SZABLON DOKUMENTU WEWNĘTRZNEGO

Polityka korzystania z systemów sztucznej inteligencji

Bezpieczna droga od pomysłu do wdrożenia

Pole	Wartość do uzupełnienia
Organizacja	[NAZWA ORGANIZACJI]
Właściciel dokumentu	[ROLA LUB JEDNOSTKA]
Zatwierdził	[IMIĘ I NAZWISKO LUB ORGAN]
Wersja	[NUMER WERSJI]
Data wejścia w życie	[DD.MM.RRRR]
Termin przeglądu	[DD.MM.RRRR]
Klasyfikacja dokumentu	[WEWNĘTRZNY PUBLICZNY INNY]

Niniejszy szablon należy dostosować do struktury organizacji, profilu działalności, klasyfikacji informacji, obowiązujących przepisów i przyjętego modelu zarządzania ryzykiem. Przed zatwierdzeniem dokument powinien zostać zweryfikowany przez właściwe jednostki prawne, bezpieczeństwa, ochrony danych i zarządzania ryzykiem.

Instrukcja użycia szablonu

- Zastąp wszystkie pola w nawiasach kwadratowych danymi organizacji.
- Usuń warianty, które nie pasują do przyjętego modelu decyzyjnego.
- Uzupełnij załącznik z zatwierdzonymi narzędziami i właścicielami.
- Powiąż dokument z polityką bezpieczeństwa informacji, ochroną danych, zakupami, SSDLC i procedurą zarządzania incydentami.

1. Cel polityki

Celem polityki jest umożliwienie pracownikom i współpracownikom bezpiecznego, zgodnego z prawem i kontrolowanego korzystania z systemów sztucznej inteligencji. Polityka określa zasady używania narzędzi AI, klasyfikowania przypadków użycia, oceny ryzyka, podejmowania decyzji, uruchamiania rozwiązań oraz reagowania na błędy i incydenty.

Organizacja nie ogranicza zarządzania AI do listy zakazanych produktów. Zapewnia pracownikom przewidywalną ścieżkę zgłoszenia pomysłu, szybkiej oceny ryzyka oraz uzyskania decyzji o eksperymencie, pilotażu lub wdrożeniu.

2. Zakres obowiązywania

Polityka obowiązuje [PRACOWNIKÓW WSPÓŁPRACOWNIKÓW DOSTAWCÓW STAŻYSTÓW INNE OSOBY] korzystających z AI w imieniu organizacji lub przy użyciu jej danych, urządzeń, kont, infrastruktury albo procesów.

Polityka obejmuje w szczególności:

- publiczne i firmowe chatboty oraz asystentów AI
- funkcje AI wbudowane w aplikacje biurowe i usługi SaaS
- generowanie oraz analizę tekstu, obrazu, dźwięku, wideo i kodu
- rozwiązania RAG, wyszukiwarki semantyczne i firmowe bazy wiedzy
- modele dostrajane, lokalne i udostępniane przez API
- automatyzacje i agentów AI wykonujących działania w systemach
- aplikacje i prototypy tworzone metodą vibe codingu lub z pomocą asystentów programistycznych

3. Definicje

Pojęcie	Znaczenie przyjęte w polityce
System AI	System maszynowy działający z różnym poziomem autonomii, który na podstawie danych wejściowych generuje wyniki takie jak predykcje, treści, rekomendacje lub decyzje.
Generative AI	System AI generujący nową treść, w tym tekst, kod, obraz, dźwięk lub wideo.
Agent AI	Rozwiązanie, które planuje kroki, korzysta z narzędzi lub danych i może wykonywać działania zmieniające stan systemów.
RAG	Architektura, w której model otrzymuje kontekst pobrany z zewnętrznego repozytorium wiedzy.
Shadow AI	Użycie systemu AI poza zatwierdzonym procesem, bez wymaganej rejestracji, oceny, umowy, kontroli lub zgody.
Vibe coding	Tworzenie kodu lub aplikacji przy intensywnym wsparciu AI, często na podstawie poleceń w języku naturalnym.
Przypadek użycia	Opis konkretnego celu, procesu, danych, odbiorców i sposobu użycia systemu AI.
Właściciel przypadku użycia	Osoba odpowiedzialna za cel biznesowy, jakość wyniku, ryzyko i cykl życia rozwiązania.

