

Bezpieczeństwo AI w firmie

Praktyczne minimum dla pracodawcy i pracowników

Dane, AI Act, RODO, zasady firmowe i najczęstsze ryzyka w codziennej pracy

AI może oszczędzać czas, ułatwiać analizę informacji i wspierać pracowników. Bez jasnych zasad może jednak prowadzić do ujawnienia danych, błędnych decyzji i problemów organizacyjnych.

Ten poradnik pokazuje, jak uporządkować korzystanie z AI w firmie krok po kroku. Bez straszenia, bez technicznego żargonu i bez obietnicy, że jeden dokument rozwiąże wszystkie problemy prawne.

Stan informacji: 23 lipca 2026 r.



02 ZANIM ZACZNIESZ: CZYM JEST TEN PORADNIK?



AI jest już używana w wielu firmach do pisania wiadomości, przygotowywania ofert, streszczania dokumentów, tworzenia grafik, analizy danych i obsługi klientów. Często dzieje się to szybciej, niż firma zdąży ustalić zasady. Pracownicy wybierają narzędzia samodzielnie, korzystają z prywatnych kont i uczą się metodą prób i błędów.

Ten poradnik ma pomóc właścicielom firm, menedżerom i pracownikom uporządkować ten obszar. Wyjaśnia, co warto sprawdzić przed dopuszczeniem narzędzia AI do pracy, jak chronić dane, kiedy wynik musi zostać sprawdzony przez człowieka oraz które zastosowania wymagają szczególnej ostrożności.

Nie jest to pełny podręcznik prawny, audyt zgodności ani indywidualna porada prawna. Przeczytanie poradnika nie oznacza automatycznie, że firma spełnia wszystkie wymagania AI Act, RODO, prawa pracy lub przepisów branżowych. Zakres obowiązków zależy od tego, jakiego narzędzia firma używa, w jakim celu, na jakich danych i jaki wpływ może mieć wynik.

Najważniejszy cel jest prosty: firma powinna wiedzieć, gdzie AI jest używana, jakie informacje trafiają do narzędzi i kto odpowiada za sprawdzenie wyniku. Dopiero później warto rozwijać bardziej rozbudowane procedury.

Bezpieczne korzystanie z AI nie zaczyna się od zakazywania narzędzi. Zaczyna się od sprawdzenia, gdzie AI jest używana, jakie dane do niej trafiają i kto odpowiada za wynik.

Najważniejsze wnioski

- AI jest już częścią codziennej pracy, nawet gdy firma formalnie jej nie wdrożyła.
- Pierwszym celem nie jest perfekcyjna zgodność, lecz widoczność i kontrola.
- Poradnik daje praktyczne minimum, ale nie zastępuje audytu ani porady prawnej.

Co zrobić w swojej firmie

Odpowiedz na trzy pytania:

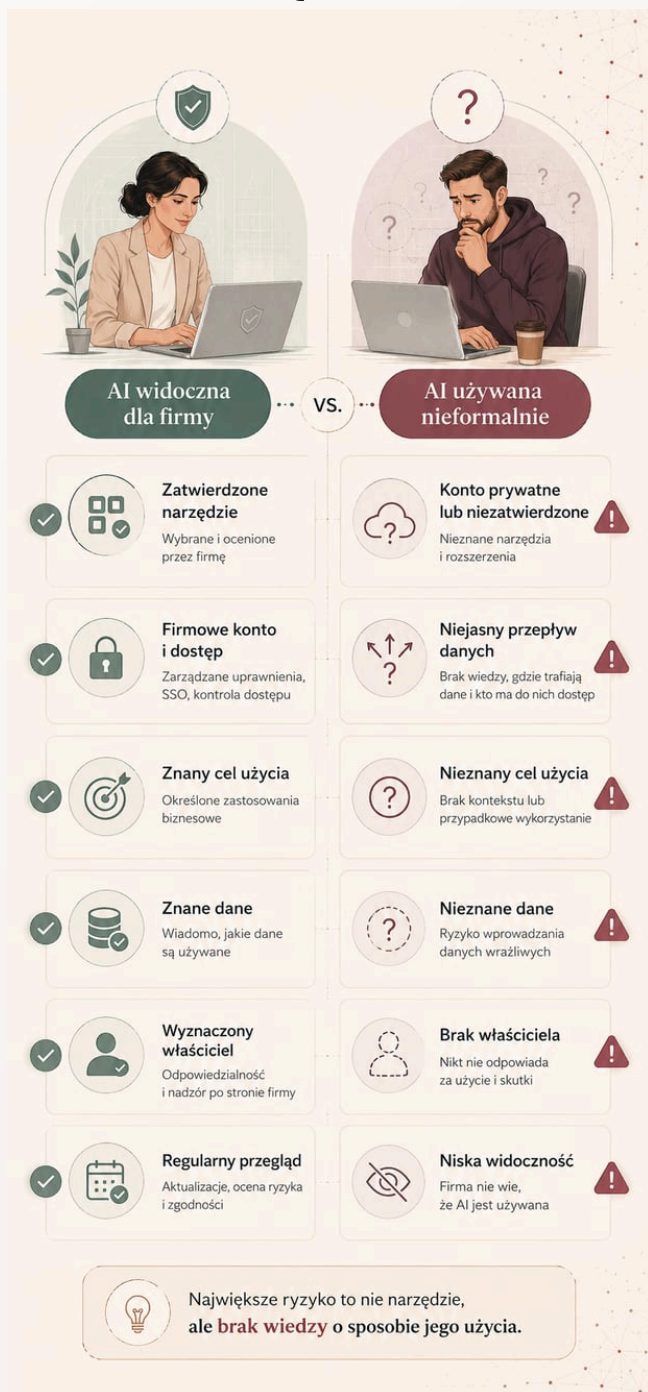
1. Kto korzysta z AI?
2. Do jakich zadań?
3. Jakie dane są wprowadzane do narzędzi?

ŹRÓDŁA

- Komisja Europejska, AI Act, podejście oparte na ryzyku
- AI Act Service Desk, art. 4, kompetencje w zakresie AI

03

AI W FIRMIE: CO, NAPRAWDĘ TRZEBA UPORZĄDKOWAĆ?



Największym problemem nie musi być samo narzędzie AI. Często jest nim brak wiedzy o tym, jak narzędzie jest używane. Zarząd może nie podjąć żadnej decyzji o wdrożeniu AI, a pracownicy i tak korzystają z niej do codziennych zadań.

Takie nieformalne używanie narzędzi nazywa się często Shadow AI. Chodzi o sytuację, w której pracownik wykorzystuje prywatne lub niezatwierdzone konto do spraw służbowych, a firma nie zna zakresu użycia, rodzaju przekazywanych danych ani ustawień usługi.

Przykład: pracownik chce szybko przygotować odpowiedź na reklamację. Kopiuje do prywatnego asystenta AI fragment korespondencji z klientem. W wiadomości znajdują się dane kontaktowe, numer zamówienia, opis problemu i niepubliczne warunki handlowe. Narzędzie przygotowuje dobry szkic, ale firma nie sprawdziła wcześniej zasad przetwarzania danych, retencji ani możliwości wykorzystania treści przez dostawcę.

Rozwiązaniem nie jest automatyczny zakaz wszystkiego. Twardy zakaz bez bezpiecznej alternatywy może jedynie przenieść korzystanie z AI poza widoczność firmy. Potrzebne są jasne zasady, lista zatwierdzonych narzędzi i prosty sposób zgłaszania nowych potrzeb.

Firma powinna prowadzić podstawową mapę zastosowań AI. Nie musi to być rozbudowany system. Na początek wystarczy tabela zawierająca nazwę narzędzia, cel użycia, użytkowników, rodzaj danych, osobę odpowiedzialną i datę ostatniego przeglądu.

