

AI RICO

METODYKA

Metodyka analizy wpływu systemu AI

Wpływ na osoby, prawa i społeczeństwo

Wersja: 1.1

Data: 11 października 2026 r.

Etap oceny: 2 z 3 – klasyfikacja → analiza wpływu → ocena ryzyka

Opracowana z uwzględnieniem wymogów ISO/IEC 42001:2023 (pkt 6.1.4, 8.4, załącznik A pkt A.5 i załącznik B pkt B.5), AI Act, RODO, Karty praw podstawowych UE, NIST AI RMF 1.0 oraz metodyki HUDERIA Rady Europy. Dokument nie jest wykładnią prawa.

SPIS TREŚCI

Streszczenie..... 3

1. Wprowadzenie..... 3

2. Koncepcja..... 4

3. Przebieg analizy 5

4. Role i odpowiedzialność..... 9

5. Kiedy wykonać i kiedy powtórzyć 9

6. Dokumentacja wyniku..... 10

Załącznik A. Skale..... 10

Załącznik B. Katalog praw i wartości..... 12

Załącznik C. Przesłanki DPIA..... 12

Załącznik D. Odniesienie do ISO/IEC 42001:2023 13

Załącznik E. Źródła 14

Załącznik F. Przykład 15

STRESZCZENIE

Analiza wpływu odpowiada na pytanie, **kogo system AI dotyka, w jaki sposób, jak dotkliwie i które prawa są w grze**. Wykonuje się ją przed pierwszym użyciem systemu wobec prawdziwych osób i powtarza przy każdej istotnej zmianie.

Metodyka dzieli pracę na siedem kroków: od opisu użycia, przez wskazanie osób i roli człowieka, po zestawienie korzyści i szkód, ocenę praw i wpływu na społeczeństwo oraz wnioski – czy wymagana jest ocena skutków dla ochrony danych (DPIA), czy użycie jest wrażliwe i czy potrzebna jest ocena skutków dla praw podstawowych (FRIA).

Trzy rozstrzygnięcia odróżniają ją od innych podejść:

- **Wpływ jest oddzielony od ryzyka.** Analiza wpływu ustala, co może się stać i jak dotkliwie to będzie. Nie szacuje prawdopodobieństwa i nie dobiera zabezpieczeń – to zadanie oceny ryzyka, która z jej wyników korzysta.
- **Dotkliwość szkody ma trzy składowe i jedną regułę złożenia:** skalę, zasięg i odwracalność ocenia się osobno, a stopień wynika z nich według stałej reguły (załącznik A). Dzięki temu dwie osoby oceniające ten sam system dochodzą do tego samego stopnia.
- **Brak informacji jest wynikiem.** Gdy faktu nie da się ustalić, pozycja dostaje wartość „do weryfikacji” z nazwaniem, czego brakuje i co zmieniłaby odpowiedź. Niepewność nie jest zastępowana domysłem.

Metodyka została opracowana z uwzględnieniem wymogów ISO/IEC 42001:2023, rozporządzenia o sztucznej inteligencji (AI Act), RODO, Karty praw podstawowych UE, NIST AI RMF 1.0 oraz metodyki HUDERIA Rady Europy. Nie jest wykładnią prawa.

1. WPROWADZENIE

1.1 Cel

Celem analizy wpływu jest dostarczenie osobom decydującym o wdrożeniu systemu AI uporządkowanej, powtarzalnej i sprawdzalnej odpowiedzi na cztery pytania:

1. Kogo dotyczą wyniki systemu – bezpośrednio i pośrednio?
2. Jaką rolę odgrywa wynik systemu w decyzjach o tych osobach i jaką rolę zachowuje człowiek?
3. Jakie korzyści i jakie szkody mogą z tego wyniknąć i jak dotkliwie byłyby szkody?
4. Które prawa i wartości są zaangażowane – także na poziomie społeczeństwa?

1.2 Dla kogo

Metodyka jest przeznaczona dla organizacji, które budują, kupują albo już używają systemów AI – niezależnie od sektora i wielkości. Posługują się nią: właściciel systemu (osoba odpowiadająca za jego użycie w procesie), osoba odpowiedzialna za zgodność w obszarze AI, inspektor ochrony danych oraz audytor.

1.3 Miejsce wśród innych ocen

Analiza wpływu jest drugim z trzech etapów oceny systemu AI.

ETAP	PYTANIE	WYNIK
1. Klasyfikacja	które przepisy AI Act dotyczą systemu i w jakiej roli występuje organizacja	kategoria ryzyka, obowiązki, wymóg FRIA
2. Analiza wpływu	kogo, jak, jak dotkliwie, jakie prawa	mapa osób, korzyści i szkód; dotkliwość; wnioski o DPIA i użyciu wrażliwym; materiał do FRIA
3. Ocena ryzyka	jak prawdopodobne są szkody, co je ogranicza, czy ryzyko jest akceptowalne	prawdopodobieństwo, zabezpieczenia, ryzyko rezydualne, rekomendacja

Granica między etapem 2 i 3 jest decyzją tej metodyki – żaden z czterech wzorców nie tnie tak ostro. Najbliższy jest podział NIST AI RMF na funkcję MAP (kontekst, osoby, korzyści i szkody) oraz funkcje MEASURE i MANAGE, ale NIST w MAP 5.1 każe udokumentować także prawdopodobieństwo każdego wpływu; HUDERIA ocenia prawdopodobieństwo jako czwartą zmienną obok trzech składowych dotkliwości, w tym samym kroku; ISO/IEC 42001 i wzór Microsoftu włączają do oceny wpływu wstępne środki zaradcze. Metodyka przenosi prawdopodobieństwo i środki do etapu 3, żeby analiza wpływu była powtarzalna z samych faktów o użyciu, a te dwa elementy dochodziły tam, gdzie ocenia się już ryzyko.

