KOMISJA
EUROPEJSKA

Bruksela, dnia 29.7.2025 r.
C(2025) 5053 final

# KOMUNIKAT KOMISJI

## Wytyczne Komisji w sprawie definicji systemu sztucznej inteligencji ustanowionej rozporządzeniem (UE) 2024/1689 (akt w sprawie sztucznej inteligencji)

## I. Cel wytycznych

1) Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 („akt w sprawie sztucznej inteligencji")[1] weszło w życie w dniu 1 sierpnia 2024 r. W akcie w sprawie sztucznej inteligencji ustanowiono zharmonizowane przepisy dotyczące rozwoju, wprowadzania do obrotu, oddawania do użytku i wykorzystywania sztucznej inteligencji („AI") w Unii[2]. Jego celem jest promowanie innowacji w dziedzinie AI i upowszechniania tej technologii, przy jednoczesnym zapewnieniu wysokiego poziomu ochrony zdrowia, bezpieczeństwa i praw podstawowych w Unii, w tym demokracji i praworządności.

2) Akt w sprawie sztucznej inteligencji nie ma zastosowania do wszystkich systemów, a jedynie do tych, które są zgodne z definicją „systemu AI" w rozumieniu art. 3 pkt 1 aktu. Definicja systemu AI ma zatem zasadnicze znaczenie dla zrozumienia zakresu stosowania aktu w sprawie sztucznej inteligencji.

3) W art. 96 ust. 1 lit. f) aktu w sprawie sztucznej inteligencji zobowiązano Komisję do opracowania wytycznych dotyczących stosowania definicji systemu AI określonej w art. 3 pkt 1 aktu. Wydając niniejsze wytyczne, Komisja pragnie pomóc dostawcom i innym odpowiednim osobom, w tym zainteresowanym uczestnikom rynku i podmiotom instytucjonalnym, w ustaleniu, czy dany system stanowi system AI w rozumieniu aktu w sprawie sztucznej inteligencji, tym samym ułatwiając skuteczne stosowanie i egzekwowanie aktu.

4) Definicję systemu AI stosuje się od dnia 2 lutego 2025 r.[3] wraz z innymi przepisami zawartymi w rozdziałach I i II aktu w sprawie sztucznej inteligencji, w szczególności art. 5 aktu dotyczącym zakazanych praktyk w zakresie AI. Ponieważ definicja systemu AI ma decydujące znaczenie dla zrozumienia zakresu aktu w sprawie sztucznej inteligencji, w tym zakazanych praktyk, niniejsze wytyczne przyjmuje się równolegle z wytycznymi Komisji w sprawie zakazanych praktyk w zakresie AI.

5) W niniejszych wytycznych uwzględniono wyniki konsultacji z zainteresowanymi stronami i konsultacji z Europejską Radą ds. Sztucznej Inteligencji.

6) Ze względu na dużą różnorodność systemów AI nie jest możliwe przedstawienie w niniejszych wytycznych wyczerpującego wykazu wszystkich potencjalnych systemów AI. Jest to zgodne z motywem 12 aktu w sprawie sztucznej inteligencji, w którym wyjaśniono, że pojęcie „systemu AI" powinno być jasno zdefiniowane przy jednoczesnym „zapewnieniu swobody umożliwiającej dostosowanie się do szybkiego rozwoju technologicznego w tej dziedzinie". Definicja systemu AI nie powinna być stosowana w sposób mechaniczny; każdy system należy oceniać na podstawie jego szczególnych cech.

[1] Rozporządzenie (UE) 2024/1689.
[2] Art. 1 aktu w sprawie sztucznej inteligencji.
[3] Art. 113 akapit trzeci lit. a).

7) Niniejsze wytyczne nie mają charakteru wiążącego. Wszelkiej autorytatywnej wykładni aktu w sprawie sztucznej inteligencji może ostatecznie dokonywać wyłącznie Trybunał Sprawiedliwości Unii Europejskiej (TSUE).

## II. Cel i główne elementy definicji systemu AI

8) W art. 3 pkt 1 aktu w sprawie sztucznej inteligencji system AI zdefiniowano w następujący sposób:
*„system AI" oznacza system maszynowy, który został zaprojektowany do działania z różnym poziomem autonomii po jego wdrożeniu oraz który może wykazywać zdolność adaptacji po jego wdrożeniu, a także który – na potrzeby wyraźnych lub dorozumianych celów – wnioskuje, jak generować na podstawie otrzymanych danych wejściowych wyniki, takie jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne lub wirtualne;*

9) Definicja ta obejmuje siedem głównych elementów: 1) system maszynowy; 2) który został zaprojektowany do działania z różnym poziomem autonomii; 3) który może wykazywać zdolność adaptacji po jego wdrożeniu; 4) a także który – na potrzeby wyraźnych lub dorozumianych celów; 5) wnioskuje, jak generować na podstawie otrzymanych danych wejściowych wyniki, 6) takie jak predykcje, treści, zalecenia lub decyzje, 7) które mogą wpływać na środowisko fizyczne lub wirtualne.

10) W definicji systemu AI przyjęto perspektywę opartą na cyklu życia obejmującą dwa główne etapy: etap przed wdrożeniem, czyli etap „tworzenia" systemu, oraz etap po wdrożeniu, czyli etap „wykorzystywania" systemu[4]. Siedem elementów wymienionych w tej definicji nie musi być stale obecnych na obu etapach tego cyklu życia. Zamiast tego w definicji uznano, że poszczególne elementy mogą pojawiać na jednym etapie, ale mogą nie utrzymywać się na obu. To podejście do zdefiniowania systemu AI jest odzwierciedleniem złożoności i różnorodności systemów AI, co gwarantuje, że definicja odpowiada celom aktu w sprawie sztucznej inteligencji poprzez uwzględnienie szerokiego zakresu systemów AI.