Najważniejsze wnioski

- Brak formalnego wdrożenia nie oznacza, że pracownicy nie korzystają z AI.
- Ryzyko rośnie, gdy firma nie zna narzędzi, danych i sposobu wykorzystania wyników.
- Widoczność jest skuteczniejsza niż zakaz, którego nikt nie przestrzega.

Co zrobić w swojej firmie

Przeprowadź krótką, nieoskarżającą ankietę w zespole:

- z jakich narzędzi AI korzystasz,
- do jakich zadań,
- jakich informacji potrzebujesz w pracy,
- czego brakuje Ci, aby korzystać z AI bezpiecznie.

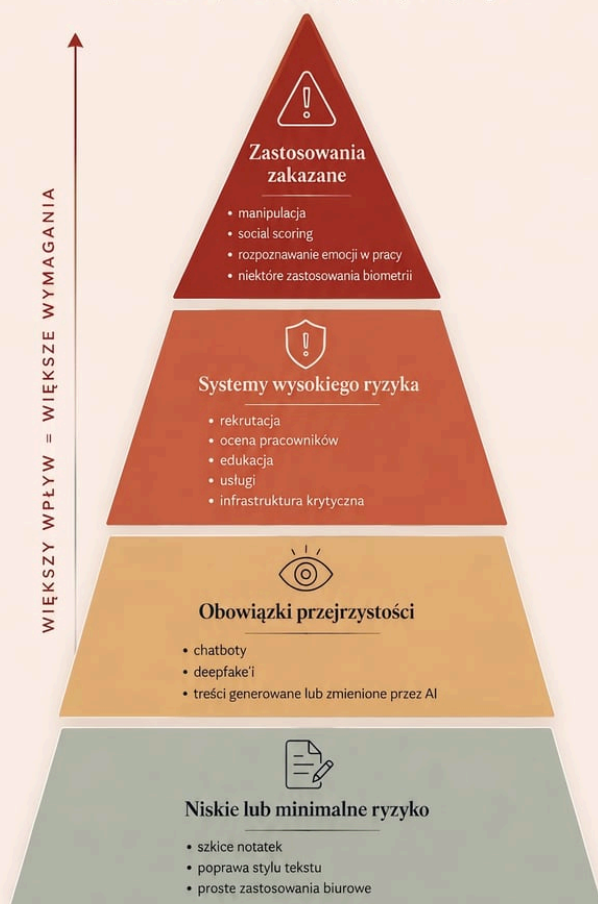
ŹRÓDŁA

- NIST, Artificial Intelligence Risk Management Framework
- OWASP, GenAI Security Project

04 AI ACT W PRAKTYCE: JAK DZIAŁA PODEJŚCIE OPARTE NA RYZYKU?

AI Act: cztery poziomy ryzyka

Im większy możliwy wpływ systemu na zdrowie, bezpieczeństwo i prawa człowieka, tym większe wymagania.



Kluczowe pytanie: nie tylko jakiego narzędzia używamy, ale przede wszystkim do czego go używamy i jaki może być skutek błędu?

AI Act to unijne rozporządzenie dotyczące sztucznej inteligencji. Nie traktuje każdego zastosowania AI tak samo. Podstawowa zasada brzmi: im większy możliwy wpływ systemu na zdrowie, bezpieczeństwo i prawa człowieka, tym większe wymagania.

Dla zwykłej firmy ważniejsze od zapamiętywania numerów artykułów jest zrozumienie czterech poziomów ryzyka.

1. Zastosowania zakazane

Są to praktyki uznane za niedopuszczalne, na przykład szkodliwa manipulacja, social scoring, niektóre formy rozpoznawania emocji w miejscu pracy oraz określone zastosowania biometrii.

2. Systemy wysokiego ryzyka

Dotyczą obszarów, w których AI może poważnie wpłynąć na człowieka. Przykłady obejmują rekrutację, ocenę pracowników, dostęp do edukacji, niektórych usług i infrastruktury krytycznej.

3. Zastosowania objęte obowiązkami przejrzystości

W określonych sytuacjach trzeba poinformować człowieka, że kontaktuje się z AI lub że treść została sztucznie wygenerowana albo zmieniona. Dotyczy to między innymi chatbotów, deepfake'ów i niektórych treści dotyczących spraw publicznych.

4. Zastosowania o niskim lub minimalnym ryzyku

W tej grupie mieści się większość prostych zastosowań biurowych. Przykładem może być szkic wewnętrznej notatki lub poprawa stylu tekstu. Niskie ryzyko nie oznacza jednak, że można ignorować RODO, tajemnicę przedsiębiorstwa, bezpieczeństwo i kontrolę jakości.

Kluczowe pytanie brzmi więc nie tylko: „jakiego narzędzia używamy?”, ale przede wszystkim: „do czego go używamy i jaki może być skutek błędu?”.

Najważniejsze wnioski

- AI Act nie zakazuje zwykłego korzystania z AI w firmie.
- Ten sam model może być użyty do zadania niskiego albo wysokiego ryzyka.
- O poziomie ostrożności decyduje cel, dane i wpływ wyniku.

Co zrobić w swojej firmie

Przypisz każdemu zastosowaniu AI jeden poziom:

- pomocnicze,
- wrażliwe,
- wpływające na decyzje dotyczące ludzi.

Zastosowania z trzeciej grupy powinny wymagać wcześniejszej oceny i zgody odpowiedzialnej osoby.

ŹRÓDŁA

- Komisja Europejska, cztery poziomy ryzyka AI Act
- AI Act, rozporządzenie (UE) 2024/1689

05 GDZIE FIRMĄ POWINNA ZACHOWAĆ SZCZEGÓLNA OSTROŻNOŚĆ?



Największej ostrożności wymagają zastosowania, w których wynik AI może wpłynąć na człowieka, jego pracę, prawa, zdrowie, pieniądze albo dostęp do ważnej usługi.

Szczególnie wrażliwy jest obszar HR. AI Act wskazuje jako wysokiego ryzyka między innymi systemy wykorzystywane do rekrutacji, filtrowania CV, oceny kandydatów, przydzielania zadań na podstawie cech lub zachowania, oceny wydajności, podejmowania decyzji o awansie oraz zakończeniu zatrudnienia. Nie oznacza to, że każda pomoc AI przy napisaniu ogłoszenia o pracę jest systemem wysokiego ryzyka. Granica pojawia się wtedy, gdy narzędzie zaczyna oceniać ludzi lub realnie wpływać na decyzję.

AI Act zakazuje również określonych praktyk. Z punktu widzenia pracodawcy szczególnie ważny jest zakaz rozpoznawania emocji w miejscu pracy, z ograniczonymi wyjątkami związanymi z medycyną lub bezpieczeństwem. Ryzykowne są też systemy manipulujące ludźmi, wykorzystujące ich podatność albo tworzące nieuzasadniony scoring społeczny.

Drugim ważnym obowiązkiem są kompetencje w zakresie AI, określane w rozporządzeniu jako AI literacy. Firma powinna zadbać, aby osoby korzystające z AI rozumiały możliwości i ograniczenia narzędzi, potrafiły chronić dane i wiedziały, kiedy wynik wymaga sprawdzenia. Przepisy nie narzucają jednej obowiązkowej liczby godzin szkolenia ani jednego certyfikatu. Zakres przygotowania powinien odpowiadać roli pracownika i ryzyku danego zastosowania.

Pracownik używający AI do prostych szkiców potrzebuje innych zasad niż osoba korzystająca z narzędzia przy rekrutacji, poradach prawnych lub analizie danych zdrowotnych.

Najważniejsze wnioski

- AI nie powinna samodzielnie decydować o zatrudnieniu, ocenie lub zwolnieniu człowieka.
- HR, zdrowie, prawo i finanse wymagają mocniejszej kontroli.
- Szkolenie powinno być dopasowane do roli, a nie identyczne dla wszystkich.

Co zrobić w swojej firmie

Wprowadź zasadę:

Każde użycie AI w rekrutacji, ocenie pracowników lub decyzjach wobec klientów wymaga wcześniejszej konsultacji z osobą odpowiedzialną za HR, prawo, dane lub bezpieczeństwo.

