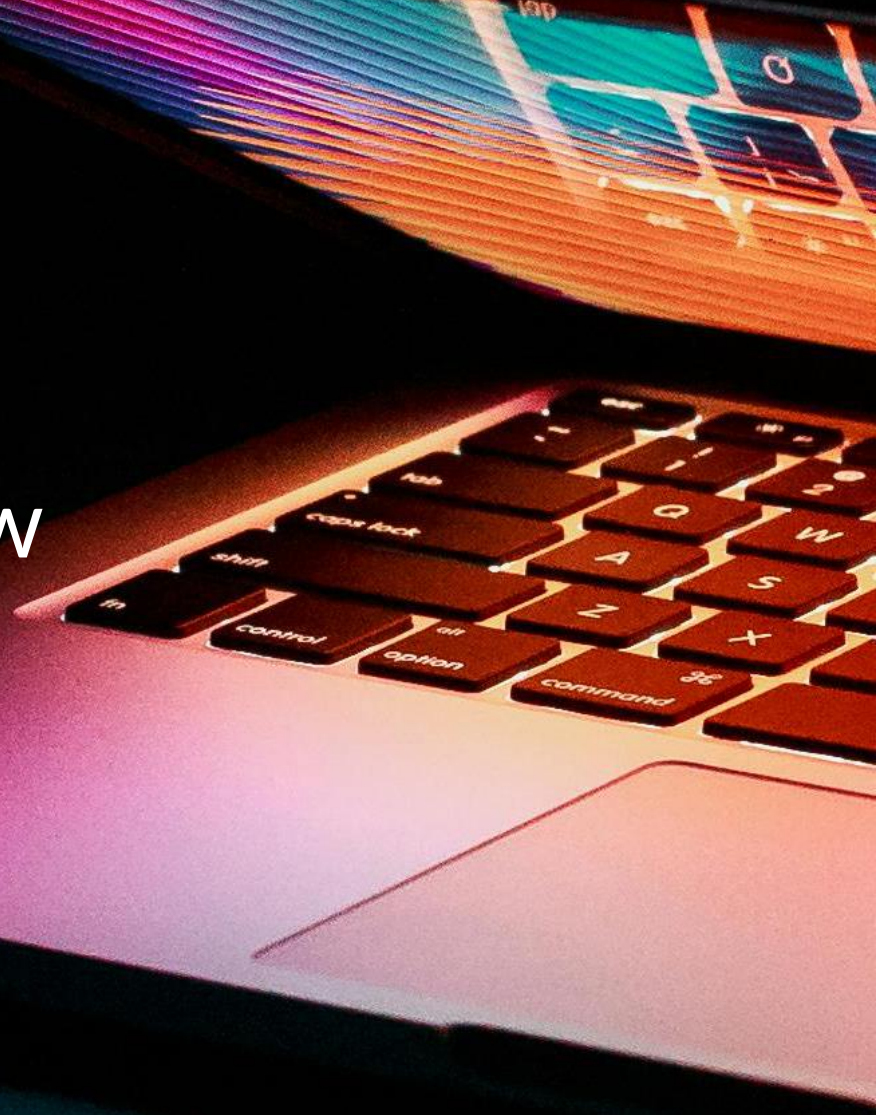
CYBERODPORNOŚĆ

Ocena skutków systemów AI dla praw podstawowych a ocena skutków dla ochrony danych

Piotr Drobek, UKSW



XVI Konferencja Bezpieczeństwo w Internecie - Cyberodporność

Warszawa 5 grudnia 2024 r.

Organizatorzy: UKSW, Ministerstwo Cyfryzacji, Naukowe Centrum Prawno-Informatyczne

Partner wspierający: NASK

Partner merytoryczny: SAMSUNG

- Ogólne Rozporządzenie o Ochronie Danych (RODO)
- Akt w sprawie sztucznej inteligencji
- Konwencja Ramowa Rady Europy w sprawie sztucznej inteligencji

Artykuł 35 RODO

- Jeżeli dany rodzaj przetwarzania – w szczególności z użyciem nowych technologii – ze względu na swój charakter, zakres, kontekst i cele z dużym prawdopodobieństwem może powodować wysokie ryzyko naruszenia praw lub wolności osób fizycznych, administrator przed rozpoczęciem przetwarzania dokonuje oceny skutków planowanych operacji przetwarzania dla ochrony danych osobowych. Dla podobnych operacji przetwarzania danych wiążących się z podobnym wysokim ryzykiem można przeprowadzić pojedynczą ocenę.
- Ocena skutków dla ochrony danych, o której mowa w ust. 1, jest wymagana w szczególności w przypadku:
- a) systematycznej, kompleksowej oceny czynników osobowych odnoszących się do osób fizycznych, która opiera się na zautomatyzowanym przetwarzaniu, w tym profilowaniu, i jest podstawą decyzji wywołujących skutki prawne wobec osoby fizycznej lub w podobny sposób znacząco wpływających na osobę fizyczną;
- b) przetwarzania na dużą skalę szczególnych kategorii danych osobowych, o których mowa w art. 9 ust. 1, lub danych osobowych dotyczących wyroków skazujących i czynów zabronionych, o czym mowa w art. 10; lub
- c) systematycznego monitorowania na dużą skalę miejsc dostępnych publicznie.