### *1. System maszynowy*

11) Termin „maszynowy" odnosi się do faktu, że systemy AI są rozwijane i działają z wykorzystaniem maszyn. Pod pojęciem „maszyny" można rozumieć zarówno komponenty sprzętowe, jak i składniki oprogramowania, które umożliwiają funkcjonowanie systemu AI. Komponenty sprzętowe odnoszą się do fizycznych elementów maszyny, takich jak procesory, pamięć, urządzenia pamięciowe, urządzenia sieciowe oraz interfejsy wejścia/wyjścia, które zapewniają infrastrukturę do obliczeń. Składniki oprogramowania obejmują kod komputerowy, instrukcje, programy, systemy operacyjne i aplikacje, które przetwarzają dane i wykonują zadania.

[4] Aby zapoznać się z etapami cyklu życia systemów AI, zob. OECD (2024), „Explanatory memorandum on the updated OECD definition of an AI system" [Nota objaśniająca dotycząca zaktualizowanej definicji systemu AI opracowanej przez OECD], OECD Artificial Intelligence Papers, nr 8, OECD Publishing, Paryż, https://doi.org/10.1787/623da898-en, s. 7.

12) Wszystkie systemy AI są maszynowe, ponieważ potrzebują maszyn, które umożliwiają im funkcjonowanie, np. trenowanie modeli, przetwarzanie danych, modelowanie predykcyjne i zautomatyzowane podejmowanie decyzji na dużą skalę. Cały cykl życia zaawansowanych systemów AI opiera się na maszynach, na które może składać się wiele komponentów sprzętowych lub elementów oprogramowania. Określenie „maszynowy" w definicji systemu AI zwraca uwagę na fakt, że systemy AI muszą opierać się na obliczeniach i operacjach maszynowych.

13) Termin „maszynowy" obejmuje wiele różnych systemów obliczeniowych. Na przykład obecnie najbardziej zaawansowane powstające kwantowe systemy obliczeniowe, istotnie różniące się od tradycyjnych systemów obliczeniowych, są systemami maszynowymi pomimo unikatowych zasad ich działania i wykorzystania w nich zjawisk mechaniki kwantowej, podobnie jak systemy biologiczne lub organiczne – o ile zapewniają zdolność obliczeniową.

### *2. Autonomia*

14) Drugi element definicji odnosi się do tego, że system został „zaprojektowany do działania z różnym poziomem autonomii". Jak wyjaśniono w motywie 12 aktu w sprawie sztucznej inteligencji, określenie „różny poziom autonomii" oznacza, że systemy AI są zaprojektowane tak, aby były „w pewnym stopniu niezależne od zaangażowania ze strony człowieka i zdolne do działania bez interwencji człowieka".

15) Pojęcia autonomii i wnioskowania idą ze sobą w parze: zdolność systemu AI do wnioskowania (tj. jego zdolność do generowania wyników, takich jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne lub wirtualne) ma kluczowe znaczenie dla zapewnienia jego autonomii.

16) Zasadnicze elementy koncepcji autonomii to „zaangażowanie ze strony człowieka" i „interwencja człowieka", a tym samym interakcja między człowiekiem a maszyną. Na jednym krańcu możliwych interakcji między człowiekiem a maszyną znajdują się systemy zaprojektowane do wykonywania wszystkich zadań za pomocą funkcji obsługiwanych ręcznie. Na drugim krańcu znajdują się systemy zdolne do działania bez zaangażowania lub interwencji człowieka, tj. systemy w pełni autonomiczne.

17) Odniesienie do „pewnego stopnia niezależności" w motywie 12 aktu w sprawie sztucznej inteligencji wyklucza systemy, które są zaprojektowane do działania wyłącznie przy pełnym zaangażowaniu człowieka i jego interwencji. Zaangażowanie i interwencja człowieka mogą mieć charakter bezpośredni, np. odbywać się w drodze sterowania ręcznego, albo pośredni, np. odbywać się za pomocą zautomatyzowanego sterowania systemowego, które pozwala ludziom na delegowanie lub nadzorowanie operacji systemowych.

18) Na przykład system, który do samodzielnego wygenerowania wyników wymaga ręcznego wprowadzenia danych wejściowych, jest systemem o „pewnym stopniu niezależności", ponieważ zaprojektowano go w sposób umożliwiający wygenerowanie

wyniku bez ręcznego sterowania tym procesem lub jego wyraźnego i dokładnego określenia przez człowieka. Podobnie system ekspercki, w przypadku którego człowiek deleguje automatyzację procesów i który jest w stanie, na podstawie danych dostarczonych przez człowieka, samodzielnie wygenerować wynik taki jak zalecenie, jest systemem charakteryzującym się „pewnym stopniem niezależności".

19) Zawarte w definicji systemu AI, w art. 3 pkt 1 aktu w sprawie sztucznej inteligencji, odniesienie do „systemu maszynowego, który został zaprojektowany do działania z różnym poziomem autonomii", podkreśla zdolność systemu do interakcji ze środowiskiem zewnętrznym, nie zaś wybór określonej techniki, takiej jak uczenie maszynowe, czy architektury modelu na potrzeby rozwoju systemu.

20) W związku z tym poziom autonomii jest warunkiem koniecznym do ustalenia, czy dany system kwalifikuje się jako system AI. Wszystkie systemy zaprojektowane tak, aby działały z pewnym racjonalnym stopniem niezależności, spełniają warunek autonomii wynikający z definicji systemu AI.