Analiza wpływu **nie zastępuje** oceny skutków dla ochrony danych (DPIA, art. 35 RODO) ani oceny skutków dla praw podstawowych (FRIA, art. 27 AI Act). Wskazuje, czy są wymagane, i dostarcza im materiału. DPIA prowadzi administrator danych z udziałem inspektora ochrony danych; FRIA – podmiot stosujący system.

1.4 Zakres

Analiza dotyczy **docelowego przeznaczenia** systemu i wszystkich zamierzonych sposobów użycia – także wtedy, gdy system jest dopiero próbą lub pilotażem. Obejmuje skutki dodatnie i ujemne: korzyści nie równoważą szkód w tej analizie, ale są wejściem do decyzji o akceptacji ryzyka w etapie 3 – bez nich tamta decyzja nie ma drugiej szali.

2. KONCEPCJA

2.1 Zasady

- 1. Podstawą są przepisy i uznane standardy.** Wewnętrzne zakazy i preferencje organizacji nie wchodzą do analizy; organizacja stosuje je osobno.
- 2. Fakty o produkcji pochodzą z dokumentacji, fakty o użyciu – od właściciela systemu.** Ograniczenia, znane błędy i zastosowania niewspierane ustala się z dokumentacji dostawcy. Skalę użycia, rolę człowieka i los osoby po niekorzystnym wyniku zna tylko organizacja.
- 3. „Do weryfikacji” jest lepsze niż błędna pewność.** Pozycja bez wystarczających danych nie dostaje werdyktu na siłę.
- 4. Wynik ma być zrozumiały dla osoby bez przygotowania prawnego.** Najważniejsze ustalenia mieszczą się na jednej stronie (karta wpływu, pkt 6.1).

5. **Dotkliwość bez prawdopodobieństwa.** Analiza mówi, jak poważna byłaby szkoda, a nie jak często wystąpi.
6. **Najpierw prawa, potem obowiązki.** Ochrona danych osobowych jest jednym z praw; wymóg DPIA jest konsekwencją stwierdzonego wpływu na to prawo, a nie punktem wyjścia analizy.

2.2 Pojęcia

Wpływ – skutek działania systemu dla osób, grup, organizacji lub społeczeństwa. Może być korzyścią albo szkodą.

Osoba, której system dotyczy – osoba, o której system coś ustala, rekomenduje lub decyduje, albo z którą wchodzi w wymianę. Nie musi być użytkownikiem systemu.

Użytkownik – osoba obsługująca system. Bywa zarazem osobą, której system dotyczy (na przykład pracownik).

Grupa w słabszym położeniu – dzieci, osoby starsze, osoby z niepełnosprawnością, pracownicy wobec pracodawcy, osoby w trudnej sytuacji ekonomicznej lub prawnej, osoby bez dostępu cyfrowego.

Dotkliwość szkody – złożenie trzech cech (za HUDERIA): **skali** (jak poważna jest szkoda dla jednej osoby), **zasięgu** (ile osób i czy są wśród nich grupy w słabszym położeniu) i **odwracalności** (czy i jak szybko da się ją naprawić). Każdą cechę ocenia się na trzech poziomach, stopień dotkliwości (niska · średnia · wysoka · krytyczna) wynika z nich według reguły z załącznika A.

Rola człowieka – to, co człowiek faktycznie robi z wynikiem systemu, zanim wynik wywoła skutek. Skala pięciostopniowa – załącznik A.

Użycie wrażliwe – użycie, w którym zachodzi co najmniej jedna z czterech okoliczności: (1) wpływ na sytuację prawną lub szanse życiowe osoby, (2) ryzyko istotnego urazu fizycznego lub psychicznego, (3) ograniczenie lub podważenie praw człowieka – prawo nie tylko „w grze”, ale z powiązaną szkodą co najmniej średnią, (4) szkoda dla społeczeństwa. Użycie wrażliwe wymaga pełnej, a nie skróconej, oceny ryzyka. Samo przetwarzanie danych osobowych nie czyni użycia wrażliwym – kieruje do wniosku o DPIA.

Wpływ na społeczeństwo – skutek wykraczający poza osoby, których system dotyczy bezpośrednio: dla środowiska, gospodarki, sfery publicznej, zdrowia i bezpieczeństwa publicznego albo norm i wartości.

3. PRZEBIEG ANALIZY

Analiza składa się z siedmiu kroków. Każdy kończy się zapisem: ustalenie, uzasadnienie, źródło oraz – gdy danych brakuje – wskazanie, czego brakuje i co zmieniłaby odpowiedź.