Artykuł 35 RODO

- Grupa Robocza Art. 29 - Wytyczne dotyczące oceny skutków dla ochrony danych oraz pomagające ustalić, czy przetwarzanie „może powodować wysokie ryzyko” do celów rozporządzenia 2016/679, (WP 248 rev. 01)
- Prezes UODO - Zmieniony wykaz rodzajów operacji wymagających oceny skutków dla ochrony danych (M.P. 2019 r. poz. 66)

Artykuł 35 RODO

Ocena zawiera co najmniej:

- systematyczny opis planowanych operacji przetwarzania i celów przetwarzania, w tym, gdy ma to zastosowanie – prawnie uzasadnionych interesów realizowanych przez administratora;
- ocenę, czy operacje przetwarzania są niezbędne oraz proporcjonalne w stosunku do celów;
- ocenę ryzyka naruszenia praw lub wolności osób, których dane dotyczą; oraz
- środki planowane w celu zaradzenia ryzyku, w tym zabezpieczenia oraz środki i mechanizmy bezpieczeństwa mające zapewnić ochronę danych osobowych i wykazać przestrzeganie niniejszego rozporządzenia, z uwzględnieniem praw i prawnie uzasadnionych interesów osób, których dane dotyczą, i innych osób, których sprawa dotyczy.

Artykuł 35 RODO

- Dokonując oceny skutków dla ochrony danych, administrator konsultuje się z inspektorem ochrony danych, jeżeli został on wyznaczony.
- W stosownych przypadkach administrator zasięga opinii osób, których dane dotyczą, lub ich przedstawicieli w sprawie zamierzonego przetwarzania, bez uszczerbku dla ochrony interesów handlowych lub publicznych lub bezpieczeństwa operacji przetwarzania.
- Nie mają zastosowania, jeżeli przetwarzanie na mocy art. 6 ust. 1 lit. c) lub e) ma podstawę prawną w prawie Unii lub w prawie państwa członkowskiego, któremu podlega administrator, i prawo takie reguluje daną operację przetwarzania lub zestaw operacji, a oceny skutków dla ochrony danych dokonano już w ramach oceny skutków regulacji w związku z przyjęciem tej podstawy prawnej – chyba że państwa członkowskie uznają za niezbędne, by przed podjęciem czynności przetwarzania dokonać oceny skutków dla ochrony danych.

Artykuł 27 Aktu w sprawie sztucznej inteligencji

- Podmioty stosujące systemy AI wysokiego ryzyka będące podmiotami prawa publicznego lub podmiotami prywatnymi świadczącymi usługi publiczne,
- oraz podmioty stosujące systemy AI wysokiego ryzyka,
  - przeznaczone do wykorzystywania do celów oceny zdolności kredytowej osób fizycznych lub ustalenia ich scoringu kredytowego, z wyjątkiem systemów AI wykorzystywanych w celu wykrywania oszustw finansowych;
  - przeznaczone do wykorzystywania przy ocenie ryzyka i ustalaniu cen w odniesieniu do osób fizycznych w przypadku ubezpieczenia na życie i ubezpieczenia zdrowotnego;
- przeprowadzają ocenę skutków w zakresie praw podstawowych, jakie może wywołać wykorzystanie takiego systemu,
- przed wdrożeniem takiego systemu, z wyjątkiem systemów AI wysokiego ryzyka przeznaczonych do stosowania w obszarze infrastruktury krytycznej.

Artykuł 27 Aktu w sprawie sztucznej inteligencji

- Obowiązek przeprowadzenia oceny skutków systemów AI dla praw podstawowych ma zastosowanie do wykorzystania systemu AI wysokiego ryzyka po raz pierwszy. W podobnych przypadkach podmiot stosujący może polegać na wcześniej przeprowadzonych ocenach skutków dla praw podstawowych lub na istniejących ocenach skutków przeprowadzonych przez dostawcę.
- Jeżeli w trakcie wykorzystania systemu AI wysokiego ryzyka podmiot stosujący uzna, że którykolwiek z elementów objętych oceną skutków uległ zmianie lub nie jest już aktualny, podmiot ten podejmuje niezbędne kroki w celu aktualizacji informacji.
- Po przeprowadzeniu oceny skutków, podmiot stosujący powiadamia organ nadzoru rynku o jej wynikach, przedkładając jako element tego powiadomienia wypełniony wzór opracowany przez Urząd ds. AI.