21) Systemy zdolne do działania przy ograniczonej interwencji człowieka lub bez interwencji człowieka w określonych kontekstach użycia, takich jak obszary wysokiego ryzyka wskazane w załącznikach I i III do aktu w sprawie sztucznej inteligencji, mogą, w pewnych okolicznościach, być obciążone dodatkowym potencjalnym ryzykiem i wymagać nadzoru ze strony człowieka. Poziom autonomii jest dla dostawcy istotnym czynnikiem przy projektowaniu np. nadzoru ze strony człowieka dla danego systemu lub środków ograniczania ryzyka w kontekście przeznaczenia tego systemu.

### *3. Zdolność adaptacji*

22) Trzeci element definicji zawartej w art. 3 ust. 1 aktu w sprawie sztucznej inteligencji przewiduje, że system „może wykazywać zdolność adaptacji po jego wdrożeniu". Autonomia i zdolność adaptacji to dwa odrębne, ale ściśle ze sobą powiązane pojęcia. Są one często omawiane razem, ale odnoszą się do różnych wymiarów funkcjonowania systemu AI. W motywie 12 aktu w sprawie sztucznej inteligencji wyjaśniono, że „zdolność adaptacji" odnosi się do zdolności do samouczenia się, która umożliwia zmianę zachowania systemu w czasie jego wykorzystywania. Nowe zachowanie systemu po adaptacji może przynieść inne wyniki niż poprzednie zachowanie systemu na bazie tych samych danych wejściowych.

23) Użycie czasownika „może" w odniesieniu do tego elementu definicji wskazuje, że aby stanowić system AI, dany system może, ale niekoniecznie musi, posiadać zdolności adaptacji lub samouczenia się po wdrożeniu. W związku z tym zdolność systemu do automatycznego uczenia się, odkrywania nowych wzorców lub identyfikowania powiązań w zakresie wykraczającym poza pierwotny trening jest fakultatywna, a zatem nie jest decydującym warunkiem dla ustalenia, czy dany system kwalifikuje się jako system AI.

### *4. Cele systemu AI*

24) Czwarty element definicji odnosi się do celów systemu AI. Systemy AI projektuje się tak, aby działały zgodnie z co najmniej jednym celem. Cele systemu mogą zostać określone w sposób wyraźny lub dorozumiany. **Cele określone wyraźnie** to jasno sformułowane cele, bezpośrednio zakodowane w systemie przez jego twórcę. Na przykład mogą one zostać określone jako optymalizacja określonej funkcji kosztów, prawdopodobieństwo lub skumulowana korzyść. **Cele dorozumiane** to cele, których nie sformułowano w sposób wyraźny, ale o których można wnioskować na podstawie zachowania systemu lub założeń leżących u podstaw systemu. Cele te mogą wynikać z danych treningowych lub z interakcji systemu AI ze środowiskiem.

25) W motywie 12 aktu o sztucznej inteligencji wyjaśniono, że „cele systemu AI mogą różnić się od przeznaczenia systemu AI w określonym kontekście". Cele systemu AI mają charakter wewnątrzsystemowy i odnoszą się do celów zadań, które mają zostać wykonane, oraz ich wyników. Na przykład celem firmowego wirtualnego asystenta AI może być udzielanie z dużą dokładnością i niskim współczynnikiem niepowodzenia odpowiedzi na pytania użytkowników dotyczące zbioru dokumentów. Natomiast **przeznaczenie** ma charakter zewnętrzny i obejmuje kontekst, do wdrożenia w którym zaprojektowano system, oraz sposób, w jaki system ten musi być użytkowany. Zgodnie z art. 3 pkt 12 aktu w sprawie sztucznej inteligencji przeznaczenie systemu AI odnosi się bowiem do „zastosowania, do jakiego system AI został przeznaczony przez jego dostawcę". Na przykład przeznaczeniem firmowego wirtualnego asystenta AI może być zapewnienie danemu działowi przedsiębiorstwa pomocy w wykonywaniu określonych zadań. Z tego względu może być niezbędne, aby dokumenty używane przez wirtualnego asystenta spełniały określone wymogi (np. długość, formatowanie) oraz aby pytania użytkowników ograniczały się do dziedziny, w której system ma funkcjonować. Wypełnianiu tego przeznaczenia służy nie tylko wewnętrzne działanie systemu zmierzające do osiągnięciu jego celów, ale także inne czynniki, takie jak włączenie systemu do szerszego obiegu zadań obsługi klienta, dane wykorzystywane przez system lub instrukcja obsługi.

### *5. Wnioskowanie o sposobie generowania wyników przy użyciu technik sztucznej inteligencji*

26) Piąty element charakteryzujący system AI to jego zdolność do wnioskowania, jak generować wyniki na podstawie otrzymanych danych wejściowych. W motywie 12 aktu w sprawie sztucznej inteligencji wyjaśniono, że „[j]edną z kluczowych cech systemów AI jest ich zdolność do wnioskowania". Jak ponadto wyjaśniono we wspomnianym motywie, pojęcie systemów AI należy odróżniać od „prostszych tradycyjnych systemów oprogramowania lub założeń programistycznych [...] i nie powinno [ono] obejmować systemów opartych na zasadach określonych wyłącznie przez osoby fizyczne w celu automatycznego wykonywania operacji". Ta zdolność do wnioskowania jest zatem kluczowym i niezbędnym warunkiem, który odróżnia systemy AI od innych rodzajów systemów.

27) W motywie 12 wyjaśniono również, że „[t]a zdolność do wnioskowania odnosi się do procesu uzyskiwania wyników, takich jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne i wirtualne, oraz do zdolności systemów AI do tworzenia modeli lub algorytmów na podstawie informacji wejściowych lub danych". Takie rozumienie pojęcia „wnioskowania" nie jest sprzeczne z normą ISO/IEC 22989, w której zdefiniowano wnioskowanie jako „rozumowanie, w którym wnioski wyprowadza się ze znanych przesłanek", przy czym norma ta zawiera uwagę dotyczącą sztucznej inteligencji, w której stwierdza się: „[w] przypadku sztucznej inteligencji przesłanką jest fakt, reguła, model, cecha lub surowe dane" [5].