KROK	NAZWA	WYNIK
1	Teoria przypadku	krótki opis tego, co czyni system groźnym lub bezpiecznym dla ludzi
2	Użycie i kontekst	zamierzone użycia, ograniczenia, skala, czas, złożoność, jurysdykcja
3	Osoby	użytkownicy, osoby dotknięte bezpośrednio i pośrednio, grupy w słabszym położeniu

KROK	NAZWA	WYNIK
4	Decyzja i człowiek	rola wyniku w decyzji, rola człowieka, skutek dla osoby, odwracalność, odwołanie
5	Korzyści i szkody	zestawienie dla każdej grupy; dotkliwość każdej szkody
6	Prawa, wartości i społeczeństwo	ocena ośmiu pozycji katalogu
7	Wnioski	DPIA, użycie wrażliwe, FRIA

Krok 1 – Teoria przypadku

Cel. Zanim zaczniesz się wypełniać pozycje, trzeba zrozumieć ten konkretny system – inaczej analiza staje się listą kontrolną.

Jak wykonać. W czterech do ośmiu zdaniach opisać: co czyni ten system groźnym albo bezpiecznym dla ludzi; co jest w nim nietypowe; co realnie łagodzi skutki. Osobno sformułować **najostrzejszą krawędź** – jeden realistyczny scenariusz szkody z łańcuchem przyczyn i skutków. Na tym etapie jest to hipoteza: krok 5 ją potwierdza albo zastępuje scenariuszem, który po pełnej analizie okazał się dotkliwszy.

Wynik. Opis przypadku i jedno zdanie o najdotkliwszym scenariuszu (hipoteza). Krok nie ma werdyktu.

Krok 2 – Użycie i kontekst

Cel. Ustalić, do czego system ma służyć, do czego nie służy i w jakich warunkach działa.

Jak wykonać.

ELEMENT	CO USTALIĆ	ŹRÓDŁO
Zamierzone użycia	kto · do czego · gdzie – dla każdego użycia osobno	opis systemu
Użycia niewspierane i ograniczenia	czego system nie potrafi i gdzie zawodzi	dokumentacja dostawcy
Skala	ilu osób rocznie dotyczą wyniki (rzęd wielkości)	właściciel systemu
Czas i częstotaż użycia	stale, okresowo czy jednorazowo; planowany okres	właściciel systemu
Złożoność systemu	rodzaj technologii, model, składniki, dostawcy	rejestr systemów AI
Jurysdykcja	kraje, w których system działa i w których są osoby, których dotyczy	opis systemu

Wynik. Lista zamierzonych użyc i ograniczeń; skala i czas użycia; opis złożoności i jurysdykcji.

Krok 3 – Osoby

Cel. Wskazać wszystkich, których system dotyka – nie tylko tych, którzy go obsługują.

Jak wykonać. Ustalić kolejno: użytkowników; osoby, których wyniki dotyczą bezpośrednio (pracownicy, kandydaci, klienci, uczniowie, pacjenci, osoby ubiegające się o świadczenia, inne); grupy w słabszym położeniu wśród nich; osoby dotknięte pośrednio – na przykład rodzinę wnioskodawcy albo współpracowników osoby ocenianej.

Wynik. Lista grup z oznaczeniem, które są w słabszym położeniu.

Krok 4 – Decyzja i człowiek

Cel. Ustalić, co dzieje się z wynikiem systemu, zanim wywoła skutek dla osoby.

Jak wykonać.

ELEMENT	CO USTALIĆ
Rola wyniku w decyzji	wynik informuje · rekomenduje · wspiera decyzję · rozstrzyga samodzielnie
Nadzór	człowiek sprawdza każdy wynik · sprawdza wyrywkowo · nie sprawdza
Rola człowieka	poziom R1–R5 według załącznika A
Skutek dla osoby	na jaką decyzję lub sytuację wpływa wynik; czy skutek jest prawny albo podobnie istotny (art. 22 RODO)
Odwracalność	czy osoba może cofnąć skutek, w jakim czasie i czym działaniem
Informacja i odwołanie	jak osoba dowiaduje się o użyciu systemu i jak może się odwołać

Wynik. Poziom roli człowieka oraz opis skutku, odwracalności i drogi odwoławczej.

Krok 5 – Korzyści i szkody

Cel. Zestawić dla każdej grupy z kroku 3, co zyskuje i co może stracić, oraz ocenić dotkliwość szkód.

Jak wykonać. Dla każdej grupy opisać korzyści i szkody. Szkody rozważyć w trzech sytuacjach:

- **błąd systemu** – osobno wynik fałszywie dodatni i fałszywie ujemny;
- **nadużycie** – zamierzone i niezamierzone, w tym użycie niezgodne z przeznaczeniem;
- **niedostępność** – co się dzieje, gdy system przestaje działać.

Jako listę kontrolną źródeł szkód przyjąć cechy godnej zaufania AI według NIST: niedokładność, niebezpieczeństwo, podatność na ataki, brak rozliczalności i przejrzystości, niewyjaśnialność, naruszenie prywatności, stronniczość. Przy każdej grupie postawić dodatkowo dwa pytania: o **skutki finansowe** dla osoby (strata, koszt, utracona korzyść) oraz o **dostępność** – czy osoba z niepełnosprawnością albo bez dostępu cyfrowego dostaje to samo co pozostali.