ŹRÓDŁA

- Komisja Europejska, przykłady systemów wysokiego ryzyka i praktyk zakazanych
- AI Act Service Desk, art. 4
- AI Act, załącznik III

06 STAN PRAWNY NA 23.07.2026: AI ACT I POLSKA USTAWA

AI Act obowiązuje w całej Unii Europejskiej, ale jego przepisy są stosowane etapami. Rozporządzenie weszło w życie 1 sierpnia 2024 r. Od 2 lutego 2025 r. stosuje się zakazy określonych praktyk oraz obowiązek dbania o kompetencje osób korzystających z AI. Od 2 sierpnia 2025 r. stosowane są przepisy dotyczące modeli AI ogólnego przeznaczenia i części systemu nadzoru. Większość pozostałych przepisów ma zastosowanie od 2 sierpnia 2026 r.

29 czerwca 2026 r. Rada UE ostatecznie przyjęła AI Omnibus, który upraszcza AI Act i przesuwając terminy stosowania przepisów dla części systemów wysokiego ryzyka. Przyjęty tekst przewiduje 2 grudnia 2027 r. dla samodzielnych systemów wysokiego ryzyka z obszarów załącznika III oraz 2 sierpnia 2028 r. dla systemów wysokiego ryzyka wbudowanych w produkty z załącznika I. Na 23 lipca 2026 r. EUR-Lex udostępnia ostateczny tekst legislacyjny PE 30 2026 REV 1, ale akt nie został jeszcze opublikowany w Dzienniku Urzędowym UE i nie wszedł w życie.

W Polsce parlamentarny etap prac został zakończony. 3 lipca 2026 r. Sejm rozpatrzył stanowisko Senatu, przyjął część poprawek i ustalił ostateczny tekst ustawy o systemach sztucznej inteligencji. Tego samego dnia ustawę przekazano Prezydentowi do podpisu. Na 23 lipca 2026 r. ustawa nie figuruje na oficjalnej liście ustaw podpisanych w lipcu i nie została odnaleziona w Dzienniku Ustaw. Do czasu podpisania i ogłoszenia polska ustawa nie obowiązuje.

Ostateczny tekst parlamentarny przewiduje utworzenie Komisji Rozwoju i Bezpieczeństwa Sztucznej Inteligencji jako głównego organu nadzoru rynku i pojedynczego punktu kontaktowego. Komisja ma między innymi prowadzić kontrole, wydawać zalecenia, decyzje, wyjaśnienia i odpłatne opinie indywidualne, zarządzać piaskownicami regulacyjnymi oraz prowadzić działania edukacyjne. Tekst przewiduje także mechanizmy naprawcze. Po kontroli Komisja może wydać zalecenia z terminem na usunięcie nieprawidłowości, a w postępowaniu może zastosować ostrzeżenie. Przewidziano również możliwość zawarcia układu, który w określonych warunkach pozwala złagodzić sankcję.

Art. 127 przewiduje wejście ustawy w życie po 14 dniach od ogłoszenia, z wyjątkiem art. 125 ust. 4, który ma wejść w życie następnego dnia po ogłoszeniu, oraz art. 8–18 i rozdziałów 3–5, 8 i 9, które mają wejść w życie po 3 miesiącach od ogłoszenia. Do tego czasu firmy powinny przygotowywać się przede wszystkim na podstawie bezpośrednio obowiązującego AI Act, RODO, prawa pracy, zasad cyberbezpieczeństwa i przepisów branżowych.

AI Act: harmonogram zmian



Najważniejsze wnioski

- Parlament zakończył prace nad polską ustawą 3 lipca 2026 r.
- Na 23 lipca ustawa oczekuje na podpis Prezydenta i nie została jeszcze ogłoszona.
- Dopóki nie nastąpi ogłoszenie, przepisy polskiej ustawy nie obowiązują.
- AI Act obowiązuje niezależnie od zakończenia polskiego procesu legislacyjnego.

Co zrobić w swojej firmie

Wyznacz osobę, która raz na kwartał sprawdzi:

- harmonogram stosowania AI Act,
- podpisanie i ogłoszenie polskiej ustawy,
- datę uruchomienia Komisji Rozwoju i Bezpieczeństwa Sztucznej Inteligencji,
- nowe wyjaśnienia Komisji Europejskiej, UODO i krajowego organu nadzoru,
- zmiany warunków korzystania z używanych narzędzi.

Nie czekaj z podstawowymi zasadami na wejście w życie polskiej ustawy. Już teraz uporządkuj narzędzia, dane, odpowiedzialność pracowników i sposób weryfikacji wyników.

ŹRÓDŁA

- Komisja Europejska, aktualna oś stosowania i upraszczania AI Act, aktualizacja 07.07.2026
- Rada UE, ostateczne przyjęcie AI Omnibus, 29.06.2026
- EUR-Lex, ostateczny tekst legislacyjny AI Omnibus, PE 30 2026 REV 1
- AI Act Service Desk, art. 113, wejście w życie i stosowanie
- Sejm RP, przebieg procesu legislacyjnego druku nr 2443
- Sejm RP, sprawozdanie komisji dotyczące poprawek Senatu, druk nr 2762
- Tekst ustawy ustalony ostatecznie po rozpatrzeniu poprawek Senatu, 3.07.2026 r., plik źródłowy: „Ustawa o systemach sztucznej inteligencji.2443_u.03.07.26.pdf”
- Prezydent RP, ustawy podpisane w lipcu 2026 r.
- Dziennik Ustaw, wykaz aktów ogłoszonych w 2026 r.

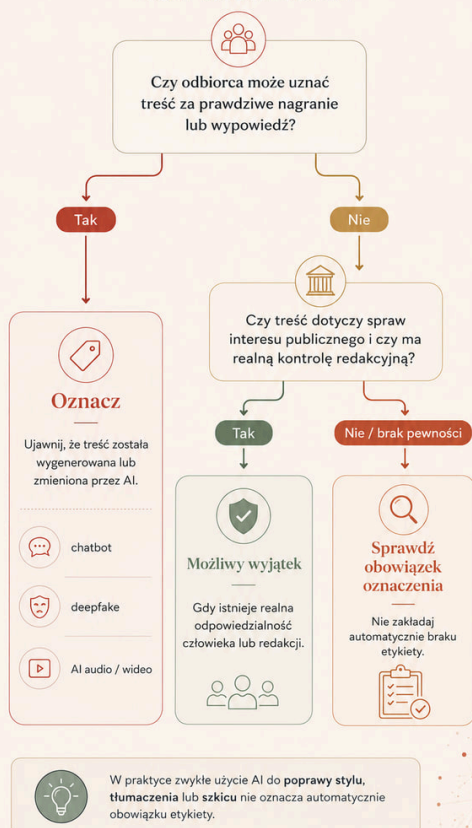
07

CZY TRZEBA OZNACZAĆ TREŚCI STWORZONE Z POMOCĄ AI?

Obowiązki przejrzystości z art. 50 stosuje się od 2 sierpnia 2026 r.

Kiedy oznaczać treści AI?

PROSTY SCHEMAT DECYZYJNY



Nie każda treść przygotowana z pomocą AI musi otrzymać etykietę „wygenerowano przez AI”. AI Act wprowadza konkretne obowiązki przejrzystości, ale ich zakres zależy od rodzaju systemu i treści.

Po pierwsze, osoba powinna zostać poinformowana, że rozmawia z systemem AI, jeżeli nie jest to oczywiste z sytuacji. Dotyczy to na przykład chatbota obsługującego klientów na stronie firmy.

Po drugie, dostawcy systemów generujących syntetyczny tekst, obraz, dźwięk lub wideo mają zapewniać techniczne oznaczenie wyników jako sztucznie wygenerowanych lub zmienionych. Jest to obowiązek po stronie dostawcy technologii, a nie wyłącznie użytkownika końcowego.