28) „Proces uzyskiwania wyników, takich jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne i wirtualne", odnosi się do zdolności systemu AI, głównie na „etapie użytkowania", do generowania wyników na podstawie danych wejściowych. „Zdolność systemów AI do tworzenia modeli lub algorytmów na podstawie informacji wejściowych lub danych" odnosi się przede wszystkim, ale nie tylko, do „etapu tworzenia" systemu, i podkreśla znaczenie technik stosowanych przy tworzeniu systemu.

29) Wyrażenie „wnioskuje, jak", użyte w art. 3 pkt 1 i doprecyzowane w motywie 12 aktu w sprawie sztucznej inteligencji, ma szerszy zakres i nie ogranicza się tylko do wąskiego rozumienia pojęcia wnioskowania jako zdolności systemu do uzyskiwania wyników na podstawie określonych danych wejściowych, a tym samym do otrzymywania rezultatu w drodze wnioskowania. W związku z tym sformułowanie użyte w art. 3 pkt 1 aktu w sprawie sztucznej inteligencji, tj. „wnioskuje, jak generować wyniki", należy rozumieć jako odnoszące się do etapu tworzenia, na którym system uzyskuje wyniki za pomocą technik sztucznej inteligencji umożliwiających wnioskowanie.

### *5.1. Techniki sztucznej inteligencji umożliwiające wnioskowanie*

30) Skupiając się konkretnie na etapie tworzenia systemu AI, w motywie 12 aktu w sprawie sztucznej inteligencji stwierdza się ponadto, że „[t]echniki, które umożliwiają wnioskowanie podczas tworzenia systemu AI, obejmują mechanizmy uczenia maszynowego, które na podstawie danych uczą się, jak osiągnąć określone cele, oraz podejścia oparte na logice i wiedzy, które polegają na wnioskowaniu na podstawie zakodowanej wiedzy lub symbolicznego przedstawienia zadania, które należy rozwiązać". Techniki te należy rozumieć jako „techniki sztucznej inteligencji".

31) W wyjaśnieniu tym podkreślono wyraźnie, że pojęcie „wnioskowania" należy rozumieć w szerszym znaczeniu jako obejmujące „etap tworzenia" systemu AI. Motyw 12 aktu w sprawie sztucznej inteligencji zawiera następnie dalsze wytyczne dotyczące technik,

[5] ISO/IEC 22989:2022, Technologie informacyjne – Sztuczna inteligencja – Pojęcia i terminologia związane ze sztuczną inteligencją.

które umożliwiają systemowi AI wnioskowanie, w jaki sposób generować wyniki. W związku z tym techniki, które można wykorzystać w celu umożliwienia wnioskowania, obejmują „mechanizmy uczenia maszynowego, które na podstawie danych uczą się, jak osiągnąć określone cele, oraz podejścia oparte na logice i wiedzy, które polegają na wnioskowaniu na podstawie zakodowanej wiedzy lub symbolicznego przedstawienia zadania, które należy rozwiązać".

32) Pierwsza kategoria technik sztucznej inteligencji, o której mowa w motywie 12 aktu w sprawie sztucznej inteligencji, to „**mechanizmy uczenia maszynowego**, które na podstawie danych uczą się, jak osiągnąć określone cele". Kategoria ta obejmuje wiele różnych mechanizmów umożliwiających systemowi „uczenie się", takich jak uczenie nadzorowane, uczenie nienadzorowane, uczenie samonadzorowane i uczenie przez wzmacnianie.

33) W przypadku **uczenia nadzorowanego** system AI uczy się na podstawie oznakowań (etykietowanych danych), przy czym dane wejściowe są sparowane z prawidłowym wynikiem. System wykorzystuje te oznakowania, aby nauczyć się odwzorowania danych wejściowych w wynikach, a następnie uogólnia je do nowych, niewidzianych wcześniej danych. Przykładem systemu uczenia nadzorowanego jest system wykrywania spamu wspomagany sztuczną inteligencją. Na etapie tworzenia system trenuje się na zbiorze danych zawierającym wiadomości e-mail, które ludzie oznaczyli jako „spam" lub „nie spam", i w ten sposób uczy się on wzorców na podstawie cech oznaczonych wiadomości e-mail. Po przejściu treningu użytkowany system jest w stanie analizować nowe wiadomości e-mail i klasyfikować je jako spam lub nie spam na podstawie wzorców, których nauczył się na podstawie danych opatrzonych etykietami.

34) Inne systemy AI oparte na uczeniu nadzorowanym to np. systemy klasyfikacji obrazów trenowane na zbiorze obrazów, przy czym każdy obraz jest opatrzony zestawem etykiet (np. przedmioty takie jak samochody), systemy diagnostyczne w wyrobach medycznych trenowane na obrazach diagnostycznych opatrzonych etykietami przez ekspertów oraz systemy wykrywania oszustw, trenowane na opatrzonych etykietami danych transakcyjnych.

35) W przypadku **uczenia nienadzorowanego** system AI uczy się na podstawie danych, które nie były etykietowane. Model jest trenowany na danych bez żadnych uprzednio zdefiniowanych etykiet ani wyników. Stosując różne techniki, takie jak analiza skupień, redukcja wymiaru, uczenie się reguł asocjacyjnych, wykrywanie anomalii czy modele generatywne, system trenuje się w zakresie wyszukiwania wzorców, struktur lub powiązań w danych bez wyraźnych wskazówek co do tego, jaki powinien być wynik. Przykładem uczenia nienadzorowanego są systemy AI wykorzystywane przez firmy farmaceutyczne do odkrywania leków. Systemy AI wykorzystują uczenie nienadzorowane (np. analizę skupień lub wykrywanie anomalii) do grupowania związków chemicznych i przewidywania potencjalnych nowych metod leczenia chorób w oparciu o ich podobieństwo do istniejących leków.