Każdej szkodzie przypisać poziom skali, zasięgu i odwracalności, a z nich – stopień dotkliwości według reguły z załącznika A. Na koniec wrócić do najostrzejszej krawędzi z kroku 1: potwierdzić ją albo zastąpić szkodą o najwyższym stopniu, jeśli analiza wskazała inną.

Na tym etapie **nie dobiera się zabezpieczeń**. Zapisuje się tylko te, które już istnieją, z zaznaczeniem, że ich skuteczność nie była oceniana.

Wynik. Zestawienie korzyści i szkód dla każdej grupy; najwyższy stopień dotkliwości jako wynik kroku; potwierdzona albo poprawiona najostrzejsza krawędź; główna korzyść, która trafia na kartę i do etapu 3.

Krok 6 – Prawa, wartości i społeczeństwo

Cel. Sprawdzić, które prawa i wartości są zaangażowane – w tym na poziomie społeczeństwa.

Jak wykonać. Przejsz przez osiem pozycji katalogu z załącznika B. Przy każdej zapisać: dotyczy / nie dotyczy / do weryfikacji, w jaki sposób i kogo. Pozycję ósmą – społeczeństwo – ocenić w pięciu obszarach, każdy osobno, z kierunkiem wpływu (korzyść, szkoda albo oba).

W większości systemów wpływ na społeczeństwo nie występuje. Nie należy go szukać na siłę.

Wynik. Lista praw i obszarów, których system dotyczy, z opisem.

Krok 7 – Wnioski

Cel. Przełożyć ustalenia na trzy rozstrzygnięcia, od których zależy dalsze postępowanie.

Wnioski wynikają z ustaleń kroków 2–6 według stałych reguł. Nie są osobną oceną ekspercką – dwie osoby z tymi samymi ustaleniami muszą dojść do tych samych wniosków.

WNIOSEK	REGUŁA	SKUTEK
DPIA	system przetwarza dane osobowe i zachodzi co najmniej jedna przesłanka z art. 35 ust. 3 RODO albo co najmniej dwa z dziewięciu kryteriów WP 248 (oba wykazy – załącznik C)	wymagana / niewymagana / do weryfikacji, ze wskazaniem przesłanek
Użycie wrażliwe	zachodzi co najmniej jedna z czterech okoliczności z pkt 2.2	tak / nie / do weryfikacji; „tak” = pełna ocena ryzyka zamiast skróconej
FRIA	wymóg wyniku z klasyfikacji (art. 27 AI Act)	gdy wymagana – ustalenia analizy stają się materiałem do sześciu elementów art. 27 ust. 1

Przypisanie okoliczności użycia wrażliwego do ustaleń:

1. sytuacja prawna lub szanse życiowe – krok 6, pozycja 5 „dostęp do dóbr i szans”, albo skutek prawny lub podobnie istotny z kroku 4;
2. uraz fizyczny lub psychiczny – krok 5, szkoda o skali S3 dotycząca zdrowia, albo krok 6, pozycja 4;
3. prawa człowieka – dowolna z pozycji 1–4, 6 lub 7 w kroku 6 „dotyczy” i powiązana z nią szkoda z kroku 5 ma dotkliwość co najmniej średnią; pozycja 2 „dotyczy” bez takiej szkody prowadzi wyłącznie do wniosku o DPIA;
4. społeczeństwo – krok 6, pozycja 8: co najmniej jeden obszar ze szkodą.

Gdy którakolwiek z tych okoliczności opiera się na ustaleniu „do weryfikacji”, a żadna inna nie zachodzi, użycie wrażliwe jest „do weryfikacji”.

Materiał do FRIA: opis procesów i czas użycia (krok 2), kategorie osób (krok 3), ryzyko szkody (krok 5), nadzór człowieka (krok 4), postępowanie w razie szkody i droga skargi (kroki 4 i 7).

3.8 Analiza skrócona

Gdy system **nie dotyczy konkretnych osób**, nie przetwarza danych osobowych, nie jest systemem wysokiego ryzyka i nie podlega obowiązkowi przejrzystości z art. 50 AI Act, analizę można zamknąć wynikiem „bez wpływu na osoby” bez przechodzenia kroków 1–6. Osoba zatwierdzająca może mimo to zlecić pełną analizę.

4. ROLA I ODPOWIEDZIALNOŚĆ

ROLA	ZADANIA
Właściciel systemu	odpowiada na pytania o sposób użycia (pkt 4.1); potwierdza zamierzone użycia
Wykonawca analizy	przechodzi kroki 1–6 na podstawie opisu systemu, dokumentacji i odpowiedzi właściciela; dopytuje tylko o fakty, od których zależy ustalenie
Recenzent	niezależnie sprawdza wynik: czy ustalenia mają oparcie w danych i czy analiza nie przekracza granicy wobec oceny ryzyka
Osoba zatwierdzająca (odpowiedzialna za zgodność AI)	ocenia uwagi recenzenta, zmienia ustalenia z uzasadnieniem, rozstrzyga pozycje „do weryfikacji” albo pozostawia je z terminem, przyjmuje lub zwraca analizę

Wykonawcą analizy i recenzentem może być człowiek albo narzędzie wspomagające – w tym model językowy. Zatwierdza zawsze człowiek. W razie potrzeby organizacja konsultuje wynik z ekspertami dziedzinowymi i przedstawicielami osób, których system dotyczy.