4. Zasady nadrzędne

- **Cel przed narzędziem** - każdy przypadek użycia musi mieć uzasadniony cel i właściciela.
- **Minimalizacja danych** - do systemu wprowadza się wyłącznie dane niezbędne do wykonania zadania.
- **Weryfikacja wyniku** - treści i rekomendacje AI nie są automatycznie traktowane jako poprawne lub wiążące.
- **Minimalne uprawnienia** - agent lub integracja otrzymuje tylko dostęp potrzebny do konkretnego zadania.
- **Nadzór człowieka** - decyzje o znaczącym wpływie na ludzi, finanse, prawa lub bezpieczeństwo podlegają właściwej kontroli człowieka.
- **Rozliczalność** - organizacja zachowuje adekwatny ślad danych, źródeł, wersji modelu, działań i zatwierdzeń.
- **Bezpieczny cykl życia** - prototyp nie staje się rozwiązaniem produkcyjnym bez spełnienia określonych wymagań.
- **Zgłaszanie bez ukrywania** - działanie w dobrej wierze i szybkie zgłoszenie błędu lub Shadow AI jest wspierane przez organizację.

5. Dane wprowadzane do systemów AI

Dopuszczalność wprowadzenia danych zależy od ich klasyfikacji, umowy z dostawcą, lokalizacji przetwarzania, retencji, użycia do trenowania modeli, kontroli dostępu oraz zatwierdzonego przypadku użycia.

Klasa danych	Domyślna zasada	Warunki lub przykłady
Publiczne	Dozwolone w zatwierdzonych narzędziach	Treści oficjalnie opublikowane lub przeznaczone do publikacji. Nadal należy sprawdzić prawa autorskie, wiarygodność i cel.
Wewnętrzne	Dozwolone warunkowo	Wyłącznie w narzędziach zatwierdzonych do tej klasy danych, z kontem służbowym i odpowiednimi ustawieniami prywatności.
Poufne	Wymaga zgłoszenia i zgody	Umowy, informacje handlowe, kod, architektura, dane klientów i niepubliczne strategie. Wymagana ocena właściciela danych i bezpieczeństwa.
Dane osobowe	Wymaga podstawy i oceny prywatności	Należy określić cel, podstawę prawną, minimalizację, retencję, role stron i ewentualną potrzebę DPIA.
Dane szczególnych kategorii	Domyślnie zabronione bez formalnej zgody	Dopuszczenie wymaga oceny DPO lub prawnej, właściciela danych i zatwierdzonego środowiska.
Sekrety i dane uwierzytelniające	Zabronione	Hasła, tokeny, klucze API, klucze prywatne, kody odzyskiwania i dane pozwalające przejąć konto lub usługę.
Dane objęte tajemnicą lub ograniczeniami sektorowymi	Zabronione bez dedykowanej decyzji	Dane regulowane, tajemnice zawodowe, informacje niejawne lub dane kontraktowo ograniczone.

Jeżeli użytkownik nie potrafi ustalić klasy danych lub warunków dostawcy, nie wprowadza danych do systemu i kontaktuje się z [WŁAŚCICIELEM DANYCH SECURITY DPO SERVICE DESK].

6. Zatwierdzone narzędzia

Aktualny wykaz narzędzi znajduje się w Załączniku 1 lub w rejestrze [NAZWA I LOKALIZACJA REJESTRU]. Umieszczenie produktu w wykazie nie oznacza zgody na każde zastosowanie. Zatwierdzenie zawsze wskazuje dozwolone konta, klasy danych, funkcje, integracje, region, retencję i ograniczenia.

