

„Może zawierać śladowe ilości AI”

Julia Padlewska-Oleksy
Monika A. Górską
Lena Marcinowska-Boulangé
Bartosz Kurzec
Krzysztof Wojdyło
Marysia Surzycka
Szymon Furi Brzozowski
Maciej Miliszkiewicz

LIPIEC 2026

**O BOWIĄZEK OZNACZANIA TREŚCI
GENEROWANYCH AI**

Wkrótce zaczną obowiązywać kolejne przepisy AI Act. Dla branży kreatywnej – choć nie tylko dla niej – oznacza to m.in. nowe, konkretne obowiązki w zakresie oznaczania treści generowanych z wykorzystaniem systemów AI.

Jak prawidłowo stosować się do tego obowiązku, jak interpretować pojęcie deepfejk, co uznać za objęte wyjątkiem tzw. „klauzuli artystycznej” i wreszcie jak konkretnie oznaczać takie materiały – to wcale nie jest proste i budzi wiele wątpliwości.

Właśnie dlatego kancelaria [Wardyński i Wspólnicy](#) wspólnie z domem produkcyjnym i agencją twórców [SHOOTME | B Corp™](#) oraz wyspecjalizowaną komórką SHOOTME AI Labs zorganizowała webinarium mające odpowiedzieć na pytanie, co obowiązek oznaczania wygenerowanych treści oznacza w praktyce m.in. dla twórców, producentów i marek. W webinarium wzięło udział blisko 250 osób. To wystarczający dowód, że temat nie jest akademicki. A mowa tylko o jednym przepisie. Dyskusja – ponad dwugodzinna – nie wyczerpała pytań.

Nasza wspólna publikacja porządkuje kluczowe wątki rozmowy i odpowiada zbiorczo na główne pytania zadane podczas spotkania. Mamy jednak świadomość, że nie wyczerpuje tematu. Po rozpoczęciu stosowania przepisów o oznaczaniu treści generowanych przez AI spotykamy się ponownie – żeby wspólnie porozmawiać o tym, co wciąż budzi kontrowersje i pytania.

Już teraz zapraszamy Was na to spotkanie.

Do zobaczenia wkrótce.

Wardyński i Wspólnicy oraz SHOOTME

Kontekst / wstęp prawny

AI Act¹ to kluczowy akt prawny, który tworzy ramy prawne dla korzystania z systemów AI w Unii Europejskiej. Jednym z celów tej regulacji jest zapewnienie odpowiedniego poziomu przejrzystości poprzez obowiązek odpowiedniego oznaczania treści wygenerowanych z użyciem systemów AI.

AI Act zawiera kilka przepisów, które dotyczą tego obowiązku. My koncentrujemy się na jednym z nich, kluczowym dla osób wykorzystujących narzędzia AI do tworzenia treści.

Chodzi o art. 50 ust. 4 AI Act w zakresie, w jakim odnosi się do obrazów, treści audio lub wideo:

Podmioty stosujące system AI, który generuje obrazy, treści audio lub wideo stanowiące treści deepfake lub który manipuluje takimi obrazami lub treściami, ujawniają, że treści te zostały sztucznie wygenerowane lub zmanipulowane.

Przepis ten nakazuje odpowiednio oznaczać treści:

- gdy jesteśmy tzw. podmiotem stosującym (zdefiniowanym w AI Act),
- gdy stosujemy tzw. systemy AI (zdefiniowane w AI Act),
- gdy generowane obrazy, treści audio lub wideo stanowią treści **deepfake** (również zdefiniowane w AI Act).

Aby więc ustalić, czy treści, które generujemy, muszą być specjalnie oznaczane, powinniśmy rozstrzygnąć trzy najważniejsze zagadnienia:

- czy spełniamy przesłanki do uznania nas za podmiot stosujący,
- czy do wytworzenia treści wykorzystaliśmy narzędzie, które kwalifikuje się jako system AI w rozumieniu AI Act,
- czy treść może być uznana za deepfake [tam, gdzie nie cytujemy wprost przepisów, posługujemy się też spolszczoną pisownią deepfejk].

I tu przydadzą się definicje z AI Actu.

Zgodnie z art. 3(4) AI Act **podmiotami stosującymi** są osoby fizyczne lub prawne, organy publiczne, agencje lub inne podmioty, które wykorzystują

¹ Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2024/1689 z dnia 13 czerwca 2024 r. w sprawie ustanowienia zharmonizowanych przepisów dotyczących sztucznej inteligencji oraz zmiany rozporządzeń (WE) nr 300/2008, (UE) nr 167/2013, (UE) nr 168/2013, (UE) 2018/858, (UE) 2018/1139 i (UE) 2019/2144 oraz dyrektyw 2014/90/UE, (UE) 2016/797 i (UE) 2020/1828 (akt w sprawie sztucznej inteligencji) (Tekst mający znaczenie dla EOG)

system AI, nad którym sprawują kontrolę (ale nie w przypadku, gdy wykorzystują system AI w ramach osobistej działalności pozazawodowej).