4.1 Pytania do właściciela systemu

Każde dopuszcza odpowiedź „jeszcze nie ustalono” – wtedy powiązane ustalenie dostaje wartość „do weryfikacji”.

1. Ilu osób rocznie dotyczą wyniki systemu (do 100 · do 1 000 · do 10 000 · więcej)?
2. Jak długo i jak często system będzie używany?
3. Co dzieje się z osobą po niekorzystnym wyniku i czy da się to cofnąć?
4. Jak osoba dowiadyuje się o użyciu systemu i jak może się odwołać?
5. Czy próba lub pilotaż działa na prawdziwych osobach albo prawdziwych danych osobowych?

5. KIEDY WYKONAĆ I KIEDY POWTÓRZYĆ

Analiza ma być przyjęta **przed pierwszym użyciem systemu wobec prawdziwych osób**.

SYTUACJA	KIEDY
System budowany we własnym zakresie	w trakcie próby (PoC), przed rozpoczęciem pilotażu; próba na prawdziwych osobach lub danych osobowych wymaga analizy już na starcie
System kupowany	w trakcie kwalifikacji, przed pilotażem
System już używany	niezwłocznie po klasyfikacji; do przyjęcia analizy system działa warunkowo

Analizę powtarza się, gdy: pojawia się nowe zamierzone użycie; zmienia się etap życia systemu; zmienia się zakres lub rodzaj przetwarzanych danych; zmienia się model będący podstawą systemu; zmienia się rola wyniku w decyzji albo sposób nadzoru człowieka; zmienia się klasyfikacja systemu; mija dwanaście miesięcy od ostatniej analizy.

6. DOKUMENTACJA WYNIKU

6.1 Karta wpływu

Najważniejsze ustalenia na jednej stronie:

POLE	TREŚĆ
Kogo dotyczy	grupy z kroku 3 i skala z kroku 2
Rola człowieka	poziom R1–R5 opisany słownie
Najdotkliwsza szkoda	jedno zdanie potwierdzone w kroku 5 i jej stopień
Główna korzyść	jedno zdanie z kroku 5 – druga szala dla decyzji w etapie 3
Prawa w grze	pozycje „dotyczy” z kroku 6
DPIA	wymagana / niewymagana / do weryfikacji
Użycie wrażliwe	tak / nie / do weryfikacji – ocena ryzyka pełna albo skrócona
FRIA	wymagana / niewymagana / do weryfikacji

6.2 Pełny zapis

Pełny zapis analizy zawiera: opis przypadku; ustalenia kroków 2–7 z uzasadnieniem, źródłami i wskazaniem braków; pytania zadane właścicielowi i odpowiedzi; uwagi recenzenta z ich oceną; zmiany wprowadzone przez osobę zatwierdzającą z uzasadnieniem; różnice wobec poprzedniej przyjętej analizy; informację, kto i kiedy analizę przyjął oraz według której wersji metodyki została wykonana.

6.3 Przechowywanie

Przyjętych analiz nie usuwa się. Kolejna analiza nie zastępuje poprzedniej – organizacja zachowuje pełną historię przez cały okres życia systemu i po jego wycofaniu.

ZAŁĄCZNIK A. SKALE

A.1 Dotkliwość szkody

Każdą szkodę opisuje się trzema składowymi, każdą na jednym z trzech poziomów:

SKŁADOWA	POZIOM 1	POZIOM 2	POZIOM 3
Skala – jak poważna jest szkoda dla jednej osoby	S1 niedogodność: strata czasu, konieczność powtórzenia czynności, irytacja	S2 istotna strata: finansowa, utrata szansy, pogorszenie dobrostanu, ujawnienie informacji prywatnych	S3 utrata lub poważne zagrożenie zdrowia, wolności, środków do życia lub bezpieczeństwa; naruszenie prawa podstawowego
Zasięg – ile osób	Z1 pojedyncze osoby (do 100 rocznie)	Z2 wiele osób (do 10 000 rocznie)	Z3 masowy (powyżej 10 000 rocznie)
Odwracalność – czy i czym działaniem da się naprawić	O1 osoba naprawia sama, bez udziału organizacji	O2 naprawia organizacja (odwołanie, korekta, rekompensata) w rozsądnym czasie	O3 trudno odwracalna albo nieodwracalna; także wtedy, gdy osoba nie wie o szkodzie i nie może zareagować

Obecność grupy w słabszym położeniu wśród osób dotkniętych podnosi zasięg o jeden poziom.

Reguła złożenia. Stopień wyjściowy wynika ze skali: S1 → niska, S2 → średnia, S3 → wysoka. Jeśli zachodzi co najmniej jedno z: odwracalność O3, zasięg Z3 – stopień rośnie o jeden (raz, nie dwa). Stopień krytyczny jest osiągalny tylko przy skali S3.

SKALA	BEZ O3 I BEZ Z3	O3 LUB Z3
S1	niska	średnia
S2	średnia	wysoka
S3	wysoka	krytyczna

Gdy którejś składowej nie da się ustalić, stopień jest „do weryfikacji”; obok podaje się stopień, jaki wynikałby z najgorszego wariantu brakującej składowej.