Po trzecie, firma wykorzystująca deepfake powinna ujawnić, że obraz, audio lub wideo zostały wygenerowane albo zmanipulowane. Dla materiałów artystycznych, satyrycznych i fikcyjnych sposób ujawnienia może być łagodniejszy, ale odbiorca nadal nie powinien być celowo wprowadzany w błąd.

Po czwarte, obowiązek może dotyczyć tekstu publikowanego w celu informowania społeczeństwa o sprawach interesu publicznego. Rozporządzenie przewiduje jednak wyjątek, gdy treść przeszła kontrolę człowieka lub redakcji, a konkretna osoba albo organizacja bierze odpowiedzialność za publikację.

W praktyce zwykłe użycie AI do poprawy stylu, skrócenia tekstu, tłumaczenia, stworzenia szkicu posta lub wariantów nagłówka nie oznacza automatycznie obowiązku dodania etykiety. Firma nadal powinna jednak sprawdzić fakty, cytaty, źródła, prawa do materiałów i zgodność przekazu z własnymi zasadami.

Najważniejsze wnioski

- Nie trzeba oznaczać każdego tekstu, przy którym pomagała AI.
- Deepfake i treści mogące wprowadzać odbiorcę w błąd wymagają szczególnej przejrzystości.
- Kontrola redakcyjna człowieka nie może być pozorna.

Co zrobić w swojej firmie

Ustal cztery zasady publikacji:

1. Każdy materiał zewnętrzny zatwierdza człowiek.
2. Deepfake jest zawsze oznaczany.
3. Treści prawne, zdrowotne i finansowe mają sprawdzone źródła.
4. Nie publikujemy fikcyjnych cytatów, ekspertów ani referencji.

ŹRÓDŁA

- AI Act Service Desk, art. 50
- Komisja Europejska, wytyczne dotyczące obowiązków przejrzystości z art. 50, 20.07.2026
- Komisja Europejska, FAQ dotyczące obowiązków przejrzystości z art. 50

AI, DANE I RODO: ZASADA MINIMUM DANYCH

Dane przed i po

Minimalizacja danych przed użyciem AI



RODO obowiązuje również wtedy, gdy dane osobowe są wpisywane do promptu, dodawane w załączniku lub udostępniane narzędziu przez integrację z pocztą, dyskiem albo systemem firmowym.

Najważniejsza zasada brzmi: do narzędzia należy przekazywać tylko te informacje, które są naprawdę potrzebne do wykonania zadania. Jeżeli AI ma poprawić styl wiadomości, zwykle nie potrzebuje imienia klienta, adresu, numeru zamówienia ani pełnej historii kontaktu. Dane można usunąć, zastąpić symbolami lub ograniczyć do niezbędnego fragmentu.

Pseudonimizacja oznacza zastąpienie bezpośrednich identyfikatorów, na przykład nazwiska określeniem „Klient A”. Nie zawsze daje jednak pełną anonimowość. Osobę nadal można rozpoznać z kontekstu, gdy tekst zawiera nazwę stanowiska, konkretną datę, opis nietypowej sytuacji albo inne charakterystyczne informacje.

Przed użyciem narzędzia firma powinna znać cel przetwarzania, podstawę prawną, zakres danych, okres przechowywania, odbiorców i zabezpieczenia. Przy rozwiązaniach mogących powodować wysokie ryzyko dla praw ludzi może być potrzebna ocena skutków dla ochrony danych. Dotyczy to szczególnie profilowania, szerokiej oceny osób, przetwarzania danych wrażliwych na dużą skalę lub rozbudowanego monitoringu.

Nie należy zakładać, że narzędzie jest bezpieczne tylko dlatego, że jest znane albo płatne. O bezpieczeństwie decydują również warunki usługi, ustawienia, umowa powierzenia danych, sposób zarządzania kontami i realny zakres dostępu.

Najważniejsze wnioski

- Prompt i załącznik mogą zawierać dane osobowe.
- Najbezpieczniejsze dane to te, których nie trzeba przekazywać.
- Pseudonimizacja zmniejsza ryzyko, ale nie zawsze oznacza anonimowość.

Co zrobić w swojej firmie

Wprowadź zasadę przed wysłaniem danych do AI:

Usuń nazwiska, dane kontaktowe, numery dokumentów i szczegóły, które nie są potrzebne do wykonania zadania.

W razie wątpliwości pracownik powinien przerwać operację i zapytać osobę odpowiedzialną za dane lub bezpieczeństwo.

ŹRÓDŁA

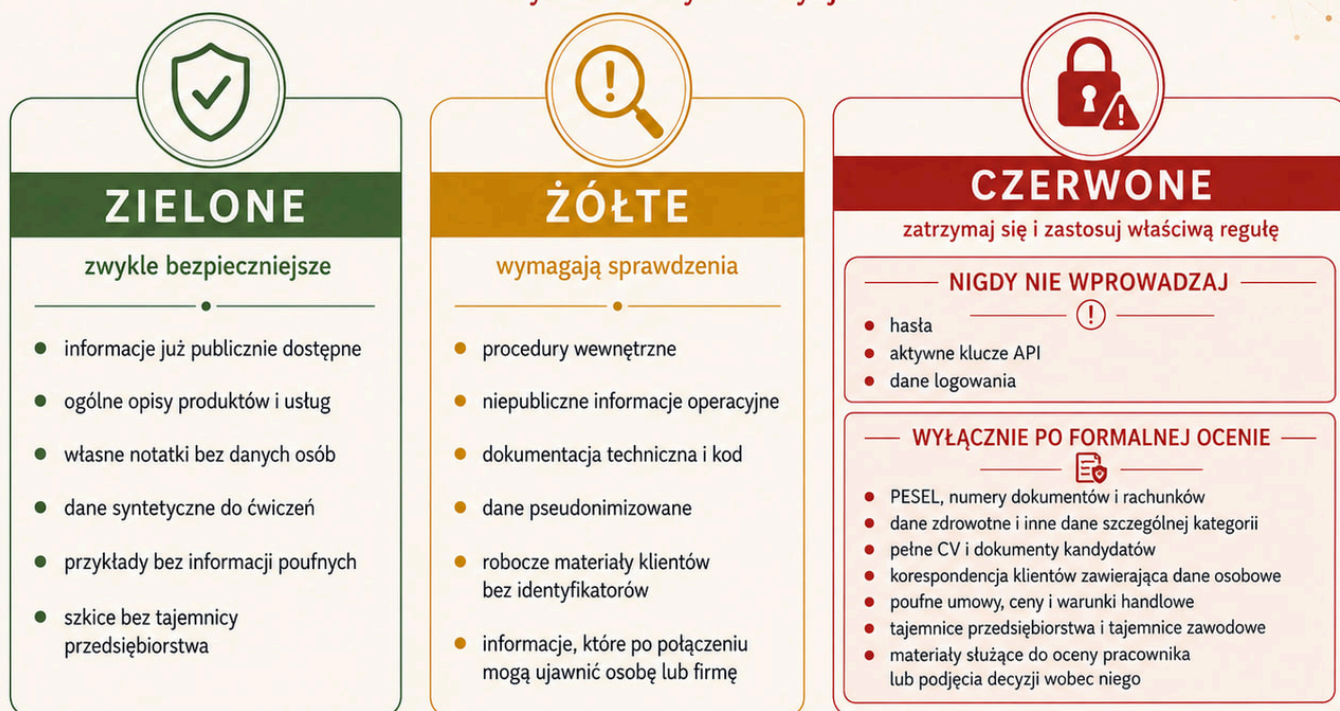
- Komisja Europejska, ochrona danych w fazie projektowania i domyślna ochrona danych
- Komisja Europejska, kiedy potrzebna jest DPIA
- EDPB, Opinion 28/2024 o danych osobowych i modelach AI

09 MACIERZ DANYCH: ZIELONE, ŻÓŁTE I CZERWONE

Prosty podział danych pomaga pracownikowi podjąć decyzję bez czytania długiej polityki. Kolory nie są oceną prawną każdego przypadku. Są praktycznym sygnałem, kiedy można działać, kiedy trzeba sprawdzić warunki, a kiedy należy przerwać i poprosić o zgodę.