Z kolei **systemy AI** zgodnie z art. 3(1) AI Act to systemy maszynowe, które zostały zaprojektowane do działania z różnym poziomem autonomii po ich wdrożeniu oraz które mogą wykazywać zdolność adaptacji po ich wdrożeniu, a także które – 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. Nie chodzi więc zatem o dowolne narzędzie informatyczne, lecz tylko takie, które spełnia warunki ze wskazanej definicji.

Zgodnie zaś z art. 3(6o) AI Act, aby dana treść była uznana za **deepfake**, musi być (i) wygenerowanym przez AI lub zmanipulowanym przez AI (ii) obrazem, treścią dźwiękową lub treścią wideo (iii) przypominającą (iv) istniejące osoby, przedmioty, miejsca, podmioty lub zdarzenia, (v) które odbiorca mógłby niesłusznie uznać za autentyczne lub prawdziwe.

Ważnym elementem art. 50 ust. 4 AI Act jest **tzew. klauzula artystyczna**, która łągodzi nieco wymogi dotyczące oznaczania treści artystycznych, nie wyłącza jednak obowiązku oznaczenia. Zgodnie z tą klauzulą: w przypadku, gdy treść stanowi część pracy lub programu o wyrażnie artystycznym, twórczym, satyrycznym, fikcyjnym lub analogicznym charakterze obowiązki w zakresie przejrzystości określone w niniejszym ustępie ograniczają się do ujawnienia istnienia takich wygenerowanych lub zmanipulowanych treści w odpowiedni sposób, który nie utrudnia wyświetlania lub korzystania z utworu.

Bardziej szczegółowe wyjaśnienia dotyczące oznaczania treści zostały zawarte w [Wytycznych Komisji Europejskiej z 20 lipca 2026 r.](#)² (dalej „**Wytyczne**”).

Innym dokumentem, do którego warto się odwołać, jest [Kodeks Praktyk](#), który zawiera szczegółowe wytyczne dotyczące technicznych aspektów oznaczania treści. Kodeks Praktyk nie jest bezwzględnie obowiązującym prawem i nie ma obowiązku jego przestrzegania. Warto jednak rozważyć formalne przystąpienie do Kodeksu, ponieważ jego przestrzeganie daje pewność, że nasze działania są zgodne z art. 50 ust. 4 AI Act.

² Annex to the Communication to the Commission. Approval of the content of the draft Communication from the Commission – Guidelines on the implementation of the transparency obligations for certain AI systems under Article 50 of Regulation (EU) 2024/1689 (the ‘AI Act’).

Na koniec uwag wstępnych warto wskazać, że za brak prawidłowego oznaczenia materiałów wygenerowanych z pomocą systemów AI grozi bardzo wysoka kara: 15 000 000 EUR lub 3% całkowitego rocznego światowego obrotu z poprzedniego roku obrotowego, w zależności od tego, która z tych wartości jest wyższa (art. 99 ust. 4 AI Act). Dlatego do obowiązku oznaczania treści radzimy podejść bardzo poważnie.

Obowiązek ten zacznie mieć zastosowanie od 2 sierpnia 2026 r.

Q 1: Czy obowiązki oznaczenia przewidziane w AI Act stosuje się także do materiałów tworzonych poza UE lub dystrybuowanych globalnie, jeśli docierają także do odbiorców w UE?

Tak. W AI Act zasadniczo decyduje nie miejsce wytworzenia treści, lecz miejsce **przeznaczenia wyniku (output)**. AI Act stosuje się zatem do materiałów tworzonych poza UE i do systemów AI zlokalizowanych w kraju trzecim, gdy wyniki (output) są przeznaczone do wykorzystania w Unii.

Sformułowanie „przeznaczone do wykorzystania w Unii” (*intended to be used in the Union*) pozostawia pewien margines interpretacyjny. Jeśli jednak treść jest aktywnie kierowana na rynek UE albo jeśli można racjonalnie przewidywać, że dana treść dotrze do odbiorców unijnych, AI Act, w tym obowiązki oznaczenia, znajdą zastosowanie. Co za tym idzie, materiały dystrybuowane globalnie również będą objęte AI Act w zakresie, w jakim trafią do odbiorców w UE.

