SZTUCZNA INTELIGENCJA

AI Act „odroczony", ale nie w całości – jakie wymogi należy teraz stosować?

Informacja o „odroczeniu AI Act" jest prawdziwa tylko częściowo. Digital Omnibus przesunął wymagania dotyczące systemów wysokiego ryzyka, ale nie ogólną datę rozpoczęcia stosowania rozporządzenia 2024/1689 (AI Act) ani obowiązku transparentności.

DANIEL NIWIŃSKI

prawnik, ekspert ds. cyberbezpieczeństwa w Kancelarii Krzysztof Rożko i Wspólnicy

Odroczenie selektywne, a nie moratorium

AI Act, czyli rozporządzenie (UE) 2024/1689, weszło w życie 1 sierpnia 2024 r., a rozpoczęcie jego stosowania jest podzielone na etapy. Od 2 lutego 2025 r. obowiązują przepisy ogólne, zakazy niedopuszczalnych praktyk oraz regulacja kompetencji w zakresie AI. Od 2 sierpnia 2025 r. stosuje się m.in. przepisy o modelach AI ogólnego przeznaczenia (GPAI), zarządzaniu i karach.

Rozporządzenie (UE) 2026/1744, określane jako Digital Omnibus on AI, weszło w życie 27 lipca 2026 r. i wprowadza nowe terminy rozpoczęcia stosowania niektórych przepisów AI Act. Precyzyjniej jest więc mówić nie o „odroczeniu wejścia w życie AI Act", lecz o przesunięciu stosowania oznaczonych przepisów rozdziału III AI Act.

Wymagania zawarte w sekcjach 1–3 tego rozdziału, dotyczące klasyfikacji, wymagań i obowiązków podmiotów związanych z systemami wysokiego ryzyka, będą stosowane:

- od 2 grudnia 2027 r. – w odniesieniu do systemów określonych w art. 6 ust. 2 i załącznika III AI Act, m.in. używanych w zatrudnieniu, edukacji, podstawowych usługach, wymiarze sprawiedliwości lub migracji;
- od 2 sierpnia 2028 r. – w odniesieniu do systemów określonych w art. 6 ust. 1 AI Act powiązanych z produktami objętymi unijnym prawodawstwem harmonizacyjnym z załącznika I.

Omnibus nie ustanowił generalnego moratorium – nie przesunięto rozpoczęcia stosowania wymogów określonych w art. 50 AI Act, a znowelizowany art. 4 nadal wymaga podjęcia środków wspierających rozwój kompetencji AI personelu i innych osób obsługujących systemy w imieniu organizacji – zmianie uległ jednak charakter tego obowiązku: z obowiązku rezultatu (zapewnienia „wystarczającego poziomu" kompetencji) na obowiązek starannego działania. Od 2 sierpnia rozpoczęto też egzekwowanie przepisów o GPAI, zakazanych praktykach, kompetencjach AI i transparentności.

Art. 50: cztery obowiązki

Art. 50 AI Act nie ogranicza się do generatywnej AI ani systemów wysokiego ryzyka. Obejmuje cztery sytuacje, a rolę dostawcy lub podmiotu stosującego ocenia się dla konkretnego systemu i zastosowania.

Po pierwsze, dostawca systemu przeznaczonego do bezpośredniej interakcji z człowiekiem powinien zaprojektować go tak, aby osoba została poinformowana, że kontaktuje się z AI. Dotyczy to m.in. chatbotów, asystentów głosowych, awatarów i niektórych agentów. Informację należy przekazać najpóźniej przy pierwszej interakcji. Wyjątek stanowi sytuacja, gdy charakter interakcji użytkownika z AI jest oczywisty dla osoby rozsądnie poinformowanej, uważnej i przezornej.

Po drugie, dostawca systemu generującego syntetyczny dźwięk, obraz, wideo lub tekst, także opartego na GPAI, powinien zapewnić maszynowo czytelne oznaczanie rezultatów i wykrywalność ich sztucznego pochodzenia. Rozwiązania mają być, w granicach technicznej wykonalności, skuteczne, interoperacyjne, odporne i niezawodne. Obowiązek nie obejmuje standardowej edycji ani zmian niemodyfikujących istotnie danych wejściowych lub ich znaczenia.

Tylko w tym drugim przypadku przewidziano okres przejściowy: dostawcy systemów wprowadzonych do obrotu przed 2 sierpnia 2026 r. mają czas na wykonanie art. 50 ust. 2 AI Act do 2 grudnia 2026 r. Przesunięcie nie dotyczy nowych systemów ani pozostałych obowiązków transparentności. Nie zwalnia również podmiotów stosujących z oznaczania deepfake'ów czy określonych publikacji tekstowych już od 2 sierpnia 2026 r.