Artykuł 27 Aktu w sprawie sztucznej inteligencji

- Jeżeli którykolwiek z obowiązków ustanowionych w niniejszym artykule został już spełniony w wyniku oceny skutków dla ochrony danych przeprowadzonej zgodnie z art. 35 rozporządzenia (UE) 2016/679 lub art. 27 dyrektywy (UE) 2016/680, ocena skutków w zakresie praw podstawowych, o której mowa w ust. 1 niniejszego artykułu, stanowi uzupełnieniem tej oceny skutków dla ochrony danych.

Artykuł 27 Aktu w sprawie sztucznej inteligencji

Ocena obejmuje:

- opis procesów podmiotu stosującego, w których system AI wysokiego ryzyka będzie wykorzystywany zgodnie z jego przeznaczeniem;
- opis okresu, w którym każdy system AI wysokiego ryzyka ma być wykorzystywany i opis częstotliwości tego wykorzystywania;
- kategorie osób fizycznych i grup, na które może mieć wpływ wykorzystywanie systemu;
- szczególne ryzyko szkody, które może mieć wpływ na zidentyfikowane kategorie osób fizycznych lub grupy osób, z uwzględnieniem informacji przekazanych przez dostawcę zgodnie z art. 13;
- opis wdrożenia środków nadzoru ze strony człowieka, zgodnie z instrukcją obsługi;
- środki, jakie należy podjąć w przypadku urzeczywistnienia się tego ryzyka, w tym ustalenia dotyczące zarządzania wewnętrznego i mechanizmów rozpatrywania skarg.

Artykuł 16 Konwencji ramowej RE w sprawie sztucznej inteligencji

Ramy zarządzania ryzykiem i skutkami

Każda ze Stron, uwzględniając zasady określone w rozdziale III, przyjmuje lub utrzymuje środki służące identyfikacji, ocenie, zapobieganiu i łagodzeniu zagrożeń stwarzanych przez systemy sztucznej inteligencji, uwzględniając rzeczywisty i potencjalny wpływ na prawa człowieka, demokrację i rządy prawa.

Każda ze Stron przyjmuje lub utrzymuje środki mające na celu zapewnienie odpowiedniego przeciwdziałania negatywnemu wpływowi systemów sztucznej inteligencji na prawa człowieka, demokrację i rządy prawa. Takie negatywne skutki i środki mające na celu zaradzenie im powinny być udokumentowane i stanowić podstawę odpowiednich środków zarządzania ryzykiem.

Każda ze Stron oceni potrzebę wprowadzenia moratorium, zakazu lub innych odpowiednich środków w odniesieniu do niektórych zastosowań systemów sztucznej inteligencji, jeżeli uzna takie zastosowania za niezgodne z poszanowaniem praw człowieka, funkcjonowaniem demokracji lub rządy prawa.

Artykuł 16 Konwencji ramowej RE w sprawie sztucznej inteligencji

Środki zarządzania ryzykiem

- w należyty sposób uwzględniają kontekst i zamierzone zastosowanie systemów sztucznej inteligencji, w szczególności w odniesieniu do zagrożeń dla praw człowieka, demokracji i rządów prawa;
- należycie uwzględniają wagę i prawdopodobieństwo potencjalnych skutków;
- uwzględniają, w stosownych przypadkach, perspektywy odpowiednich zainteresowanych stron, w szczególności osób, których prawa mogą zostać naruszone;
- mają być stosowane iteracyjnie w trakcie działań w ramach cyklu życia systemu sztucznej inteligencji;
- obejmują monitorowanie ryzyka i negatywnego wpływu na prawa człowieka, demokrację i rządy prawa;
- obejmują dokumentację ryzyka, rzeczywistych i potencjalnych skutków oraz podejście do zarządzania ryzykiem; oraz
- wymagają, w stosownych przypadkach, testowania systemów sztucznej inteligencji przed udostępnieniem ich do pierwszego użycia oraz w przypadku ich znaczącej modyfikacji.

HUDERIA METHODOLOGY

COMMITTEE ON ARTIFICIAL INTELLIGENCE (CAI) – 28.11.2024 r.

METHODOLOGY FOR THE RISK AND IMPACT ASSESSMENT OF ARTIFICIAL INTELLIGENCE SYSTEMS FROM THE POINT OF VIEW OF HUMAN RIGHTS, DEMOCRACY AND THE RULE OF LAW (HUDERIA METHODOLOGY)

„HUDERIA" to wytyczne, które zapewniają ustrukturyzowane podejście do oceny ryzyka i wpływu systemów sztucznej inteligencji, specjalnie dostosowane do ochrony i promowania praw człowieka, demokracji i rządów prawa.