36) **Uczenie samonadzorowane** to podkategoria uczenia nienadzorowanego, w ramach której system AI uczy się na podstawie nieetykietowanych danych w sposób nadzorowany, wykorzystując same dane do tworzenia własnych etykiet lub określania celów. Systemy AI oparte na uczeniu samonadzorowanym wykorzystują różne techniki, takie jak autokodery, generatywne sieci przeciwstawne lub uczenie kontrastowe. Przykładem systemu AI opartego na uczeniu samonadzorowanym jest system rozpoznawania obrazów, który uczy się rozpoznawać przedmioty poprzez przewidywanie brakujących pikseli w obrazie. Inne przykłady to modele językowe, które uczą się przewidywać kolejny token w zdaniu, czy systemy rozpoznawania mowy, które uczą się rozpoznawać wypowiadane słowa poprzez przewidywanie kolejnej cechy akustycznej w sygnale dźwiękowym.

37) Systemy AI oparte na **uczeniu przez wzmacnianie** opierają się na danych gromadzonych w ramach własnego doświadczenia za pośrednictwem funkcji „nagradzania". W przeciwieństwie do systemów AI, które uczą się na podstawie danych etykietowanych (uczenie nadzorowane) lub wzorców (uczenie nienadzorowane), systemy AI oparte na uczeniu przez wzmacnianie bazują na doświadczeniu. System nie otrzymuje wyraźnych etykiet, lecz uczy się w drodze prób i błędów, doskonaląc swoją strategię na podstawie informacji zwrotnych uzyskiwanych z otoczenia. Przykładem systemu AI opartego na uczeniu przez wzmacnianie jest wspomagane sztuczną inteligencją ramię robota, który może wykonywać zadania takie jak chwytanie przedmiotów. Uczenie przez wzmacnianie można również wykorzystać np. do optymalizacji zaleceń dotyczących zindywidualizowanych treści w wyszukiwarkach oraz działania pojazdów autonomicznych.

38) **Uczenie głębokie** to podkategoria uczenia maszynowego, w której wykorzystuje się architekturę warstwową (sieci neuronowe) do uczenia reprezentacji. Systemy AI oparte na uczeniu głębokim mogą automatycznie uczyć się cech na podstawie surowych danych, co eliminuje potrzebę ręcznej inżynierii cech. Ze względu na liczbę warstw i parametrów trenowanie systemów AI opartych na uczeniu głębokim zazwyczaj wymaga dużych ilości danych, ale po otrzymaniu wystarczających danych potrafią one nauczyć się rozpoznawać wzorce i sporządzać prognozy z dużą dokładnością. Systemy AI oparte na uczeniu głębokim są szeroko stosowane i stanowią technologię leżącą u podstaw wielu niedawnych przełomów w dziedzinie sztucznej inteligencji.

39) Obok różnych omówionych powyżej mechanizmów uczenia maszynowego druga kategoria technik, o których mowa w motywie 12 aktu w sprawie sztucznej inteligencji, to „**podejścia oparte na logice i wiedzy**, które polegają na wnioskowaniu na podstawie zakodowanej wiedzy lub symbolicznego przedstawienia zadania, które należy rozwiązać". Zamiast wnioskować na podstawie danych, systemy AI uczą się na podstawie wiedzy, w tym reguł, faktów i powiązań zakodowanych przez ekspertów. Na bazie zakodowanej przez ekspertów wiedzy systemy te mogą „rozumować" za pomocą silników dedukcyjnych lub inferencyjnych bądź operacji takich jak sortowanie, wyszukiwanie, dopasowywanie, tworzenie łańcuchów. Wyciągając logiczne wnioski, takie systemy stosują w odniesieniu do nowych sytuacji logikę formalną, uprzednio określone reguły lub ontologie. Podejścia oparte na logice i wiedzy obejmują np.

reprezentację wiedzy, programowanie indukcyjne (indukcyjne programowanie logiczne), bazy wiedzy, silniki inferencyjne i dedukcyjne, rozumowanie (symboliczne), systemy eksperckie oraz metody wyszukiwania i optymalizacji. Na przykład klasyczne modele przetwarzania języka oparte na wiedzy gramatycznej i semantyce logicznej bazują na strukturze języka, identyfikując syntaktyczne i gramatyczne składniki zdań, aby wyodrębnić znaczenie danego tekstu. Innym ważnym przykładem systemów AI opartych na logice i wiedzy są systemy eksperckie wczesnej generacji przeznaczone do diagnostyki medycznej, rozwijane poprzez kodowanie wiedzy szeregu ekspertów medycznych i mające wyciągać wnioski na podstawie zbioru objawów danego pacjenta.

### *5.2. Systemy nieobjęte zakresem definicji systemu AI*

40) W motywie 12 wyjaśniono także, że definicja systemu AI powinna odróżniać systemy AI od „prostszych tradycyjnych systemów oprogramowania lub założeń programistycznych [...] i nie powinn[a] obejmować systemów opartych na zasadach określonych wyłącznie przez osoby fizyczne w celu automatycznego wykonywania operacji".