Status	Znaczenie	Działanie użytkownika
Zatwierdzone	Narzędzie może być używane w opisanym zakresie.	Przestrzegaj warunków wpisu i zgłoś nowe zastosowanie, jeśli spełnia kryteria z rozdziału 8.
Warunkowe	Dozwolone tylko dla wskazanego zespołu, danych, pilotażu lub okresu.	Potwierdź, że Twój przypadek mieści się w zatwierdzonym zakresie.
Do oceny	Brakuje zakończonej oceny lub umowy.	Nie używaj danych organizacji. Złóż zgłoszenie.
Zabronione	Narzędzie lub funkcja nie może być używana w organizacji.	Nie używaj. Jeżeli narzędzie już jest używane, zgłoś ten fakt zgodnie z rozdziałem 14.

7. Zabronione zastosowania

Zabrania się korzystania z AI w celu lub w sposób, który:

- narusza prawo, prawa osób, zobowiązania umowne, regulacje sektorowe lub wewnętrzne zasady organizacji
- realizuje praktyki zakazane na podstawie właściwych przepisów AI Act lub innych regulacji
- wymaga podania sekretów, danych uwierzytelniających lub obejścia zabezpieczeń
- umożliwia niejawne profilowanie, dyskryminację, manipulację albo monitorowanie osób bez właściwej podstawy i zgody
- generuje lub rozpowszechnia treści podszywające się pod osoby albo organizację bez autoryzacji i wymaganego oznaczenia
- automatycznie podejmuje lub wykonuje decyzje o znaczącym wpływie bez zatwierdzonego nadzoru człowieka
- pozwala agentowi wykonywać płatności, zwroty, wysyłkę, usuwanie danych, zmiany uprawnień lub inne działania wysokiego wpływu bez kontroli po stronie usługi
- łączy dostęp do poufnych danych z nieograniczonym przeglądaniem internetu lub wysyłaniem danych na zewnętrzne adresy
- uruchamia produkcyjnie aplikację stworzoną metodą vibe codingu poza zatwierdzonym SSDLC
- wykorzystuje dane lub materiały bez wymaganej licencji, podstawy prawnej albo uprawnienia

8. Kiedy należy zgłosić przypadek użycia

Zgłoszenie jest wymagane przed rozpoczęciem pilotażu lub użycia, jeżeli występuje co najmniej jedna z poniższych przesłanek:

- nowe narzędzie, model, dostawca, konto, wtyczka lub rozszerzenie
- dane inne niż publiczne albo dane dotyczące osób

- RAG, indeks wektorowy, dostrajanie modelu lub własna baza wiedzy
- integracja z pocztą, plikami, CRM, ERP, HR, kodem, bazą danych lub API
- agent wykonujący działania albo podejmujący wieloetapowe decyzje
- wpływ na zatrudnienie, edukację, dostęp do usług, kredyt, zdrowie, bezpieczeństwo lub prawa osób
- automatyczne generowanie komunikacji zewnętrznej, ofert, umów, kodu produkcyjnego lub decyzji
- użycie biometrii, analizy emocji, danych dzieci lub grup szczególnie narażonych
- istotna zmiana celu, danych, modelu, uprawnień, integracji, autonomii lub skali istniejącego rozwiązania

Zgłoszenie nie jest wymagane dla zastosowań wymienionych w zatwierdzonym katalogu niskiego ryzyka, o ile użytkownik spełnia wszystkie warunki katalogu. Katalog i wyjątki prowadzi [WŁAŚCICIEL REJESTRU].

9. Bezpieczna droga od pomysłu do wdrożenia

Każdy zgłoszony przypadek przechodzi ścieżkę proporcjonalną do ryzyka. Celem procesu jest szybkie udzielenie odpowiedzi: zgoda, zgoda warunkowa, potrzeba dodatkowych testów albo odmowa z uzasadnieniem.

