

BEZPIECZEŃSTWO DANYCH

Najważniejsze wytyczne dotyczące narzędzi i rozwiązań AI, obejmują zapewnienie ich zgodności z poniższymi normami:

1. National Institute of Standards and Technology's ("NIST") AI Risk Management Framework, ([AI Risk Management Framework | NIST](#))
2. the EU AI Act ("AIA"), ([Rozporządzenie - UE - 2024/1689 - EN - EUR-Lex \(europa.eu\)](#))
3. the International Organization for Standardization's guidance on AI risk management ("ISO/IEC 23894") (<https://www.iso.org/obp/ui/#iso:std:iso-iec:23894:ed-1:v1:en>)

W McDonald's dozwolone jest użycie narzędzi zatwierdzonych przez Generative AI Center of Excellence (Gen AI COE).

Przed wdrożeniem systemu AI McDonald's przeprowadzi ocenę systemu („Ocena AI”), która może obejmować ocenę wpływu AI („AIIA”) w zależności od klasyfikacji ryzyka systemu AI.

Celem Oceny AI jest pomoc w identyfikacji i rozwiązywaniu ryzyk związanych z takim systemem AI oraz zapewnienie zgodności z Zasadami AI.

Ocena AI zawiera:

1. Opisanie zakresu zamierzonych zastosowań i funkcji systemu AI;
2. Opisanie planowanego zakresu geograficznego i czasowego systemu AI;
3. Wyjaśnienie źródeł i charakteru danych treningowych dla systemu AI, w tym wszelkich danych objętych prawami autorskimi użytych do trenowania systemu;
4. Identyfikację i wyjaśnienie komponentu nadzoru ludzkiego nad systemem AI oraz wszelkich procedur obsługi skarg i zadośćuczynienia;
5. Identyfikację kategorii osób lub grup, które mogą być dotknięte przez system AI, w tym czy system AI będzie używany do podejmowania decyzji przez lub o ludziach, a jeśli tak, to kto będzie używał wyników generowanych przez AI do podejmowania takich decyzji;
6. Opisanie i ocenę potencjalnych korzyści z używania systemu AI;
7. Wskazanie odpowiedniego etapu cyklu życia systemu AI, który obejmowałby: (i) planowanie; (ii) projektowanie; (iii) rozwój; (iv) testowanie; (v) wdrożenie i uruchomienie; lub (vi) plug and play;
8. Wyjaśnienie, jak system AI został przeszkolony, w tym logiki i metod statystycznych użytych do stworzenia systemu AI;
9. Oceny dokładności i niezawodności systemu AI;
10. Opisanie i oceny wszelkich potencjalnie wysokich ryzyk dla prywatności ludzi i podstawowych praw, z uwzględnieniem uprzedzeń i dyskryminacyjnych skutków;
11. Opisanie i oceny przewidywalnego wpływu systemu AI na środowisko (jeśli dotyczy);
12. Wskazania, jak potencjalnie wysokie ryzyka (w tym ryzyko szkód i negatywnego wpływu na podstawowe prawa) będą łagodzone; oraz
13. Oceny, czy jakiegokolwiek wysokie ryzyka pozostają nawet po wdrożeniu środków łagodzących.

Zasady McDonald's dotyczące AI w Praktyce (wyciąg):

Wykorzystanie systemów AI w McDonald's musi uwzględniać następujące zasady AI McDonald's:

- **Ważność i niezawodność.** Wszelkie wyniki generowane przy pomocy AI powinny być weryfikowane, aby upewnić się, że zamierzony wynik został osiągnięty. Wyniki powinny być również testowane pod kątem niezawodności na podstawie spójności połączonych rezultatów.
- **Bezpieczne.** Bezpieczne korzystanie z AI zależy od odpowiedzialnego projektowania i wdrażania w sposób minimalizujący niezamierzone i nieoczekiwane szkody, a także od dokumentacji i komunikacji na temat potencjalnych ryzyk, zwłaszcza jeśli dotyczą one potencjalnego ryzyka dla bezpieczeństwa ludzi.
- **Zabezpieczone i odporne.** Systemy AI zaprojektowane w celu zapobiegania nieautoryzowanemu użyciu i dostępowi są bardziej zabezpieczone i odporne. Systemy AI powinny być opracowywane tak, aby wytrzymywały nieoczekiwane, wrogie zdarzenia, minimalizując zakłócenia w działalności McDonald's i chroniąc nasze dane.
- **Agencja ludzka i nadzór.** Systemy AI powinny być opracowywane i używane jako narzędzie służące ludziom, szanujące autonomię osobistą i mogące być odpowiednio nadzorowane i kontrolowane przez ludzi.
- **Odpowiedzialne i przejrzyste.** Systemy AI powinny być budowane z uwzględnieniem przejrzystości i dokładności już na etapie projektowania. Zespoły odpowiedzialne za dowód koncepcji i rozwój oraz właściciele powinni być przygotowani do wyjaśnienia, zarówno wewnątrz, jak i zewnątrz, jakie dane wejściowe są używane w ich systemach AI oraz kto jest odpowiedzialny, gdy wyniki są nieprawidłowe, szkodliwe lub niespójne z wartościami McDonald's, zasadami AI i politykami. Ludzie powinni być świadomi, że mają do czynienia z systemem AI oraz znać możliwości i ograniczenia systemu AI.
- **Wyjaśnialne i interpretowalne.** Systemy AI powinny zawierać wystarczającą ilość informacji, aby użytkownicy końcowi i operatorzy mogli zrozumieć projekt, funkcję, cele i potencjalne skutki dla praw człowieka i bezpieczeństwa. Innymi słowy, należy podkreślić, co, jak i dlaczego system AI działa.
- **Zrównoważone.** Systemy AI powinny być opracowywane i używane w sposób zrównoważony i przyjazny dla środowiska, a także w sposób korzystny dla ludzi, szanujący prawa człowieka i uwzględniający bezpieczeństwo dzieci i młodzieży. Długoterminowy wpływ systemów AI na społeczeństwo i jednostki powinien być regularnie monitorowany i oceniany.
- **Wzmocniona prywatność i zarządzanie danymi.** Systemy AI powinny być projektowane od samego początku z uwzględnieniem funkcji wzmacniających prywatność, które zmniejszają uprzedzenia i poprawiają bezpieczeństwo oraz przejrzystość, przetwarzając dane spełniające wysokie standardy jakości i integralności. W miarę możliwości i odpowiedniości należy stosować metody ochrony prywatności danych, takie jak deidentyfikacja, agregacja i szyfrowanie.
- **Sprawiedliwe – z zarządzaniem szkodliwymi uprzedzeniami.** Systemy AI powinny być opracowywane i wdrażane w celu minimalizowania problemów związanych ze sprawiedliwością oraz szkodliwymi uprzedzeniami lub dyskryminacją (np. oferowanie

określonych usług lub produktów tylko niektórym osobom lub wyróżnianie niektórych osób do oceny wyników).

- **Inkluzywne.** Wykorzystanie i rozwój systemów AI powinny obejmować opinie i wkład wewnętrznych i zewnętrznych aktorów o różnorodnym pochodzeniu i dyscyplinach oraz promować równy dostęp, równość płci i różnorodność kulturową.
- **Dynamiczne.** Wykorzystanie systemów AI powinno przewidywać dynamiczny i zmieniający się charakter ryzyk, wymagań regulacyjnych i prawnych oraz oczekiwań klientów dotyczących AI. Dlatego dynamiczne zarządzanie ryzykiem powinno być integralną częścią korzystania z AI.
- **Ciągłe doskonalenie.** Systemy AI powinny być ciągle doskonalone w miarę pojawiania się nowych ryzyk.

Dodatkowo rozwiązanie powinno spełnić ogólne kryteria bezpieczeństwa tj.:

1. **Zgodność z lokalnymi regulacjami:** narzędzie spełnia wymagania prawne i regulacyjne, takie jak RODO w Europie.
2. **Szyfrowanie danych:** narzędzie powinno oferować silne mechanizmy szyfrowania danych zarówno w trakcie przesyłania, jak i przechowywania.
3. **Kontrola dostępu:** uwierzytelnienie dwuskładnikowa 2FA/MFA.
4. **Audyt, monitorowanie i logowanie działań:** narzędzie powinno umożliwiać audyty i monitorowanie aktywności, aby wykrywać i reagować na potencjalne zagrożenia bezpieczeństwa.
5. **Transparentność:** dostawca narzędzia powinien być transparentny w kwestii swoich praktyk dotyczących bezpieczeństwa i prywatności.
6. **Wsparcie techniczne:** ważne jest, aby dostawca oferował wsparcie techniczne w przypadku problemów związanych z bezpieczeństwem.