41) Niektóre systemy mają zdolność do wnioskowania w wąskim zakresie, ale mogą jednak nie wchodzić w zakres definicji systemu AI ze względu na swoją ograniczoną zdolność do analizowania wzorców i autonomicznego dostosowywania wyników. Systemy takie mogą obejmować:

***Systemy służące ulepszaniu optymalizacji matematycznej***

42) Systemy wykorzystywane do ulepszania optymalizacji matematycznej lub przyspieszania i przybliżania tradycyjnych, ugruntowanych metod optymalizacji, takich jak metody regresji liniowej lub logistycznej, nie wchodzą w zakres definicji systemu AI. Wynika to z faktu, że chociaż modele te mają zdolność do wnioskowania, nie wykraczają one poza „podstawowe przetwarzanie danych". Wskazówką, że system nie wykracza poza podstawowe przetwarzanie danych, może być fakt, że używa się go w ten sam sposób od wielu lat[6]. Dotyczy to np. modeli opartych na uczeniu maszynowym, które służą do przybliżania funkcji lub parametrów w przypadku problemów z optymalizacją przy jednoczesnym utrzymaniu skuteczności. Systemy te mają na celu poprawę efektywności algorytmów optymalizacji stosowanych w problemach obliczeniowych. Pomagają one np. przyspieszyć wykonanie zadań związanych z optymalizacją dzięki wyuczonym przybliżeniom, heurystyce lub strategiom wyszukiwania.

43) Na przykład w systemach modelujących zjawiska fizyczne techniki uczenia maszynowego mogą być stosowane do poprawy wydajności obliczeniowej, przyspieszając tradycyjne symulacje oparte na prawach fizyki lub szacując parametry, które są następnie wykorzystywane w sprawdzonych modelach dotyczących zjawisk fizycznych. Systemy te nie wchodziłyby w zakres definicji systemu AI. W tym przykładzie modele uczenia maszynowego przybliżają złożone procesy atmosferyczne,

[6] W każdym razie systemy, które już wprowadzono do obrotu lub oddano do użytku przed dniem 2 sierpnia 2026 r., objęte są zasadą praw nabytych zapisaną w art. 111 ust. 2 aktu w sprawie sztucznej inteligencji.

takie jak mikrofizyka chmury lub turbulencja, co pozwala na szybsze i bardziej efektywne pod względem obliczeń przygotowanie prognoz.

44) Inny przykład systemu, który nie wchodzi w zakres definicji, to satelitarny system telekomunikacyjny służący optymalizacji przydziału szerokości pasma i zarządzania zasobami. W łączności satelitarnej reagowanie w czasie rzeczywistym na zapotrzebowanie ruchu sieciowego przy zastosowaniu tradycyjnych metod optymalizacji może być trudne, zwłaszcza gdy konieczne jest dostosowanie się do różnych poziomów zapotrzebowania ze strony użytkowników w poszczególnych regionach. Modele uczenia maszynowego mogą być wykorzystywane np. do przewidywania ruchu sieciowego i optymalizacji przydziału zasobów, takich jak moc i szerokość pasma, dla transponderów satelitarnych, ze skutecznością podobną, jak w przypadku ugruntowanych metod stosowanych w tej dziedzinie.

45) Chociaż systemy te mogą obejmować mechanizmy automatycznej samokorekty, korekty takie mają na celu optymalizację funkcjonowania systemów poprzez poprawę ich wydajności obliczeniowej, a nie np. umożliwienie dostosowania ich modeli decyzyjnych w inteligentny sposób. Pod takimi warunkami można je wyłączyć z zakresu definicji systemu AI.

***Podstawowe przetwarzanie danych***

46) System podstawowego przetwarzania danych to system wykonujący uprzednio określone, wyraźne instrukcje lub operacje. Systemy te opracowuje się i wdraża w celu wykonywania zadań w oparciu o ręcznie wprowadzane dane wejściowe lub reguły, bez „uczenia się, rozumowania lub modelowania" na którymkolwiek z etapów cyklu życia systemu. Generują one wyniki, działając w oparciu o ustalone reguły zaprogramowane przez człowieka, bez wykorzystywania technik sztucznej inteligencji, takich jak uczenie maszynowe czy wnioskowanie oparte na logice. Do tych systemów podstawowego przetwarzania danych należą np. systemy zarządzania bazami danych wykorzystywane do sortowania lub filtrowania danych w oparciu o określone kryteria (np. „znajdź wszystkich klientów, którzy w ostatnim miesiącu zakupili dany produkt"), standardowe aplikacje oferujące arkusze kalkulacyjne, które nie obejmują funkcji wspomaganych sztuczną inteligencją, oraz oprogramowanie obliczające na podstawie badania średnią populacyjną, którą następnie wykorzystuje się w kontekście ogólnym.

47) W zakres definicji systemu AI nie wchodzą również systemy służące wyłącznie do analizy opisowej, testowania hipotez i wizualizacji. Na przykład w przypadku oprogramowania do wizualizacji sprawozdań ze sprzedaży można zastosować metody statystyczne do stworzenia tablicy wskaźników sprzedaży, pokazującej całkowitą sprzedaż, średnią sprzedaż w podziale na regiony i tendencje w zakresie sprzedaży w czasie. Dzięki metodom statystycznym dane te można podsumowywać i obrazować na wykresach i diagramach. System nie przedstawia jednak zaleceń dotyczących tego, w jaki sposób poprawić sprzedaż ani jakie produkty promować. Innym przykładem jest system oprogramowania, który stosuje techniki statystyczne do badań opinii publicznej lub do wyników ankiet w celu określenia ich ważności, wiarygodności, korelacji i znaczenia statystycznego. W przypadku takich systemów nie ma mowy o „uczeniu się,

rozumowaniu lub modelowaniu" – po prostu przedstawiają one dane o charakterze informacyjnym.