Etap	Wymagany rezultat	Warunek przejścia
1. Pomysł	Cel, właściciel, użytkownicy, oczekiwany efekt i skutek błędu.	Kompletne zgłoszenie z Załącznika 2.
2. Triage	Wstępny próg ryzyka i lista zaangażowanych ról.	Wyznaczona ścieżka i termin decyzji.
3. Klasyfikacja	Rola organizacji, klasyfikacja AI Act, prywatność, dane, bezpieczeństwo i ryzyko biznesowe.	Udokumentowane wymagania i właściciel ryzyka.
4. PoC	Test w ograniczonym środowisku z kryteriami sukcesu i zatrzymania.	Brak nieakceptowalnych wyników i znane ograniczenia.
5. Zabezpieczenia	Kontrole techniczne, organizacyjne, prawne i procesowe.	Spełniona checklista przed uruchomieniem.
6. Decyzja	Zgoda, zgoda warunkowa, stop lub zakaz.	Decyzja osoby posiadającej mandat.
7. Produkcja	Kontrolowane uruchomienie, monitoring, instrukcja, rollback i termin przeglądu.	Odbiór i zapis w rejestrze.
8. Utrzymanie	Monitoring jakości, kosztu, incydentów, zmian i zgodności.	Okresowa ponowna ocena lub wycofanie.

10. Progi ryzyka i terminy decyzji

Próg	Przykładowe cechy	Decyzja i orientacyjny termin
Niski	Dane publiczne lub zwykłe wewnętrzne, brak działania w systemach, wynik weryfikowany przez człowieka.	Automatyczna zgoda z katalogu albo właściciel procesu. [1 DZIEŃ ROBOCZY].
Podwyższony	Dane poufne, RAG, rekomendacje biznesowe, integracje tylko do odczytu lub ograniczony wpływ.	Właściciel biznesowy i danych po konsultacji z IT, Security, DPO lub Legal. [5 DNI ROBOCZYCH].
Wysoki	Dane wrażliwe, decyzje dotyczące ludzi, operacje zapisu, działania finansowe, wysoka autonomia lub trudny do odwrócenia skutek.	Komitet lub osoba akceptująca ryzyko po formalnej ocenie i testach. [15 DNI ROBOCZYCH].

Próg	Przykładowe cechy	Decyzja i orientacyjny termin
Niedopuszczalny	Praktyka zakazana prawem albo ryzyko, którego nie można zredukować do akceptowalnego poziomu.	Zakaz lub stop. Eskalacja do [ORGAN DECYZYJNY].

Wartości i terminy należy dostosować do skali organizacji. Każdy próg musi mieć jedną wskazaną rolę posiadającą prawo do ostatecznej decyzji. Brak odpowiedzi w terminie nie oznacza automatycznej zgody.

11. Role i odpowiedzialności

Rola	Decyduje lub odpowiada za	Nie przerzuca dalej
Zarząd lub sponsor	Apetyt na ryzyko, zasoby i nadzór nad systemem zarządzania AI.	Odpowiedzialności za ustanowienie skutecznych zasad.
Właściciel przypadku użycia	Cel, miernik sukcesu, jakość wyniku, użytkowników i skutek błędu.	Odpowiedzialności za rezultat biznesowy.
Właściciel danych	Klasyfikację, dostęp, minimalizację, retencję i dopuszczalne użycie danych.	Decyzji o zakresie danych.
IT lub Product	Architekturę, integracje, utrzymanie, wersje, monitoring i wycofanie.	Własności operacyjnej rozwiązania.
Security	Model zagrożeń, wymagane kontrole, testy i ryzyko cyberbezpieczeństwa.	Akceptacji ryzyka biznesowego.
DPO lub Legal	Prywatność, podstawę prawną, umowy, AI Act i inne obowiązki.	Decyzji technicznych poza swoim mandatem.
Procurement lub Vendor Management	Ocenę dostawcy, warunki umowne, podwykonawców, region i zakończenie usługi.	Oceny rzeczywistego przypadku użycia.
Risk Owner	Akceptację ryzyka rezydualnego w granicach przyznanego mandatu.	Decyzji bez zapisania warunków i terminu przeglądu.
Użytkownik	Korzystanie zgodnie z zakresem, weryfikację wyników i zgłaszanie problemów.	Odpowiedzialności za świadome użycie narzędzia.