Po trzecie, podmioty stosujące systemy rozpoznawania emocji lub kategoryzacji biometrycznej powinny informować osoby narażone na ich działanie. Organizacje stosujące wyżej wskazane systemy nadal oczywiście muszą przestrzegać pozostałych wymogów regulacyjnych, w tym RODO lub innych właściwych przepisów: komunikat o AI nie legalizuje przetwarzania danych biometrycznych ani wnioskowania o emocjach. Po czwarte, podmiot stosujący system generujący lub modyfikujący deepfake powinien ujawnić sztuczne pochodzenie materiału. Dla dzieł ewidentnie artystycznych, satyrycznych lub fikcyjnych informację można podać tak, aby nie zakłócała odbioru. Pamiętać przy tym należy, iż nie wszystkie fikcyjne treści będą zawsze stanowiły deepfake. Jeśli dana postać jest na tyle abstrakcyjna, że nie ma potencjału, aby została uznana za autentyczną lub prawdziwą (np. słoń prowadzący samochód), wówczas taka treść nie powinna zostać uznana za deepfake, a przez to obowiązek oznaczania takiej treści nie będzie miał zastosowania. Obowiązek obejmuje także tekst publikowany w celu informowania społeczeństwa o sprawach interesu publicznego. Oznaczenie nie jest wymagane po rzeczywistej kontroli człowieka lub redakcji, jeżeli osoba fizyczna lub prawna ponosi odpowiedzialność redakcyjną. Korekta językowa czy pobieżna akceptacja nie wystarczą do wykazania zgodności.

> Podmiot stosujący system generujący lub modyfikujący deepfake powinien ujawnić sztuczne pochodzenie materiału

Informacje powinny być jasne, odróżnialne, dostępne i przekazane najpóźniej przy pierwszej interakcji lub ekspozycji. Naruszenie może podlegać karze do 15 mln euro albo do 3 proc. całkowitego rocznego światowego obrotu – w zależności od tego, która z kwot jest wyższa – z uwzględnieniem szczególnych zasad dla mniejszych podmiotów. W Polsce funkcję krajowego organu nadzoru rynku w tym zakresie pełni Komisja Rozwoju i Bezpieczeństwa Sztucznej Inteligencji (KRiBSI), powołana ustawą z 27 lipca 2026 r. o systemach sztucznej inteligencji. Faktyczna działalność KRiBSI (wybór przewodniczącego i ukonstytuowanie składu) rozpocznie się jednak dopiero jesienią 2026 r., co oznacza przejściową lukę w krajowym nadzorze nad stosowaniem art. 50 AI Act.

> Podmioty stosujące systemy rozpoznawania emocji lub kategoryzacji biometrycznej powinny informować osoby narażone na ich działanie

Wytyczne i ikony

Komisja Europejska 20 lipca 2026 r. przyjęła wytyczne do art. 50 AI Act. Dokument wyjaśnia m.in. role dostawcy i podmiotu stosującego, pojęcia bezpośredniej interakcji, treści syntetycznej, standardowej edycji, deepfake'u, interesu publicznego i realnej kontroli człowieka. Zawiera też przykłady sytuacji objętych i nieobjętych przepisem. Wytyczne rozstrzygają też kwestię treści powstałych przed 2 sierpnia 2026 r. – nie ma obowiązku ich retroaktywnego oznaczania, choć KE zaleca to w miarę możliwości.

Komisja Europejska udostępniła również zestaw unijnych ikon informujących o wykorzystaniu AI podczas generowania danego materiału. Ich użycie jest fakultatywne, podczas gdy samo oznaczanie pozostaje obowiązkowe. Ikona nie tworzy automatycznego domniemania zgodności: liczą się m.in. jej widoczność, moment prezentacji, zrozumiały język, dostępność oraz zachowanie informacji po udostępnieniu lub pobraniu treści.

> Z punktu widzenia zgodności regulacyjnej kluczowym jest wykrywanie tzw. Shadow AI, czyli niezatwierdzonych przez podmiot narzędzi AI wykorzystywanych przez pracowników/współpracowników podmiotu

Jak zorganizować zgodność w praktyce?

Punktem wyjścia jest inwentaryzacja szersza niż lista zakupionych aplikacji. Powinna objąć funkcje AI w pakietach biurowych i SaaS, rozszerzenia, API, rozwiązania wewnętrzne i narzędzia dostawców działających w imieniu organizacji. W rejestrze warto ująć właściciela biznesowego, dostawcę i datę wdrożenia, użytkowników, dozwolone dane wejściowe (np. na podstawie klasyfikacji informacji), odbiorców i sposób weryfikacji przez człowieka.