Jakie dane można wprowadzać do AI?

— • Szybka matryca decyzji • —



Hasła, aktywne klucze API i dane logowania: nigdy. Pozostałe dane czerwone można przetwarzać wyłącznie w zatwierdzonym rozwiązaniu, po formalnej ocenie prawnej i bezpieczeństwa oraz z ograniczeniem danych do niezbędnego minimum.

Najważniejsze wnioski

- Największe ryzyko powstaje przy kopiowaniu całych dokumentów „dla wygody”.
- Niepubliczne nie zawsze oznacza zakazane, ale wymaga sprawdzenia.
- Hasła, aktywne klucze API i dane logowania są zawsze wyłączone.
- Pozostałe dane czerwone wymagają formalnej oceny, zatwierdzonego narzędzia i ograniczenia danych do minimum.

Co zrobić w swojej firmie

Wydrukuj jednostronicową listę danych zielonych, żółtych i czerwonych. W czerwonej kategorii wyraźnie oddziel dane, których nie wolno wprowadzać nigdy, od danych wymagających formalnej oceny. Umieść listę w polityce AI, intranecie i materiałach szkoleniowych.

ŹRÓDŁA

- RODO, zasada minimalizacji danych, art. 5
- Komisja Europejska, data protection by design and by default
- EDPB, Opinia 28/2024

10 DARMOWE KONTO, PŁATNA SUBSKRYPCJA CZY NARZĘDZIE FIRMOWE?

Cena subskrypcji nie mówi jeszcze, czy narzędzie nadaje się do pracy z danymi firmowymi. Płatne konto indywidualne może oferować więcej funkcji, ale nadal działać na warunkach usługi konsumenckiej. Rozwiązanie firmowe zwykle daje dodatkowe umowy, ustawienia administracyjne i kontrolę nad użytkownikami. Warto rozróżnić trzy poziomy.

Bezpłatne konto indywidualne

Jest tworzone i kontrolowane przez pracownika. Firma zwykle nie zarządza dostępem, nie może łatwo odebrać konta po odejściu pracownika i nie ma pełnej kontroli nad ustawieniami.

Płatne konto indywidualne

Może zapewniać lepszy model i więcej funkcji, ale nie staje się automatycznie usługą firmową. Nadal trzeba sprawdzić regulamin, zasady użycia danych, możliwość wyłączenia treningu, retencję i sposób usuwania treści.

Business, Enterprise, Workspace lub API

Nazwy różnią się zależnie od dostawcy. Takie usługi mogą oferować umowę powierzenia danych, centralną administrację, zarządzanie użytkownikami, uwierzytelnianie wieloskładnikowe, logi, kontrolę retencji i dodatkowe zabezpieczenia. Nadal nie zwalnia to firmy z prawidłowej konfiguracji i ograniczenia danych.

Przykładowo OpenAI wskazuje, że treści z usług indywidualnych mogą być wykorzystywane do ulepszania modeli, chyba że użytkownik zrezygnuje z tego w ustawieniach. Dla usług biznesowych i API firma deklaruje brak domyślnego trenowania na danych klientów oraz udostępnia dokumenty dotyczące przetwarzania danych. To przykład pokazujący, dlaczego trzeba sprawdzać konkretny produkt, a nie tylko nazwę dostawcy.

Przed zatwierdzeniem narzędzia firma powinna znać odpowiedzi na kilka podstawowych pytań.

Tabela porównawcza:

Obszar	Konto indywidualne	Konto firmowe
Właściciel konta	pracownik	organizacja
Zarządzanie użytkownikami	ograniczone	centralne
DPA	nie zawsze	często dostępne
Ustawienia danych	indywidualne	organizacyjne
Odebranie dostępu	utrudnione	zarządzane przez administratora

Najważniejsze wnioski

- Płatne konto indywidualne nie jest tym samym co konto firmowe.
- DPA, czyli umowa lub dodatek regulujący powierzenie i zasady przetwarzania danych, oraz funkcje administracyjne są ważniejsze niż sama cena subskrypcji.
- Nawet dobre narzędzie może zostać źle skonfigurowane.

Co zrobić w swojej firmie

Dla każdego narzędzia sprawdź i zapisz:

1. Czy dane są używane do trenowania modeli?
2. Czy dostępna jest umowa DPA?
3. Jak długo przechowywane są treści?
4. Gdzie dane są przetwarzane i czy są transferowane poza EOG?
5. Kto jest podwykonawcą dostawcy?
6. Czy administrator może dodawać i usuwać użytkowników?
7. Czy można włączyć MFA, ograniczyć funkcje i prowadzić audyt?
8. Jak usunąć lub wyeksportować dane po zakończeniu współpracy?

ŹRÓDŁA

- OpenAI, sposób używania danych w usługach indywidualnych i biznesowych
- OpenAI, Enterprise Privacy
- OpenAI, Data Processing Addendum

11 NAJCZĘSTSZE RYZYKA AI W FIRMIE, CZĘŚĆ 1

Większość problemów z AI w firmach nie wynika z zaawansowanych ataków. Częściej chodzi o codzienne skróty: prywatne konto, wklejenie zbyt dużej ilości danych albo wysłanie niesprawdzonej odpowiedzi klientowi.

1 Shadow AI

Pracownicy używają narzędzi bez wiedzy firmy. Organizacja nie wie wtedy, jakie dane są przekazywane, kto ma dostęp do historii rozmów i czy konto można odebrać po odejściu pracownika.

2 Ujawnienie danych poufnych lub osobowych

Do promptu trafia korespondencja klienta, umowa, CV, cennik, kod, notatka ze spotkania albo dane logowania. Nawet gdy narzędzie nie wykorzystuje danych do trenowania modelu, firma nadal musi ocenić podstawę przetwarzania, retencję, dostęp i zabezpieczenia.

3 Halucynacje i błędne informacje

Model może podać nieistniejący przepis, badanie, cytaty, funkcję produktu lub źródło. Odpowiedź bywa płynna i przekonująca, dlatego błąd może pozostać niezauważony.

4 Nadmierne zaufanie do wyniku

Pracownik przestaje samodzielnie oceniać odpowiedź, bo narzędzie zwykle działa dobrze. Ryzyko rośnie, gdy wynik wpływa na klienta, umowę, finanse, zdrowie albo decyzję wobec pracownika.

5 Prawa autorskie i poufność materiałów osób trzecich

Do AI mogą trafiać dokumenty, zdjęcia, projekty i teksty, których firma nie powinna dalej udostępniać. Problemem może być zarówno przekazanie materiału do narzędzia, jak i późniejsze wykorzystanie wygenerowanej treści bez sprawdzenia jej pochodzenia.

Nie każde ryzyko wymaga rozbudowanego systemu. Potrzebne są jednak jasne zasady, zatwierdzone narzędzia i obowiązek kontroli wyniku przed użyciem zewnętrznym.

5 najczęstszych ryzyk AI w firmie



Najczęściej zawodzi nie technologia, lecz codzienne skróty i brak zasad.

Najważniejsze wnioski

- Najczęstsze problemy wynikają z pośpiechu i braku zasad.
- Dobra odpowiedź AI nadal może zawierać poważny błąd.
- Narzędzie nie powinno otrzymywać więcej danych, niż potrzebuje do zadania.

Co zrobić w swojej firmie

Wprowadź trzy proste reguły:

- Nie używamy prywatnych kont do danych firmowych.
- Nie wklejamy pełnych dokumentów, jeżeli wystarczy krótki, oczyszczony fragment.
- Treść wysyłaną klientowi zawsze zatwierdza człowiek.