12. Wymagania przed uruchomieniem

Zakres wymagań jest proporcjonalny do ryzyka, lecz przed uruchomieniem rozwiązania należy udokumentować co najmniej:

☐ właściciela przypadku, właściciela danych i właściciela utrzymania
☐ cel, użytkowników, dane wejściowe, wyniki i niedozwolone zachowania
☐ rolę organizacji i wynik klasyfikacji regulacyjnej
☐ ocenę dostawcy, umowy, retencji, regionu, podwykonawców i użycia danych do trenowania
☐ architekturę, granice zaufania, integracje, uprawnienia i przepływy danych
☐ model zagrożeń i testy obejmujące błędne, złośliwe i nieoczekiwane wejścia
☐ mechanizm weryfikacji źródeł i ograniczenia halucynacji

☐ nadzór człowieka adekwatny do skutku decyzji lub działania
☐ logowanie wersji modelu, źródeł, wywołań narzędzi, decyzji i zatwierdzeń
☐ limity kosztu, liczby kroków, czasu działania i częstotliwości operacji
☐ monitoring jakości, bezpieczeństwa, kosztu, zmian modelu i incydentów
☐ procedurę awaryjnego wyłączenia, rollbacku, odtworzenia i wycofania
☐ instrukcję użytkownika, wymagane szkolenie i kryteria ponownej oceny

13. Dodatkowe wymagania dla agentów AI

- Treści użytkownika, dokumentów, stron internetowych i wiadomości są traktowane jako niezaufane dane, a nie instrukcje systemowe.
- Każde narzędzie i usługa niezależnie weryfikuje tożsamość, zakres dostępu, własność obiektu i uprawnienie do operacji.
- Uprawnienia są ograniczone do konkretnego zadania, rekordu, czasu i środowiska.
- Operacje o wysokim wpływie wymagają podglądu oraz świadomego zatwierdzenia przez upoważnioną osobę.
- Dostęp do internetu jest oddzielony od sekretów i danych poufnych; miejsca docelowe i parametry wyjściowe podlegają kontroli.
- Agent posiada limit kroków, kosztu i czasu oraz techniczny wyłącznik awaryjny.
- Logi pozwalają odtworzyć wejście, pobrane źródła, decyzje, wywołania narzędzi, wynik i osobę zatwierdzającą.

14. Zgłaszanie błędów i incydentów

Każda osoba niezwłocznie zgłasza podejrzenie ujawnienia danych, błędnej decyzji, nieautoryzowanej operacji, manipulacji wynikiem, dyskryminacji, nieprawidłowego oznaczenia treści, nadmiernego kosztu, utraty kontroli nad agentem albo użycia AI poza zatwierdzonym zakresem.

Kanał	Dane do uzupełnienia
Formularz lub system zgłoszeń	[ADRES LUB NAZWA SYSTEMU]
Adres e mail	[ADRES ZESPOŁU]
Telefon dla zdarzeń pilnych	[NUMER TELEFONU]
Właściciel procesu	[ROLA LUB ZESPÓŁ]
Docelowy czas reakcji	[CZAS WG KRYTYCZNOŚCI]

Zgłaszający powinien, o ile jest to bezpieczne, przerwać działanie, zachować identyfikator rozmowy, czas, model, prompt, załączniki, wynik i wykonane akcje. Nie należy usuwać dowodów ani samodzielnie kontaktować się z osobą podejrzewaną o atak.

Zgłoszenie w dobrej wierze, w tym ujawnienie niezatwierdzonego zastosowania AI, nie może być zniechęcane. Priorytetem jest ograniczenie szkody, poznanie rzeczywistego użycia i bezpieczne uporządkowanie procesu, z zastrzeżeniem odpowiedzialności za świadome naruszenia.