Trzy składowe i ich nazwy pochodzą z HUDERIA (tam: scale, scope, reversibility). HUDERIA nie narzuca sposobu ich łączenia – dopuszcza metody ilościowe, półilościowe, macierze albo reguły i każe uwzględnić zarówno skutki o małym zasięgu i dużej skali, jak i o dużym zasięgu i małej skali; powyższa reguła jest jedną z takich metod. Dwie różnice wobec HUDERIA: tam zasięg obejmuje także ramy czasowe skutku i poziom ekspozycji grupy (jaka część grupy jest dotknięta) – metodyka ujmuje czas użycia w kroku 2, a ekspozycję grupy przez podniesienie zasięgu o jeden poziom; oraz HUDERIA ocenia w tym samym kroku prawdopodobieństwo, które metodyka przenosi do etapu 3.

A.2 Rola człowieka

POZIOM	OPIS
R1	człowiek reaguje tylko na alarmy
R2	system działa samodzielnie i przekazuje sprawę człowiekowi, gdy sam tak uzna
R3	system działa samodzielnie, a przekazanie sprawy człowiekowi wymusza reguła
R4	człowiek widzi wynik i może go zatrzymać; bez reakcji system działa dalej
R5	nic nie dzieje się bez decyzji człowieka

ZAŁĄCZNIK B. KATALOG PRAW I WARTOŚCI

NR	POZYCJA	NA CO ZWRÓCIĆ UWAGĘ
1	Niedyskryminacja i równe traktowanie (art. 21 i 26 Karty)	różna jakość usługi dla różnych grup; nierówny dostęp do zasobów i szans; stereotypizacja i pomijanie; dostępność dla osób z niepełnosprawnością
2	Prywatność i ochrona danych osobowych (art. 7–8)	zakres i rodzaj danych; podstawa wniosku o DPIA
3	Godność i autonomia (art. 1 i 3)	manipulacja, nadmierne poleganie na systemie, decyzje bez udziału człowieka
4	Bezpieczeństwo fizyczne i psychiczne (art. 2 i 3)	uraz, zagrożenie zdrowia, obciążenie psychiczne
5	Dostęp do dóbr i szans (art. 14, 15, 34)	zatrudnienie, edukacja, kredyt, świadczenia, mieszkanie, ubezpieczenie
6	Prawa pracownicze (art. 27–31)	monitorowanie, ocena, warunki pracy; wpływ na zatrudnienie i wymagane kompetencje
7	Prawa dziecka (art. 24)	każde użycie, które dotyczy dzieci albo jest dla nich dostępne
8	Społeczeństwo	pięć obszarów – niżej

Obszary wpływu na społeczeństwo (pozycja 8):

- **środowisko** – zużycie energii i zasobów, gdy skala przetwarzania jest istotna;
- **gospodarka** – dostęp do usług finansowych, rynek pracy, handel;
- **sfera publiczna** – stanowienie prawa, wybory, wymiar sprawiedliwości, bezpieczeństwo państwa;
- **zdrowie i bezpieczeństwo publiczne** – dostęp do opieki, diagnoza i leczenie, szkody na dużą skalę;
- **normy i wartości** – dezinformacja, utrwalanie uprzedzeń, treści wprowadzające w błąd.

Prawo do dobrej administracji i do skutecznego środka prawnego (art. 41 i 47 Karty) ocenia się w kroku 4, przy informacji i odwołaniu; wolność wypowiedzi (art. 11) – w pozycji 3.

ZAŁĄCZNIK C. PRZESŁANKI DPIA

DPIA jest wymagana, gdy system przetwarza dane osobowe i zachodzi co najmniej jedna przesłanka z art. 35 ust. 3 RODO:

- systematyczna, kompleksowa ocena osób oparta na zautomatyzowanym przetwarzaniu, w tym profilowaniu, będąca podstawą decyzji wywołujących skutki prawne lub podobnie istotne;
- przetwarzanie na dużą skalę szczególnych kategorii danych albo danych o wyrokach i czynach zabronionych;
- systematyczne monitorowanie na dużą skalę miejsc dostępnych publicznie;

albo co najmniej **dwa** z dziewięciu kryteriów wytycznych Grupy Roboczej Art. 29 (WP 248):

1. ocena lub punktacja osób, w tym profilowanie i przewidywanie;

2. automatyczne podejmowanie decyzji o skutku prawnym lub podobnie istotnym;
3. systematyczne monitorowanie;
4. dane wrażliwe lub o charakterze wysoce osobistym;
5. przetwarzanie na dużą skalę;
6. dopasowywanie lub łączenie zbiorów danych;
7. dane osób w słabszym położeniu;
8. innowacyjne zastosowanie technologii;
9. przetwarzanie, które utrudnia osobie skorzystanie z prawa, usługi lub umowy.

Organizacja uwzględnia także wykaz rodzajów operacji wymagających DPIA ogłoszony przez właściwy organ nadzorczy.

ZAŁĄCZNIK D. ODNIESIENIE DO ISO/IEC 42001:2023

Norma zalicza do oceny wpływu także prawdopodobieństwo, decyzję o akceptacji i środki zaradcze. W tej metodyce te trzy elementy należą do oceny ryzyka; norma wymaga, aby wynik oceny wpływu był w niej uwzględniony (pkt 6.1.4). Wymagania normy pokrywają więc analiza wpływu i ocena ryzyka łącznie.