Etykieta danego narzędzia AI jako „generator tekstu" nie przesądza o obowiązkach w odniesieniu do wykorzystania tego narzędzia. Kluczowym może okazać się zbadanie procesu: czy AI kontaktuje się z człowiekiem, wspiera decyzję o osobie, publikuje wynik, dotyczy interesu publicznego, używa wizerunku lub głosu oraz zachowuje oznaczenia dostawcy. Podmiot korzystający z pozyskanego z zewnątrz narzędzia będzie zwykle podmiotem stosującym, lecz może stać się dostawcą, gdy oferuje system pod własną nazwą, zleca jego opracowanie albo przejmuje tę odpowiedzialność wskutek wprowadzonych w narzędziu zmian.

Analizę regulacyjną należy połączyć z analizą biznesową – dla skutecznego wdrożenia nowych wymogów należy przede wszystkim ustalić cel, korzyść, alternatywy, poziom automatyzacji, nadzór człowieka i konsekwencje błędu. Na tej podstawie można zaprojektować kontrolę redakcyjną, zatwierdzanie publikacji, obowiązki informacyjne i wymagania umowne wobec dostawców (takie jak m.in. zachowanie metadanych, audyt, zgłaszanie zmian i incydentów oraz wsparcie zgodności).

Z punktu widzenia zgodności regulacyjnej kluczowym jest również wykrywanie tzw. Shadow AI, czyli niezatwierdzonych przez podmiot narzędzi AI wykorzystywanych przez pracowników/współpracowników podmiotu. Wykrywanie Shadow AI, dla zapewnienia skuteczności, nie powinno opierać się wyłącznie na ankiecie lub zakazie uregulowanym w treści polityki lub procedury. Warsztaty biznesowe warto połączyć z proporcjonalną analizą logów SSO, proxy i narzędzi bezpieczeństwa, zakupów, dodatków i integracji. Jeżeli dany podmiot pragnie wdrożyć narzędzia służące do monitorowania pracowników i analizy wykorzystywanych przez nich narzędzi, organizacja musi zaprojektować proces w taki sposób, aby jednocześnie respektować zasady wynikające z RODO i prawa pracy. Ryzyko wykorzystania nieautoryzowanych narzędzi AI ograniczają zwykle: zatwierdzony katalog narzędzi, szybka ścieżka oceny zastosowań i bezpieczna alternatywa dostarczona przez organizację.

Specyfika wymogów dot. AI, w tym w szczególności wymogów dot. sposobu wykorzystania narzędzi AI, zarządzania katalogiem dopuszczonych do użytkowania narzędzi AI, wymogów w odniesieniu do transparentności oraz odpowiedniej wiedzy i kompetencji osób korzystających z narzędzi AI, sprawia, że dla skutecznego dostosowania podmiotu do nowych wymogów regulacyjnych potrzebny jest skutecznie wdrożony system zarządzania (AI Governance). Skuteczny system zarządzania AI w organizacji będzie zawierał co najmniej takie elementy, jak wyznaczony właściciel AI governance, dokumentacja dot. dopuszczania i ponownej oceny narzędzi AI, zasady klasyfikacji i wykorzystania danych podmiotu w narzędziach AI, zasady publikacji treści generowanych przez AI, szkolenia, kontrola dostawców AI, obsługa incydentów związanych z AI i dowody wykonania obowiązków. Przy przetwarzaniu danych osobowych wymagane jest przeprowadzenie oceny podstawy i celu przetwarzania danych, zastosowanie zasady minimalizacji, transparentności, w uzasadnionych przypadkach powierzenie przetwarzania danych, odpowiednie uregulowanie transferu poza EOG i konieczność przeprowadzenia DPIA/FRIA. Dla podmiotów objętych DORA lub NIS2/UKSC AI może być zarazem usługą ICT oraz elementem ryzyka ze strony zewnętrznego dostawcy, ciągłości działania i bezpieczeństwa informacji.

Przesunięcie rozpoczęcia stosowania niektórych wymogów daje czas na kompleksowe wdrożenie mechanizmów zapewniających compliance w sposób holistyczny, uwzględniając cały krajobraz regulacyjny dotyczący bezpieczeństwa i ochrony danych zarówno osobowych, jak i nieosobowych. Organizacja wdrażająca nowe wymagania powinna wykazać, jakie systemy AI stosuje, w jakiej roli i dlaczego obowiązki transparentności wykonała albo uznała za niestosowane. /℗