15. Ponowna ocena i wycofanie

Każdy wpis w rejestrze zawiera datę kolejnej oceny. Domyślna częstotliwość wynosi [12 MIESIĘCY DLA NISKIEGO RYZYKA], [6 MIESIĘCY DLA PODWYŻSZONEGO] i [3 MIESIĄCE DLA WYSOKIEGO] albo częściej, jeżeli wymagają tego przepisy, umowa lub decyzja właściciela ryzyka.

Ponowna ocena jest wymagana także po:

- zmianie modelu, dostawcy, warunków usługi, regionu lub zasad retencji
- zmianie celu, grupy użytkowników, skali, danych, integracji, narzędzi lub autonomii
- istotnym incydencie, błędzie, skardze, wyniku audytu lub sygnale o dyskryminacji
- nowej podatności, zmianie profilu zagrożeń albo zmianie zabezpieczeń
- zmianie prawa, wytycznych, klasyfikacji AI Act lub roli organizacji
- spadku jakości, wzroście kosztu, braku właściciela albo utracie możliwości utrzymania rozwiązania

Rozwiązanie jest wycofywane, gdy nie ma właściciela, nie spełnia warunków decyzji, nie można skutecznie monitorować ryzyka, przestało realizować cel albo jego ryzyko przekracza przyjęty apetyt organizacji.

16. Wyjątki

Wyjątek od polityki wymaga pisemnego wniosku wskazującego cel, zakres, czas obowiązywania, właściciela, ryzyko, kontrole kompensacyjne i plan zakończenia. Wyjątek zatwierdza [ROLA LUB ORGAN] po konsultacji z właściwymi jednostkami. Wyjątek nie może zezwalać na praktykę zakazaną prawem.

17. Kompetencje i AI literacy

Organizacja zapewnia szkolenia odpowiednie do roli, wiedzy, sposobu korzystania z AI oraz wpływu systemu na osoby. Minimalny program obejmuje ochronę danych, weryfikację wyników, ograniczenia modeli, prompt injection, bezpieczne korzystanie z narzędzi, zgłaszanie incydentów i zasady tej polityki.

Właściciele przypadków użycia, programiści, administratorzy, osoby zatwierdzające wyniki i członkowie organów decyzyjnych otrzymują szkolenie rozszerzone, dopasowane do ich odpowiedzialności.

18. Zgodność z AI Act i innymi wymaganiami

Dla każdego zgłaszanego rozwiązania organizacja ustala, czy rozwiązanie spełnia definicję systemu AI, jaką rolę pełni organizacja, do jakiego przypadku użycia system jest przeznaczony oraz jakie obowiązki prawne mają zastosowanie. Analiza obejmuje w szczególności praktyki zakazane, systemy wysokiego ryzyka, wymagania przejrzystości, modele ogólnego przeznaczenia, obowiązki podmiotów stosujących oraz ewentualną ocenę wpływu na prawa podstawowe.

Klasyfikacja AI Act nie zastępuje oceny ryzyka cyberbezpieczeństwa, prywatności i ryzyka biznesowego. System spoza kategorii wysokiego ryzyka może nadal powodować znaczącą szkodę dla organizacji lub osób.

Stan prawny i harmonogram stosowania przepisów należy zweryfikować w dniu przyjęcia oraz podczas każdego przeglądu polityki. Na dzień 10.09.2026 r. AI Act jest stosowany etapowo, a harmonogram uwzględnia zmiany wprowadzone przez AI Omnibus.

19. Egzekwowanie i odpowiedzialność

Naruszenie polityki może skutkować ograniczeniem dostępu, wycofaniem rozwiązania, działaniami naprawczymi albo konsekwencjami przewidzianymi w regulacjach wewnętrznych i prawie. Ocena uwzględnia charakter naruszenia, intencję, skutek, współpracę przy ograniczeniu szkody i szybkość zgłoszenia.