ŹRÓDŁA

- NIST, Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile
- OWASP, GenAI Security Project i Top 10 dla aplikacji LLM
- EDPB, Opinia 28/2024 dotycząca danych osobowych i modeli AI

12 NAJCZĘSTSZE RYZYKA AI W FIRMIE, CZĘŚĆ 2

Ryzyko rośnie, gdy AI nie tylko tworzy tekst, ale także czyta pliki, korzysta z internetu, łączy się z pocztą, dyskiem, CRM albo wykonuje działania w imieniu użytkownika.

6 Prompt injection i złośliwe treści

Dokument, strona internetowa lub wiadomość może zawierać ukrytą instrukcję skierowaną do systemu AI. Narzędzie może potraktować ją jak polecenie i ujawnić dane, zmienić wynik albo podjąć niepożądane działanie. Samo napisanie w instrukcji „ignoruj złośliwe polecenia” nie daje pełnej ochrony.

7 Nadmierne uprawnienia agentów i konektorów

Asystent połączony z pocztą, dyskiem lub systemem sprzedażowym może mieć dostęp znacznie szerszy, niż potrzebuje. Błąd lub złośliwa instrukcja może wtedy wywołać skutki poza samą rozmową, na przykład wysłanie wiadomości, zmianę pliku albo pobranie danych.

8 Stronniczość i niesprawiedliwe decyzje

AI może powielać błędy obecne w danych albo stosować niejasne kryteria. Jest to szczególnie ważne w rekrutacji, ocenie pracowników, przyznawaniu benefitów, ofertowaniu i obsłudze reklamacji.

9 Deepfake, podszywanie się i oszustwa

AI ułatwia tworzenie wiarygodnych wiadomości, głosu i obrazu. Pracownik może otrzymać fałszywe polecenie przelewu, prośbę o hasło albo nagranie rzekomego przełożonego.

10 Zależność od dostawcy

Model, cena, limit, regulamin lub sposób działania usługi może się zmienić. Narzędzie może mieć awarię, wycofać funkcję albo pogorszyć jakość wyników. Krytyczny proces nie powinien zależeć od jednego produktu bez planu awaryjnego.

Bezpieczeństwo nie polega na założeniu, że system nigdy się nie pomyli. Polega na ograniczeniu skutków błędu.

Gdy AI ma więcej uprawnień



Najważniejsze wnioski

- Im więcej dostępu ma AI, tym poważniejsze mogą być skutki błędu.
- Agent wykonujący działania wymaga mocniejszych ograniczeń niż zwykły chatbot.
- Firma potrzebuje planu na zmianę ceny, modelu, funkcji lub dostawcy.

Co zrobić w swojej firmie

Dla każdego narzędzia połączanego z innymi systemami sprawdź:

- do jakich danych ma dostęp,
- jakie działania może wykonać,
- czy każda ważna operacja wymaga potwierdzenia człowieka,
- jak szybko można odebrać dostęp,
- co zrobicie podczas awarii lub zmiany warunków usługi.

ŹRÓDŁA

- OWASP, Top 10 dla aplikacji LLM, w tym prompt injection, sensitive information disclosure, excessive agency i overreliance
- NCSC, Guidelines for Secure AI System Development
- ENISA, Securing Machine Learning Algorithms
- ENISA, Cybersecurity in the Frontier AI Era, 7.07.2026

13 HUMAN IN THE LOOP: KIEDY CZŁOWIEK MUSI SPRAWDZIĆ WYNIK?

Human in the loop oznacza, że człowiek pozostaje realną częścią procesu. Nie chodzi o samo kliknięcie „zatwierdź”. Osoba nadzorująca powinna rozumieć zadanie, znać ograniczenia narzędzia, mieć czas na sprawdzenie wyniku oraz możliwość jego odrzucenia, poprawienia albo zatrzymania całego procesu.

AI Act wymaga skutecznego nadzoru człowieka w systemach wysokiego ryzyka. Osoba nadzorująca powinna potrafić monitorować działanie systemu, rozpoznać nieprawidłowości, właściwie zinterpretować wynik i nie ulegać automatycznemu zaufaniu do rekomendacji AI. Powinna też móc pominąć wynik, odwrócić decyzję albo bezpiecznie przerwać działanie systemu.

W zwykłych zastosowaniach biurowych warto użyć prostego podziału.

Poziom wpływu	Przykład	Wymagana kontrola
Niski	szkic notatki wewnętrznej, lista pomysłów	szybkie sprawdzenie sensu i zgodności z zadaniem
Średni	oferta, post, odpowiedź klientowi, analiza dokumentu	obowiązkowa akceptacja człowieka przed użyciem
Wysoki	treść prawną, finansową, zdrowotną, HR lub dotyczącą bezpieczeństwa	kontrola kompetentnej osoby oraz źródeł
Krytyczny	decyzja o zatrudnieniu, świadczeniu, zdrowiu, dostępie lub bezpieczeństwie	AI nie podejmuje samodzielnej decyzji; potrzebna jest formalna procedura i możliwość odwołania

Najważniejsza zasada brzmi: im większy wpływ wyniku na człowieka, pieniądze, prawa, zdrowie lub reputację firmy, tym silniejszy powinien być nadzór.



Najważniejsze wnioski

- Nadzór człowieka musi być realny, a nie pozorny.
- Osoba zatwierdzająca potrzebuje wiedzy, czasu i uprawnień do odmowy.
- Wysokiego ryzyka nie da się ograniczyć samym dopiskiem „wynik wygenerowała AI”.

Co zrobić w swojej firmie

Przy każdym zastosowaniu zapisz:

- Kto sprawdza wynik?
- Co dokładnie ma sprawdzić?
- Z jakich źródeł korzysta?
- Kiedy wynik trzeba odrzucić?
- Kto podejmuje ostateczną decyzję?

ŹRÓDŁA

- AI Act Service Desk, art. 14, human oversight
- AI Act Service Desk, art. 26, obowiązki podmiotów stosujących systemy wysokiego ryzyka
- NIST, AI RMF Generative AI Profile

14 MINIMALNY ZESTAW ZASAD AI W FIRMIE

Mała firma nie potrzebuje na początek wielostronicowego systemu zarządzania AI. Potrzebuje kilku zasad, które pracownicy rozumieją i rzeczywiście stosują.

1. Rejestr narzędzi i zastosowań

Zapisz nazwę narzędzia, cel użycia, użytkowników, rodzaj danych, właściciela procesu i datę ostatniego przeglądu. Rejestr ma pokazać, gdzie AI jest używana, a nie tworzyć dokumentację dla samej dokumentacji.

2. Lista narzędzi dozwolonych, warunkowych i niedozwolonych

Pracownik powinien wiedzieć, z czego może korzystać bez pytania, co wymaga zgody oraz czego nie wolno używać do pracy firmowej.

3. Zasady dotyczące danych

Określ, czego nie wolno wprowadzać, kiedy trzeba usunąć identyfikatory i kiedy potrzebna jest konsultacja z osobą odpowiedzialną za dane, bezpieczeństwo lub prawo.

4. Role i odpowiedzialność

Wyznacz właściciela każdego ważniejszego zastosowania AI. Jedna osoba powinna odpowiadać za ogólne zasady, ale odpowiedzialność za wynik pozostaje również po stronie właściciela danego procesu.

5. Zasady kontroli wyniku

Ustal, które treści wymagają zwykłego sprawdzenia, które zatwierdzenia eksperta, a które nie mogą być tworzone lub podejmowane automatycznie.

6. Kompetencje pracowników

Szkolenie powinno obejmować narzędzia używane w danej roli, ochronę danych, ograniczenia modeli, sposób sprawdzania wyników i zgłaszanie problemów. Nie musi być jednakowe dla wszystkich.

Dobra polityka jest krótka, konkretna i regularnie aktualizowana. Jeżeli pracownik nie potrafi znaleźć odpowiedzi w ciągu minuty, dokument jest prawdopodobnie zbyt skomplikowany.