Mają one odgrywać wyjątkową i kluczową rolę na styku międzynarodowych standardów praw człowieka i istniejących ram technicznych dotyczących zarządzania ryzykiem w kontekście sztucznej inteligencji.

HUDERIA może być wykorzystywana zarówno przez podmioty publiczne, jak i prywatne, aby pomóc w identyfikacji i przeciwdziałaniu ryzyka i wpływowi na prawa człowieka, demokrację i rządy prawa przez cały cykl życia systemów sztucznej inteligencji.

HUDERIA METHODOLOGY

The Context-Based Risk Analysis (COBRA)

The Stakeholder Engagement Process (SEP)

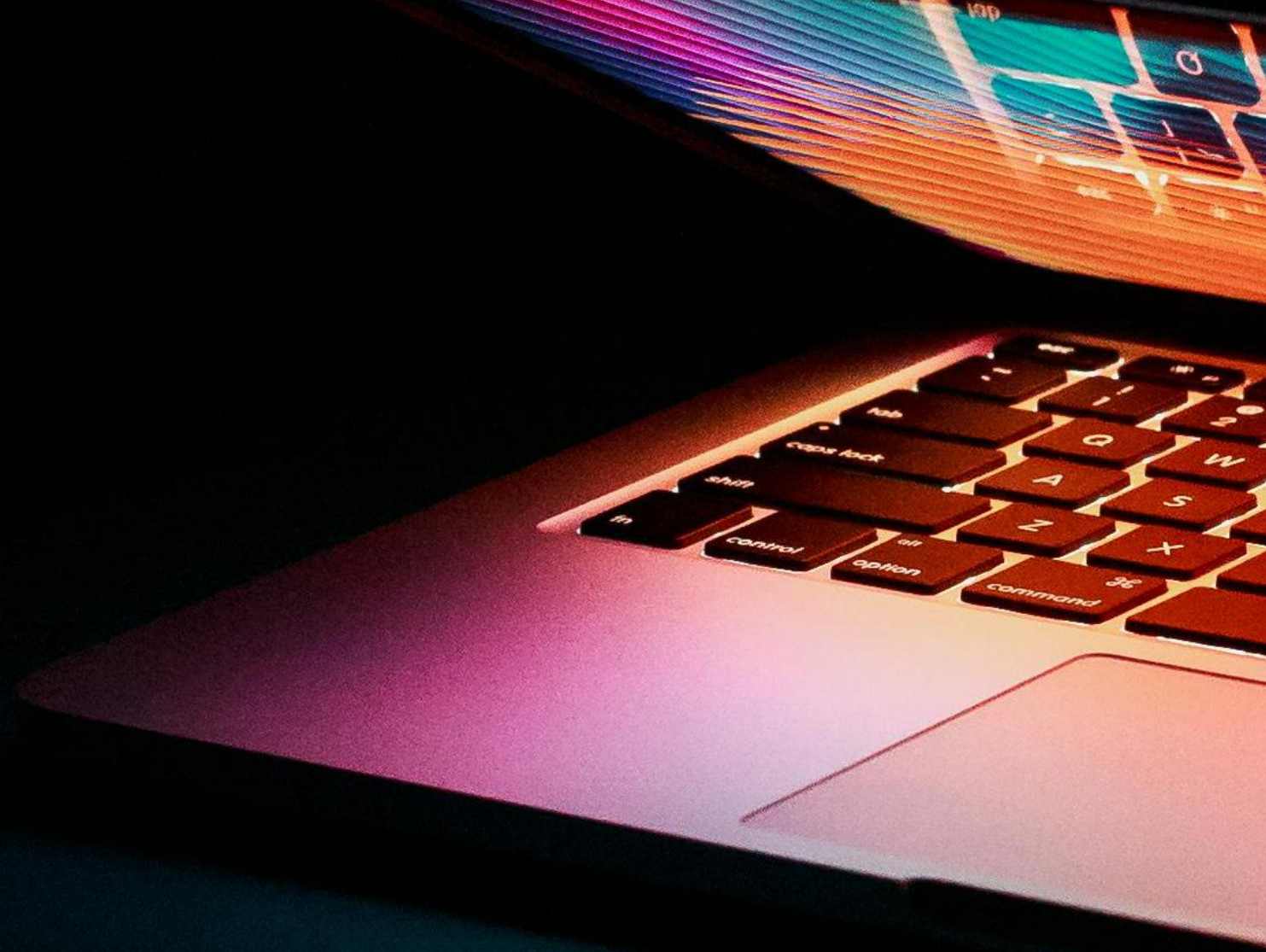
The Risk and Impact Assessment

Mitigation Plan

Iterative Review

CYBERODPORNOŚĆ

Dziękuję za uwagę!



UKSW

Ministerstwo Cyfryzacji

Naukowe Centrum Prawno-informatyczne

SAMSUNG

NASK