20. Zatwierdzenie polityki

Rola	Imię i nazwisko	Decyzja	Data
Właściciel polityki	[UZUPEŁNIJ]	[AKCEPTACJA]	[DATA]
Security	[UZUPEŁNIJ]	[AKCEPTACJA]	[DATA]
DPO lub Legal	[UZUPEŁNIJ]	[AKCEPTACJA]	[DATA]
Zarząd lub organ zatwierdzający	[UZUPEŁNIJ]	[AKCEPTACJA]	[DATA]

Załącznik 1. Rejestr zatwierdzonych narzędzi

Narzędzie i właściciel	Status i zakres	Dane i konta	Kontrole i przegląd
[NAZWA WERSJA DOSTAWCA] Właściciel [ROLA]	[ZATWIERDZONE WARUNKOWE ZABRONIONE] Funkcje [ZAKRES]	Klasy danych [ZAKRES] Konta [SŁUŻBOWE INNE]	Retencja [OKRES] Trening [TAK NIE] Przegląd [DATA]
[UZUPEŁNIJ]	[UZUPEŁNIJ]	[UZUPEŁNIJ]	[UZUPEŁNIJ]
[UZUPEŁNIJ]	[UZUPEŁNIJ]	[UZUPEŁNIJ]	[UZUPEŁNIJ]

Załącznik 2. Formularz zgłoszenia przypadku użycia

Nazwa przypadku użycia

.....

.....

Właściciel biznesowy i jednostka

.....

.....

Problem i oczekiwany rezultat

.....

.....

Użytkownicy i osoby dotknięte wynikiem

.....

.....

Dane wejściowe i ich klasyfikacja

.....

.....

Wynik systemu i sposób jego wykorzystania

.....

.....

Narzędzie model dostawca i wersja

.....

.....

Integracje uprawnienia i możliwe działania

.....

.....

Czy system korzysta z RAG pamięci internetu lub narzędzi

.....

.....

Możliwy skutek błędu lub nadużycia

.....

.....

Proponowany nadzór człowieka

.....

.....

Planowany termin i skala pilotażu

.....

.....

Załącznik 3. Karta analizy ryzyka AI

Obszar	Pytania kontrolne	Wynik i dowód
Regulacja	Czy to system AI? Jaka jest rola organizacji? Jaki przypadek użycia? Czy występuje praktyka zakazana, high risk, przejrzystość, GPAI lub FRIA?	[UZUPEŁNIJ]
Prawa i prywatność	Kogo dotyczy wynik? Jakie dane są przetwarzane? Czy potrzebna jest DPIA? Czy istnieje ryzyko dyskryminacji lub ograniczenia praw?	[UZUPEŁNIJ]
Cyberbezpieczeństwo	Jakie są granice zaufania, dane, uprawnienia, integracje, podatności, scenariusze nadużyć, monitoring i reakcja?	[UZUPEŁNIJ]
Biznes	Jaki jest skutek błędu, zasięg szkody, odwracalność, koszt, właściciel ryzyka i warunki pilotażu?	[UZUPEŁNIJ]

Skala oceny

Ocena	Prawdopodobieństwo	Wpływ
1	Rzadkie lub mało prawdopodobne	Nieznaczny i łatwo odwracalny
2	Mało prawdopodobne	Ograniczony
3	Możliwe	Znaczący
4	Prawdopodobne	Poważny
5	Bardzo prawdopodobne lub oczekiwane	Krytyczny lub trudny do odwrócenia

Rejestr ryzyka

ID	Scenariusz i aktywo	P	W	Wynik	Kontrole właściciel i termin
R1	[ZAGROŻENIE PROWADZI DO SKUTKU]	[1 5]	[1 5]	[P X W]	[KONTROLE ROLA DATA]
R2	[UZUPEŁNIJ]				
R3	[UZUPEŁNIJ]				
R4	[UZUPEŁNIJ]				