Minimalny system AI w firmie

6 ELEMENTÓW, KTÓRE PORZĄDKUJĄ KORZYSTANIE Z AI



Dobra polityka jest **krótka, konkretna i łatwa do znalezienia.**

Najważniejsze wnioski

- Najpierw zasady, które da się stosować, później rozbudowana dokumentacja.
- Każde ważniejsze zastosowanie powinno mieć właściciela.
- Polityka bez szkolenia i bezpiecznego narzędzia nie rozwiązuje problemu.

Co zrobić w swojej firmie

Przygotuj jednostronicową instrukcję dla pracowników zawierającą:

- dozwolone narzędzia,
- dane zakazane,
- obowiązek sprawdzania wyników,
- obszary wymagające zgody,
- sposób zgłaszania incydentu lub wątpliwości.

ŹRÓDŁA

- AI Act Service Desk, art. 4, kompetencje w zakresie AI
- NIST, Artificial Intelligence Risk Management Framework
- Komisja Europejska, ochrona danych w fazie projektowania i domyślna ochrona danych

15 OCENA NARZĘDZIA, INCYDENT I CIĄGŁOŚĆ DZIAŁANIA

Bezpieczeństwo AI trzeba sprawdzać przed wdrożeniem, podczas używania i po wystąpieniu problemu.

Przed zatwierdzeniem narzędzia

Firma powinna sprawdzić regulamin, politykę prywatności, DPA, zasady trenowania na danych klientów, retencję, podwykonawców, lokalizację przetwarzania, możliwość usunięcia danych, funkcje administracyjne i sposób odebrania dostępu. Przy integracji z pocztą, dyskiem lub CRM trzeba dodatkowo ograniczyć uprawnienia do niezbędnego minimum.

Gdy do AI trafiły niewłaściwe dane

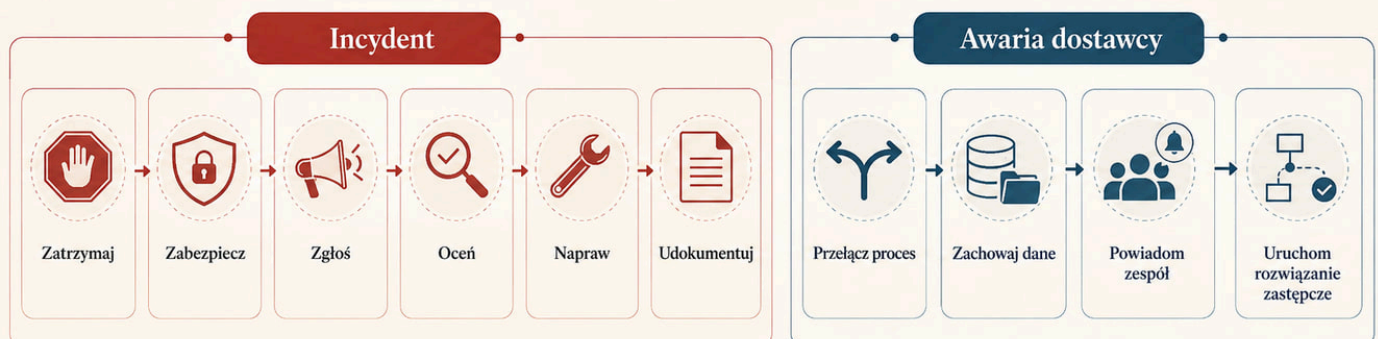
Pracownik powinien wiedzieć, że problem należy zgłosić od razu, bez ukrywania błędu. Pierwsze działania to zatrzymanie dalszego użycia, ustalenie zakresu ujawnionych informacji, usunięcie rozmowy lub pliku, jeżeli jest to możliwe, odebranie dostępów oraz zmiana ujawnionych haseł lub kluczy.

Następnie firma ocenia, czy doszło do naruszenia ochrony danych osobowych, tajemnicy przedsiębiorstwa albo innych obowiązków. W przypadku naruszenia danych osobowych, które może powodować ryzyko dla praw lub wolności osób, administrator zgłasza je organowi nadzorcemu bez zbędnej zwłoki, nie później niż w ciągu 72 godzin od stwierdzenia. Nie oznacza to, że każdy błąd automatycznie wymaga zgłoszenia. Każdy powinien jednak zostać oceniony i udokumentowany.

Plan awaryjny

Firma powinna założyć, że usługa może przestać działać, zmienić model, cenę, limit lub warunki. Krytyczne procesy potrzebują ręcznej ścieżki zastępczej, możliwości eksportu danych i jasno określonego sposobu zmiany dostawcy.

Dwa procesy postępowania



Nie każdy błąd wymaga zgłoszenia, ale każdy powinien zostać **oceniony i udokumentowany**.



Proces krytyczny powinien mieć ścieżkę zastępczą.

Najważniejsze wnioski

- Incydent należy zgłosić szybko, a nie dopiero wtedy, gdy pojawią się skutki.
- Dostęp agenta do poczty lub dysku powinien być możliwie wąski.
- Firma potrzebuje sposobu pracy bez danego narzędzia.

Co zrobić w swojej firmie

Przygotuj dwie krótkie karty:

Karta incydentu: kogo powiadomić, co zapisać, jak ograniczyć skutki.

Karta dostawcy: właściciel, DPA, retencja, podwykonawcy, uprawnienia, koszt, plan wyjścia.

ŹRÓDŁA

- Komisja Europejska, postępowanie w przypadku naruszenia ochrony danych
- NCSC, Guidelines for Secure AI System Development
- OWASP, GenAI Security Project
- NIST, AI RMF Generative AI Profile

16

PLAN MINIMUM NA PIERWSZE 30 DNI

Nie trzeba wdrażać wszystkiego jednocześnie. Najlepszym początkiem jest krótki plan, który daje firmie widoczność, ogranicza największe ryzyka i nie zatrzymuje użytecznych zastosowań AI.

Tydzień 1: sprawdź stan obecny

- Zapytaj pracowników, z jakich narzędzi korzystają i do jakich zadań.
- Zidentyfikuj użycie prywatnych kont oraz integracje z pocztą, dyskiem i innymi systemami.
- Oznacz zastosowania dotyczące HR, klientów, prawa, zdrowia, finansów i bezpieczeństwa.
- Zapisz narzędzia i zastosowania w prostym rejestrze.

Tydzień 2: ustal podstawowe zasady

- Wybierz zatwierdzone narzędzia i konta firmowe.
- Określ dane zielone, żółte i czerwone.
- Wprowadź obowiązek kontroli treści wysyłanych na zewnątrz.
- Ustal, które zastosowania wymagają wcześniejszej zgody.

Tydzień 3: przygotuj ludzi i dostęp

- Przeprowadź krótkie szkolenie dopasowane do ról.
- Włącz MFA i usuń niepotrzebne konta.
- Ogranicz dostęp konektorów oraz agentów do niezbędnego minimum.
- Wyznacz osoby odpowiedzialne za zasady, dane i konkretne procesy.

Tydzień 4: sprawdź dokumenty i gotowość na problem

- Zweryfikuj regulaminy, DPA, retencję i podwykonawców najważniejszych dostawców.
- Przygotuj kartę zgłoszenia incydentu.
- Ustal ręczną ścieżkę awaryjną dla ważnych procesów.
- Zaplanuj przegląd zasad za trzy miesiące.

Po 30 dniach firma nie będzie miała zamkniętego całego tematu. Będzie jednak wiedziała, gdzie używa AI, które ryzyka są najważniejsze i kto odpowiada za dalsze działania.

Najważniejsze wnioski

- Najpierw widoczność, później doskonalenie.
- Jedna realna zasada jest lepsza niż polityka, której nikt nie czyta.
- Przegląd trzeba powtarzać, bo narzędzia i warunki szybko się zmieniają.

Co zrobić w swojej firmie

Wyznacz jedną osobę odpowiedzialną za koordynację planu oraz właściciela każdego zadania. Do każdego działania przypisz termin i dowód wykonania, na przykład listę narzędzi, instrukcję, protokół szkolenia lub kartę dostawcy.