***Systemy oparte na klasycznej heurystyce***

48) Klasyczna heurystyka obejmuje techniki rozwiązywania problemów bazujące na metodach opartych na doświadczeniu, mające na celu skuteczne znajdowanie przybliżonych rozwiązań. Techniki heurystyczne stosuje się powszechnie w kontekście programowania, gdy znalezienie dokładnego rozwiązania jest niepraktyczne ze względu na ograniczenia czasowe lub związane z zasobami. Klasyczna heurystyka zazwyczaj obejmuje podejścia oparte na regułach, rozpoznawanie wzorców lub strategie prób i błędów, a nie uczenie się oparte na danych. W przeciwieństwie do nowoczesnych systemów uczenia maszynowego, które dostosowują swoje modele w oparciu o relacje między danymi wejściowymi a wynikami, klasyczne systemy heurystyczne poszukują rozwiązań, stosując uprzednio zdefiniowane reguły lub algorytmy. Na przykład program szachowy wykorzystujący algorytm minimax z heurystycznymi funkcjami oceny może oceniać pozycje na szachownicy bez wcześniejszego uczenia się na podstawie danych. Chociaż metody heurystyczne są skuteczne w wielu zastosowaniach, mogą wykazywać niezdolność do dostosowania się i uogólniania w porównaniu z systemami AI, które wyciągają wnioski z doświadczenia.

***Proste systemy prognostyczne***

49) Wszystkie systemy maszynowe, które wymaganą skuteczność mogą osiągnąć dzięki podstawowej zasadzie uczenia się statystycznego, można wprawdzie sklasyfikować pod względem technicznym jako oparte na mechanizmach uczenia maszynowego, ale ze względu na swoją skuteczność nie wchodzą w zakres definicji systemu AI.

50) Na przykład w prognozach finansowych (podstawowa analiza porównawcza) takich systemów maszynowych można używać do przewidywania przyszłych cen akcji poprzez zastosowanie estymatora wykorzystującego strategię „średniej" w celu ustalenia prognozy bazowej (np. zawsze przewidując średnią cenę historyczną). Takie podstawowe metody analizy porównawczej pomagają ocenić, czy bardziej zaawansowane modele uczenia maszynowego mogłyby wnieść wartość dodaną. Inny przykład to wykorzystanie średniej temperatury z minionego tygodnia do przewidywania temperatury w kolejnym dniu. Taki podstawowy system służy wyłącznie do szacowania średnich, ale nie osiąga on skuteczności bardziej złożonych systemów prognozowania szeregów czasowych, co wymagałoby bardziej zaawansowanych modeli.

51) Systemy szacowania statycznego, takie jak system przewidywania czasu reakcji w obsłudze klienta, które wykorzystują szacunki statyczne do przewidywania średniego czasu rozwiązania problemu na podstawie danych z przeszłości, oraz typowe predyktory, np. służące prognozowaniu zapotrzebowania w sklepie w celu przewidzenia, ile sztuk danego produktu sklep będzie sprzedawał każdego dnia, to inne przykłady rozwiązań pomagających w ustaleniu poziomu bazowego lub poziomu odniesienia, np. poprzez przewidywanie średniej.

### *6. Wyniki, które mogą wpływać na środowisko fizyczne lub wirtualne*

52) Zgodnie z szóstym elementem definicji systemu AI w art. 3 pkt 1 aktu w sprawie sztucznej inteligencji system taki wnioskuje, „jak generować [...] wyniki, takie jak predykcje, treści, zalecenia lub decyzje, które mogą wpływać na środowisko fizyczne lub wirtualne". Zdolność systemu do generowania wyników, takich jak predykcje, treści i zalecenia, na podstawie otrzymanych danych wejściowych i z wykorzystaniem uczenia maszynowego oraz podejść opartych na logice i wiedzy, ma zasadnicze znaczenie dla tego, co robią systemy AI i co odróżnia te systemy od innych form oprogramowania. Zdolność do generowania wyników i rodzaj wyników, jakie system może generować, mają kluczowe znaczenie dla zrozumienia funkcjonowania i wpływu systemu AI.

53) Wyniki działania systemów AI należą do czterech szerokich kategorii wymienionych w art. 3 pkt 1 aktu w sprawie sztucznej inteligencji: predykcji, treści, zaleceń i decyzji. Każda kategoria różni się pod względem poziomu interwencji człowieka.

54) **Predykcje** należą do najczęstszych wyników generowanych przez system AI i wymagają najmniejszej interwencji człowieka. Predykcja to oszacowanie wartości nieznanej (wyniku) na podstawie znanych wartości dostarczonych do systemu (dane wejściowe). Systemy oprogramowania wykorzystuje się do generowania predykcji od dziesięcioleci. Systemy AI wykorzystujące uczenie maszynowe są w stanie generować predykcje, które ujawniają złożone wzorce w danych i oferują dokładne prognozy w bardzo dynamicznych i złożonych środowiskach.

55) Na przykład systemy AI stosowane w samochodach autonomicznych projektuje się w taki sposób, aby dokonywały w czasie rzeczywistym prognoz w niezwykle złożonym i dynamicznym środowisku, z wieloma rodzajami czynników i interakcji oraz praktycznie nieskończoną liczbą możliwych sytuacji, a także aby podejmowały decyzje o odpowiednim dostosowaniu swojego zachowania. Systemy, które nie bazują na sztucznej inteligencji, zazwyczaj oparte na danych historycznych, danych naukowych lub uprzednio określonych regułach, takie jak niektóre niewykorzystujące sztucznej inteligencji systemy eksperckie w wyrobach medycznych, nie są w stanie uporać się z takim stopniem złożoności. Podobnie systemy AI kontrolujące zużycie energii mają na celu oszacowanie jej zużycia poprzez analizę danych z inteligentnych liczników, prognoz pogody i wzorców zachowań konsumentów. Opierając się na mechanizmach uczenia maszynowego, systemy AI projektuje się w taki sposób, aby znajdować między tymi zmiennymi złożone korelacje, co pozwoli tworzyć dokładniejsze prognozy zużycia energii.