Decyzja	Treść
Wynik początkowy	[NISKI PODWYŻSZONY WYSOKI NIEDOPUSZCZALNY]
Ryzyko rezydualne	[WYNIK PO KONTROLACH]
Decyzja	[ZGODA ZGODA WARUNKOWA STOP ZAKAZ]
Warunki i ograniczenia	[ZAKRES ŚRODOWISKO UŻYTKOWNICY DANE CZAS]
Właściciel ryzyka	[IMIĘ ROLA]
Termin ponownej oceny	[DATA LUB ZDARZENIE]

Załącznik 4. Checklista przed uruchomieniem

☐ Przypadek użycia ma właściciela biznesowego, danych i utrzymania.
☐ Cel, użytkownicy, dane, źródła i oczekiwany wynik są opisane.
☐ Rola organizacji i klasyfikacja AI Act zostały zapisane.
☐ Ocena prywatności, umowy i dostawcy została zakończona.
☐ Architektura, granice zaufania i przepływy danych są udokumentowane.
☐ Uprawnienia odpowiadają konkretnym zadaniom i są technicznie egzekwowane.
☐ Przeprowadzono testy funkcjonalne, bezpieczeństwa i testy specyficzne dla AI.
☐ Zweryfikowano prompt injection, wycieki, poisoning, halucynacje i nadużycia narzędzi.
☐ Ustalono nadzór człowieka i działania wymagające zatwierdzenia.
☐ Logi pozwalają odtworzyć źródła, wywołania, wynik i decyzję.
☐ Działają limity kosztu, kroków, czasu i częstotliwości.
☐ Istnieje monitoring, procedura incydentowa, wyłącznik i rollback.
☐ Użytkownicy otrzymali instrukcję i szkolenie.
☐ Ryzyko rezydualne zostało zaakceptowane przez właściwą osobę.
☐ W rejestrze zapisano zakres, warunki i termin ponownej oceny.

Załącznik 5. Formularz zgłoszenia błędu lub incydentu

Pole	Informacja
Data i osoba zgłaszająca	[UZUPEŁNIJ]
Narzędzie system model i wersja	[UZUPEŁNIJ]
Identyfikator rozmowy lub operacji	[UZUPEŁNIJ]
Opis oczekiwanego i rzeczywistego działania	[UZUPEŁNIJ]
Dane źródła prompt i załączniki	[UZUPEŁNIJ LUB WSKAŻ LOKALIZACJĘ DOWODÓW]
Działania wykonane przez system	[UZUPEŁNIJ]
Osoby systemy i dane objęte zdarzeniem	[UZUPEŁNIJ]
Działania podjęte natychmiast	[UZUPEŁNIJ]
Czy rozwiązanie zostało zatrzymane	[TAK NIE NIE DOTYCZY]
Właściciel obsługi i status	[UZUPEŁNIJ]

Załącznik 6. Źródła i dokumenty powiązane

- AI Act Service Desk Harmonogram wdrażania AI Act
- AI Act Service Desk Artykuł 4 AI literacy
- Rozporządzenie UE 2024 1689 w EUR Lex
- NIST AI Risk Management Framework Generative AI Profile
- OWASP Top 10 for LLM and Generative AI Applications
- OWASP AI Agent Security Cheat Sheet
- NIST Secure Software Development Framework

Dokumenty organizacji do powiązania

- [POLITYKA BEZPIECZEŃSTWA INFORMACJI]
- [POLITYKA KLASYFIKACJI I OCHRONY DANYCH]
- [PROCEDURA ZARZĄDZANIA INCYDENTAMI]
- [STANDARD SSDLC I ZARZĄDZANIA ZMIANĄ]
- [POLITYKA ZAKUPÓW I OCENY DOSTAWCÓW]
- [POLITYKA OCHRONY DANYCH OSOBOWYCH]
- [REJESTR SYSTEMÓW I PRZYPADKÓW UŻYCIA AI]