Plan 4 tygodni

KRÓTKI PLAN WDROŻENIA BEZPIECZNEGO KORZYSTANIA Z AI

TYDZIEŃ 1



Inwentaryzacja

- ✓ narzędzia
- ✓ zadania
- ✓ prywatne konta
- ✓ rejestr

TYDZIEŃ 2



Zasady

- ✓ zatwierdzone narzędzia
- ✓ dane zielone / żółte / czerwone
- ✓ kontrola treści
- ✓ zgoda

TYDZIEŃ 3



Ludzie i dostęp

- ✓ szkolenie
- ✓ MFA
- ✓ ograniczenie dostępu
- ✓ właściciele procesów

TYDZIEŃ 4



Umowy i gotowość

- ✓ DPA i retencja
- ✓ karta incydentu
- ✓ ścieżka awaryjna
- ✓ przegląd za 3 miesiące



Po 30 dniach firma ma **widoczność, priorytety i odpowiedzialność**.

ŹRÓDŁA

- NIST, AI Risk Management Framework
- AI Act Service Desk, art. 4
- NCSC, Guidelines for Secure AI System Development

17 CHECKLISTA: BEZPIECZNE KORZYSTANIE Z AI W FIRMIE

Ta checklista nie jest audytem zgodności. Służy do szybkiego sprawdzenia, czy firma wdrożyła podstawowe zabezpieczenia. Odpowiedź „nie” nie oznacza automatycznie naruszenia prawa. Pokazuje obszar, który warto uporządkować.

Narzędzia i widoczność

- ☐ Wiemy, z jakich narzędzi AI korzystają pracownicy.
- ☐ Prowadzimy prosty rejestr zastosowań AI.
- ☐ Mamy listę narzędzi dozwolonych, warunkowych i niedozwolonych.
- ☐ Do pracy firmowej używamy kont zarządzanych przez organizację.

Dane i dostęp

- ☐ Pracownicy wiedzą, jakich danych nie wolno wprowadzać do niezatwierdzonych narzędzi.
- ☐ Stosujemy zasadę minimum danych i usuwamy zbędne identyfikatory.
- ☐ Sprawdziliśmy DPA, retencję, podwykonawców i zasady używania danych.
- ☐ Mamy MFA i regularnie odbieramy niepotrzebne dostępy.
- ☐ Konektory i agenci mają tylko niezbędne uprawnienia.

Ludzie i odpowiedzialność

- ☐ Osoby korzystające z AI zostały przeszkolone odpowiednio do swojej roli.
- ☐ Każde ważniejsze zastosowanie ma właściciela.
- ☐ Treści zewnętrzne są sprawdzane przez człowieka.
- ☐ Zastosowania HR, prawne, zdrowotne i finansowe mają dodatkową kontrolę.
- ☐ Pracownik może zgłosić wątpliwość lub błąd bez obawy przed ukrywaniem problemu.

Incydenty i ciągłość

- ☐ Mamy prostą procedurę zgłaszania incydentu AI.
- ☐ Wiemy, co zrobić po ujawnieniu danych, hasła lub klucza API.
- ☐ Krytyczne procesy mają ręczną ścieżkę zastępczą.
- ☐ Ustalamy budżety, limity i osoby zatwierdzające koszty.
- ☐ Przeglądamy narzędzia i zasady co najmniej raz na kwartał.

WŁAŚCICIEL PRZEGLĄDU

DATA

NASTĘPNY PRZEGLĄD

Jak czytać wynik

0–6 pól: zacznij od narzędzi, danych i szkolenia.

7–13 pól: masz podstawy, ale potrzebujesz spójnych zasad i odpowiedzialności.

14–19 pól: dobra baza, skup się na testowaniu, aktualizacji i obszarach wysokiego wpływu.

Wynik jest orientacyjny i nie stanowi oceny prawnej.

Co zrobić w swojej firmie

Wybierz trzy niezaznaczone pola, które ograniczą największe ryzyko. Przypisz właściciela i termin wykonania w ciągu najbliższych 30 dni.

ŹRÓDŁA

Checklista podsumowuje zalecenia i źródła przedstawione na stronach 2–16. Nie jest samodzielnym standardem zgodności.

18

GDZIE SZUKAĆ AKTUALNYCH INFORMACJI?

Prawo, funkcje narzędzi i warunki dostawców zmieniają się szybciej niż większość firm aktualizuje dokumenty. Dlatego przed podjęciem decyzji należy sprawdzać informacje bezpośrednio w aktualnych źródłach właściwej instytucji lub dostawcy.

Prawo i wytyczne Unii Europejskiej

- EUR-Lex: AI Act i RODO
- AI Act Service Desk
- Komisja Europejska: wdrażanie AI Act
- EDPB: ochrona danych i AI

Polska

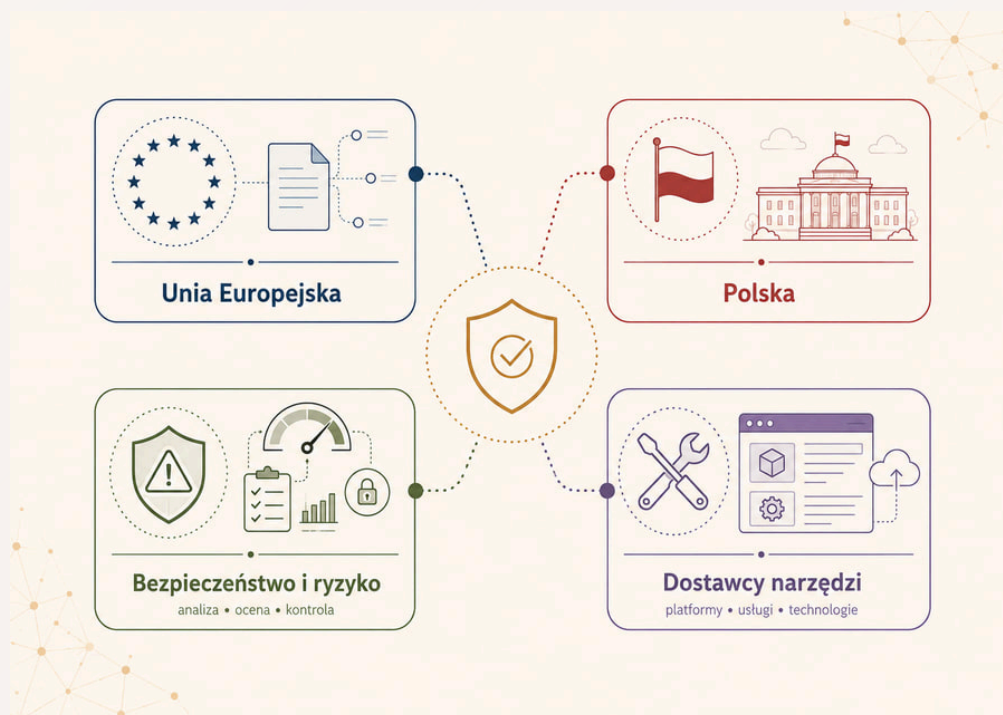
- Sejm RP: proces legislacyjny
- Dziennik Ustaw
- UODO
- Ministerstwo Cyfryzacji

Bezpieczeństwo i zarządzanie ryzykiem

- NIST AI Risk Management Framework
- ENISA: bezpieczeństwo AI
- OWASP GenAI Security Project
- NCSC: Secure AI System Development

Dostawcy narzędzi

Sprawdź wyłącznie aktualne dokumenty konkretnego produktu: regulamin, politykę prywatności, DPA, trust center, listę podwykonawców, retencję, lokalizację danych i informacje o używaniu treści do trenowania modeli.



Najważniejsze wnioski

- Nie opieraj decyzji na starym artykule lub zrzucie ekranu.
- Zawsze sprawdzaj konkretny produkt i rodzaj konta.