PUNKT NORMY	WYMAGANIE	MIEJSCE W METODYCE
6.1.4	proces oceny skutków dla osób, grup i społeczeństwa	cała metodyka
6.1.4	zamierzone użycie i przewidywalne nadużycie	kroki 2 i 5
6.1.4	kontekst techniczny i społeczny, jurysdykcje	kroki 2, 3 i 6
6.1.4	wynik udokumentowany	pkt 6
6.1.4	wynik uwzględniony w ocenie ryzyka	krok 7 – użycie wrażliwe; dotkliwość z kroku 5
8.4	w zaplanowanych odstępach i przy istotnych zmianach	pkt 5
8.4, A.5.3	przechowywanie wyników	pkt 6.3
A.5.2, B.5.2 lit. a	okoliczności wykonania oceny	pkt 5
B.5.2 lit. b	identyfikacja i analiza skutków	kroki 2–6
B.5.2 lit. b	prawdopodobieństwo, akceptacja, środki	ocena ryzyka (etap 3)
B.5.2 lit. c	kto wykonuje ocenę	pkt 4
B.5.2 lit. d	do czego służy wynik	krok 7
B.5.2	sytuacja prawna i szanse, dobrostan, prawa człowieka, społeczeństwa	cztery okoliczności użycia wrażliwego
B.5.3	użycie i nadużycie, wpływ dodatni i ujemny, przewidywalne awarie	kroki 2 i 5
B.5.3	grupy demograficzne · złożoność · rola człowieka	kroki 3 · 2 · 4

PUNKT NORMY	WYMAGANIE	MIEJSCE W METODYCE
B.5.3	zatrudnienie i kompetencje pracowników	krok 6, pozycja 6
A.5.4, B.5.4	potrzeby dzieci, osób z niepełnosprawnością, starszych, pracowników	krok 3; krok 6, pozycje 6 i 7
B.5.4	sprawiedliwość, rozliczalność, przejrzystość, prywatność, zdrowie, skutki finansowe, dostępność, prawa człowieka	kroki 4, 5 i 6
B.5.4	konsultacja z ekspertami	pkt 4
A.5.5, B.5.5	wpływ na społeczeństwo	krok 6, pozycja 8

ZAŁĄCZNIK E. ŹRÓDŁA

ŹRÓDŁO	CO Z NIEGO PRZYJĘTO
Rozporządzenie (UE) 2024/1689 (AI Act), art. 14, 26, 27, 50	elementy FRIA jako materiał wynikowy; nadzór człowieka
Rozporządzenie (UE) 2016/679 (RODO), art. 22 i 35; wytyczne WP 248	reguła wymogu DPIA
Karta praw podstawowych Unii Europejskiej	katalog praw
ISO/IEC 42001:2023, pkt 6.1.4, 8.4, załącznik A pkt A.5, załącznik B pkt B.5	wymagania wobec procesu i zapisu; wpływ na społeczeństwo
NIST AI Risk Management Framework 1.0 (AI 100-1, styczeń 2023), pkt 1.1, 1.2.3, funkcja MAP	definicje wpływu i ryzyka; reguła priorytetu dla systemów wchodzących w interakcję z ludźmi, na danych wrażliwych albo o wynikach dotyczących ludzi; podział funkcji jako najbliższy wzorzec granicy wobec oceny ryzyka; cechy godnej zaufania AI jako lista kontrolna źródeł szkód
Microsoft Responsible AI Impact Assessment (wzór i przewodnik, czerwiec 2022)	układ: zamierzone użycie jako kto · do czego · gdzie → interesariusze z korzyściami i szkodami → użycia niewspierane i ograniczenia → skutek błędu (fałszywie dodatni i ujemny) i nadużycia; pięć poziomów roli człowieka; trzy cele sprawiedliwości; trzy okoliczności użycia wrażliwego; aktualizacja co najmniej raz w roku, przy nowym użyciu i przed kolejnym etapem wydania
HUDERIA – Rada Europy, „HUDERIA Methodology and Model” (metodyka zatwierdzona 26 lutego 2025 r., model COBRA 25 lutego 2026 r.)	dotkliwość jako złożenie skali, zasięgu i odwracalności (za Konwencją ramową Rady Europy i Wytycznymi ONZ dotyczącymi biznesu i praw człowieka); definicje tych trzech składowych; wskazanie, że metodę ich łączenia ustala użytkownik; triaż – system bez wpływu kończy proces, korzyści ważone z ryzykiem; hierarchia środków avoid → mitigate → restore → compensate (dla etapu 3)

Wzór Microsoftu jest formularzem z przewodnikiem do pracy zespołowej: nie ma skali dotkliwości ani oceny prawdopodobieństwa (o prawdopodobieństwo pyta raz, jakościowo – czy awaria dotyka grup marginalizowanych częściej lub mocniej), za to obejmuje wstępne środki łagodzące, ocenę przydatności

systemu do celu, gotowość technologii, złożoność zadania i środowiska oraz wymagania wobec danych. Tych elementów metodyka nie obejmuje: środki i ocena przydatności należą do etapu 3, pozostałe są materiałem dla etapu 3, nie dla analizy wpływu. Skalę dotkliwości metodyka bierze z HUDERIA, nie z wzoru Microsoftu.