56) **Treści** odnoszą się do generowania przez system AI nowych materiałów. Mogą to być teksty, obrazy, nagrania wideo, muzyka i inne rodzaje wyników. Rośnie liczba systemów AI, w których do generowania treści wykorzystuje się modele uczenia maszynowego (na przykład oparte na technologiach wstępnie wytrenowanego transformatora generatywnego (GPT)). Chociaż z technicznego punktu widzenia treści, jako kategorię wyników, można rozumieć jako sekwencję „predykcji" lub „decyzji", ze względu na

rozpowszechnienie tych wyników w systemach generatywnej AI wymieniono je w motywie 12 aktu w sprawie sztucznej inteligencji jako oddzielną kategorię wyników.

57) **Zalecenia** odnoszą się do sugestii dotyczących konkretnych działań, produktów lub usług, które to sugestie przedstawiane są użytkownikom odpowiednio do ich preferencji, zachowań lub innych danych wejściowych. Podobnie jak w przypadku predykcji, zarówno systemy oparte na sztucznej inteligencji, jak i systemy nieoparte na sztucznej inteligencji można projektować w taki sposób, aby generowały zalecenia. Na przykład systemy rekomendacji oparte na sztucznej inteligencji mogą wykorzystywać duże zbiory danych, dostosowywać się do zachowań użytkowników w czasie rzeczywistym, zapewniać wysoce spersonalizowane zalecenia i efektywnie zwiększać swoją skalę w miarę rozrastania się zbioru danych – są to funkcje rzadko zapewniane przez systemy nieoparte na sztucznej inteligencji, które wykorzystują statyczne mechanizmy bazujące na regułach i ograniczone ilości danych. W innych przypadkach zalecenia odnoszą się do potencjalnych decyzji, np. dotyczących zatrudnienia kandydata w systemie rekrutacji, które będą oceniane przez ludzi. Jeżeli zalecenia te stosuje się automatycznie, stają się decyzjami.

58) **Decyzje** odnoszą się do wniosków wyciągniętych przez system lub dokonanych przez niego wyborów. System AI, który generuje decyzje, automatyzuje procesy, które tradycyjnie opierają się na osądzie człowieka. Taki system oznacza pełną automatyzację procesu, w wyniku czego w środowisku otaczającym system generowany jest określony wynik bez interwencji człowieka.

59) Podsumowując, systemy AI, w tym systemy oparte na mechanizmach uczenia maszynowego oraz podejściach opartych na logice lub wiedzy, różnią się od systemów nieopartych na AI pod względem zdolności generowania wyników, takich jak predykcje, treści, zalecenia i decyzje, ponieważ potrafią radzić sobie ze złożonymi relacjami i wzorcami w danych. Systemy AI mogą zasadniczo generować wyniki bardziej zniuansowane niż inne systemy, na przykład wykorzystując wzorce wyuczone podczas treningu lub stosując reguły określone przez ekspertów do podejmowania decyzji, co pozwala na bardziej zaawansowane rozumowanie w ustrukturyzowanych środowiskach.

### 7. Interakcje ze środowiskiem

60) Zgodnie z siódmym elementem definicji systemu AI wyniki generowane przez ten systemu „**mogą wpływać na środowisko fizyczne lub wirtualne**”. Element ten należy rozumieć jako podkreślenie faktu, że systemy AI nie są pasywne, lecz mają aktywny wpływ na środowisko, w którym są stosowane. Odniesienie do „środowiska fizycznego lub wirtualnego” wskazuje, że system AI może mieć wpływ zarówno na materialne obiekty fizyczne (np. ramię robota), jak i na środowisko wirtualne, w tym przestrzenie cyfrowe, przepływy danych i ekosystemy oprogramowania.

## III. Uwagi końcowe

61) Definicja systemu AI obejmuje szerokie spektrum systemów. Przy ustalaniu, czy system oprogramowania jest systemem AI, należy opierać się na konkretnej architekturze i funkcjach danego systemu i uwzględniać siedem elementów definicji zawartej w art. 3 pkt 1 aktu w sprawie sztucznej inteligencji.

62) Nie jest możliwe automatyczne określanie systemów, które wchodzą lub nie wchodzą w zakres definicji systemu AI, ani sporządzenie wyczerpujących wykazów takich systemów.

63) Obowiązkom regulacyjnym i nadzorowi na mocy aktu w sprawie sztucznej inteligencji podlegają jedynie niektóre systemy AI. Podejście oparte na analizie ryzyka, o którym mowa w akcie w sprawie sztucznej inteligencji, oznacza, że zakazami określonymi w art. 5 aktu, regulacjami dotyczącymi systemów AI wysokiego ryzyka, o których mowa w art. 6 aktu, oraz wymogami w zakresie przejrzystości dotyczącymi ograniczonej liczby uprzednio określonych systemów AI, ustanowionymi w art. 50 aktu, będą objęte jedynie te systemy, które stwarzają najpoważniejsze ryzyko dla podstawowych praw i wolności. Zdecydowana większość systemów – nawet jeśli w rozumieniu art. 3 pkt 1 aktu w sprawie sztucznej inteligencji kwalifikują się one jako systemy AI – nie będzie podlegać żadnym wymogom regulacyjnym na mocy aktu.

64) Akt w sprawie sztucznej inteligencji ma również zastosowanie do modeli AI ogólnego przeznaczenia, których dotyczy jego rozdział V. Analiza różnic między systemami AI a modelami AI ogólnego przeznaczenia nie wchodzi w zakres niniejszych wytycznych.

ODPIS UWIERZYTELNIONY
W imieniu Sekretarza Generalnego

Martine DEPREZ
Dyrektor
Proces Decyzyjny i Kolegialność
KOMISJA EUROPEJSKA