Warto mieć w pamięci, że celem AI Act jest skuteczna ochrona odbiorców w UE, co wymaga zapobiegania obchodzeniu przepisów. Dlatego przepisy AI Act stosują się również do podmiotów z krajów trzecich i treści przez nie tworzonych, jeśli mają trafić do UE. Koncepcja ta jest analogiczna do chyba już powszechnie znanych zasad RODO – i celowo spójna z tamtą logiką.

Q 2: Co z materiałami opublikowanymi przed 2 sierpnia 2026 r., ale emitowanymi, przedłużanymi albo ponownie wykorzystywanymi po tej dacie?

W tym zakresie wskazówkę interpretacyjną zawierają Wytyczne, zgodnie z którymi materiały wytworzone i opublikowane przed 2 sierpnia 2026 r. nie podlegają formalnemu obowiązkowi oznaczania. Mimo to Wytyczne

zachęcają do oznaczenia takich materiałów przez podmioty, które je wytworzyły lub są w ich posiadaniu.

Jeżeli jednak jakieś materiały zostały wygenerowane przed 2 sierpnia, ale nie zostały przed tą datą opublikowane (pierwsza publikacja ma nastąpić w dniu 2 sierpnia 2026 r. lub po tej dacie), należy oznaczyć je zgodnie z regułami wynikającymi z AI Act.

Q 3: Jak traktować materiały hybrydowe: realni aktorzy/ modelki + wygenerowane tło, scenografia albo pojedyncze elementy AI?

Aby odpowiedzieć na to pytanie, trzeba w pierwszej kolejności ustalić, czy mamy do czynienia z deepfejkem.

Deepfejk to wygenerowane lub zmanipulowane przez AI obrazy, treści dźwiękowe lub treści wideo, które przypominają istniejące osoby, przedmioty, miejsca, podmioty lub zdarzenia, które odbiorca mógłby niesłusznie uznać za autentyczne lub prawdziwe.

Cztery warunki muszą być zatem spełnione **łącznie**, aby wygenerowany obraz, dźwięk lub wideo można było uznać za deepfejk:

- [obraz, dźwięk, wideo] przypomina – przez co rozumie się podobieństwo (nie musi być identyczność),
- istniejące – przez co rozumie się, że treść (kontent) obrazu, dźwięku, wideo przypomina osoby lub rzeczy, które istnieją lub mogłyby istnieć w rzeczywistości,
- osoby, obiekty, miejsca, podmioty lub zdarzenia – w odniesieniu do pojęcia osoby rozumie się realistyczne istoty ludzkie (w tym cyfrowe repliki prawdziwych osób, realistyczne awatary lub postacie ludzkie wygenerowane przez AI, a także cechy charakterystyczne, takie jak wizerunek, głos, zachowanie itp.),
- które fałszywie sprawiałyby wrażenie autentycznego lub prawdziwego dla osoby oglądającej – przez co rozumie się wprowadzenie w błąd osoby oglądającej co do prawdziwości treści, zawartości obrazu lub wideo.

deepfejk	nie deepfejk
 Obraz wygenerowany przez Midjourney (źródło: Forbes). Przypomina istniejącą osobę – papieża i fałszywie sprawia wrażenie autentycznego dla osoby go oglądającej – bo można byłoby założyć, że zimą papież chodzi w białym puchowym płaszczu	 Obraz wygenerowany przez Gemini (opracowanie własne). Przypomina istniejące miejsce – Paryż, nie sprawia jednak wrażenia autentycznego dla osoby go oglądającej – bo nie istnieją sfinksy, a tym bardziej latające sfinksy w Paryżu.
Wideo wygenerowane przez AI, przedstawiające osobę przypominającą konkretnego polityka, wygłaszającego przemówienie przed publicznością.	Kreskówka wygenerowana przez AI na podstawie istniejącego już obrazu przedstawiającego wydarzenie historyczne.
Wideo wygenerowane przez AI, przedstawiające wygenerowanego przez AI influencera lub celebrytę w kontekście reklamowym lub promocyjnym.	Wideo wygenerowane przez AI, przedstawiające myszy kłócące się w ludzkim języku o to, jaki ser jest najlepszy, w ramach kampanii reklamowej producenta serów.

Dla materiałów hybrydowych, tj. łączących elementy wygenerowane w AI i realne obiekty lub realnych (żywych) modeli, ważna jest przede wszystkim możliwość wywołania fałszywego wrażenia autentyczności.

Niestety nie ma ustalonej „miarki” lub proporcji przesądzającej, „ile AI” w obrazie lub wideo powoduje, że mamy do czynienia z deepfejkiem – każdy przypadek musi być oceniany indywidualnie.

Poniżej omawiamy kilka przypadków dotyczących materiałów o hybrydowym charakterze:

Realna osoba + wygenerowane tło lub scenografia

Sam fakt, że osoba (model) jest prawdziwa, nie wyklucza obowiązku oznaczenia dla deepfejków. Znaczenie ma to, czy osoba oglądająca obraz lub wideo może zostać wprowadzona w błąd co do autentyczności przedstawionej sceny.