Metodyka nie korzysta z tekstu ISO/IEC 42005:2025; jej układ jest spójny z ogólnie znanym zakresem tej normy. Wzór kwestionariusza FRIA przewidziany w art. 27 ust. 5 AI Act nie został opublikowany do dnia wydania tej wersji – po publikacji materiał wynikowy zostanie dopasowany do jego układu.

ZAŁĄCZNIK F. PRZYKŁAD

System. Narzędzie dostawcy zewnętrznego szereguje kandydatów według dopasowania do ogłoszenia o pracę. Rekruter widzi listę posortowaną od najlepszego dopasowania i przegląda górne trzydzieści pozycji; pozostali kandydaci po czternastu dniach otrzymują odmowę, jeśli rekruter nie zareaguje. Około 5 000 kandydatów rocznie, Polska. Klasyfikacja: system wysokiego ryzyka (załącznik III pkt 4 lit. a AI Act), organizacja jest podmiotem stosującym.

Krok 1 – teoria przypadku. System nie odrzuca nikogo wprost, ale ustawia kolejność, a kolejność rozstrzyga, kto w ogóle zostanie obejrany. Osoba poniżej trzydziestej pozycji nie dowiaduje się, że o jej losie zdecydował model. Najostrzejsza krawędź (hipoteza): kandydat z nietypową ścieżką – przerwa opiekuńcza, dyplom zagraniczny – trafia na dół listy i dostaje odmowę, choć żaden człowiek nie spojrzał na jego zgłoszenie.

Krok 2 – użycie i kontekst. Zamierzone użycie: rekruterzy · wstępna selekcja zgłoszeń · rekrutacje zewnętrzne. Ograniczenia z dokumentacji dostawcy: model trenowany na danych z rynku anglojęzycznego; dostawca zastrzega, że narzędzie „wspiera, nie zastępuje” decyzji. Skala: do 10 000 rocznie. Użycie stałe.

Krok 3 – osoby. Użytkownicy: rekruterzy. Dotknięci bezpośrednio: kandydaci. Grupy w słabszym położeniu: osoby po przerwie zawodowej, osoby z niepełnosprawnością, cudzoziemcy. Pośrednio: zespoły, do których trafiają zatrudnieni.

Krok 4 – decyzja i człowiek. Wynik rekomenduje; nadzór wyrywkowy (górne trzydzieści pozycji); rola człowieka R4 – bez reakcji rekrutera odmowa wychodzi sama. Skutek: brak zaproszenia na rozmowę – skutek podobnie istotny w rozumieniu art. 22 RODO. Odwracalność: kandydat nie wie o użyciu systemu, więc nie ma jak zareagować. Informacja i odwołanie: brak informacji o użyciu systemu; odwołanie formalnie możliwe, praktycznie nieużywane.

Krok 5 – korzyści i szkody. Korzyści: kandydaci dostają odpowiedź szybciej; rekruterzy oszczędzają czas na rekrutacjach masowych. Szkody – najdotkliwsza: wynik fałszywie ujemny (dobry kandydat nisko na liście) → utrata szansy zatrudnienia: skala S2, zasięg Z2 podniesiony do Z3 przez grupy w słabszym położeniu, odwracalność O3 (kandydat nie wie) → **wysoka**. Ta sama szkoda z innej przyczyny: stronniczość modelu wobec przerw w karierze – stopień ten sam, ale dotyczy wskazanych grup w większym stopniu. Niedostępność systemu: rekrutacja ręczna, S1 → niska. Nadużycie: rekruter traktuje ranking jako jedyne kryterium – już uwzględnione w R4. Krawędź z kroku 1 potwierdzona. Główna korzyść: krótszy czas odpowiedzi dla kandydatów.

Krok 6 – prawa, wartości i społeczeństwo. 1 niedyskryminacja – dotyczy (jakość oceny różna dla osób z nietypową ścieżką); 2 prywatność – dotyczy (profilowanie na danych z CV); 3 godność i autonomia – dotyczy (decyzja bez udziału człowieka dla osób poniżej trzydziestej pozycji); 4 bezpieczeństwo – nie dotyczy; 5 dostęp do dóbr i szans – dotyczy (zatrudnienie); 6 prawa pracownicze – nie dotyczy (kandydaci nie są jeszcze pracownikami); 7 prawa dziecka – nie dotyczy; 8 społeczeństwo – nie dotyczy w żadnym z pięciu obszarów.

Krok 7 – wnioski. DPIA: dane osobowe, systematyczna ocena z profilowaniem i skutkiem podobnie istotnym (art. 35 ust. 3 lit. a) → **wymagana**. Użycie wrażliwe: okoliczność (1) – pozycja 5 „dotyczy” → **tak**, pełna ocena ryzyka. FRIA: wymóg według klasyfikacji – zależy od statusu organizacji w rozumieniu art. 27 AI Act.

Karta wpływu. Kogo dotyczy: kandydaci, do 10 000 rocznie, w tym osoby po przerwie zawodowej, z niepełnosprawnością, cudzoziemcy · Rola człowieka: R4 · Najdotkliwsza szkoda: dobry kandydat z nietypową ścieżką odrzucony bez spojrzenia człowieka – wysoka · Główna korzyść: szybsza odpowiedź dla kandydatów · Prawa w grze: niedyskryminacja, prywatność, godność i autonomia, dostęp do szans · DPIA: wymagana · Użycie wrażliwe: tak · FRIA: według klasyfikacji.