Przykładowo: modelka + realistyczne (podmienione przez AI) tło kawiarni w Paryżu to prawdopodobnie deepfejk wymagający oznaczenia, natomiast modelka + futurystyczne (wygenerowane przez AI) tło w stylu sci-fi (jednoznaczna fikcja) prawdopodobnie nim nie jest.

Drobne, pojedyncze elementy AI

Jeżeli element wygenerowany przez AI jest marginalny, nie zmienia znaczenia przekazu i nie wpływa na odbiór autentyczności, argument za obowiązkiem oznaczania przewidzianego w AI Act jest słabszy, ale nie wykluczony. Inaczej jest, gdy pojedynczy element jest kluczowy dla przekazu – np. broń w ręku osoby, tłum demonstrantów, zniszczenia wojenne czy egzotyczna lokalizacja. W takim przypadku nawet „tylko jeden element” może całkowicie zmieniać odbiór materiału i wywoływać obowiązek odpowiedniego oznaczenia.

Q 4: Kiedy użycie AI w postprodukcji oznacza treść wygenerowaną lub zmodyfikowaną przez AI, a kiedy jest tylko standardową obróbką?

Przepisy AI Act w tym zakresie są mocno enigmatyczne.

AI Act nie rozstrzyga, czy czynności składające się na edycję zdjęcia (np. podniesienie rozdzielczości, wyostrenie, usunięcie elementów tła czy poprawa kolorów) sprawiają, że obrazy, treści dźwiękowe lub treści wideo stają się deepfejkiem, choć odpowiedzi kryją się w definicji deepfejka. Aby doprowadzić do powstania deepfejka, czynności edytorskie musiałyby prowadzić do wygenerowania lub zmanipulowania treści. Manipulowanie treścią polegałoby na przekłamaniu treści oryginalnej w sposób, który sugeruje inną sytuację niż pierwotna, zmieniając sens i wagę emocjonalną dla odbiorcy.

- Prosta edycja ograniczająca się do np. odsumienia obrazu, której stopień nie powoduje „dotworzenia” do obrazu nowych elementów, które potencjalnie nie istnieją na oryginalnym obrazie, **nie tworzy deepfejka**.
- Usuwanie osób ze zdjęć (pomijając klauzulę artystyczną) w zakresie, w jakim nie powoduje to przekłamania przekazu płynącego ze zdjęcia, **nie musi tworzyć deepfejka**. Ale np. usuwanie osób z plaży, tak by nie było nikogo

w tle, już **może stanowić deepfejk**, jeśli stwarza wrażenie, że miejsce jest puste, a naturalnie jest bardzo tłoczne.

- Korekcja kolorów nie będzie deepfejkiem w zakresie, w jakim nie zmieniamy odbioru treści (np. „podkręcenie” nasycenia krajobrazu, aby uwydatnić jego piękno lub nałożenie LUT-u, by dopasować kolory do reszty filmu). Jako przykład deepfejka w tym zakresie można wskazać np. zdjęcie/ nagranie spotkania polityków, w którym korekcja kolorów powoduje odczucie grozy i nieprzyjemnej atmosfery, choć realna sytuacja była zgoła odmienna, lub uwydatnia cechy polityka, przekłamując jego stan zdrowia.
- Problematiczną kwestią jest „rozszerzanie obrazów” przy pomocy AI. Bezpiecznie byłoby przyjąć, że rozszerzanie otoczenia (które nie jest łatwo rozpoznawalne, np. łąki) jest dopuszczalne do momentu, w którym rozszerzamy je o podobne elementy (więcej łąki, więcej lasu, więcej identycznej tekstury w tle). Taka edycja nie stanowi deepfejka. Deepfejkiem może być jednak rozszerzanie obrazu o nowe elementy, które mogą mylić odbiorcę (np. rozszerzenie łąki tak, by było widać Zamek Królewski na Wawelu).

Jak widać nie jest to sytuacja zero-jedynkowa. Należy brać pod uwagę kontekst swoich działań i stopień zmian, które powodują.

Q 5: Jak powinno wyglądać oznaczenie treści AI w zależności od formatu: video, internet, TV, radio, piosenka, teledysk?

Zgodnie z Wytycznymi niezależnie od zastosowanego formatu oznaczenie powinno być czytelne i wyraźne w odbiorze (dźwiękowym lub wizualnym) oraz sprawiać, że odbiorca jest w stanie dowiedzieć się, że treść jest deepfejkiem w pierwszym momencie kontaktu z treścią. Informacja o tym, że fragment treści stanowi deepfejk, powinna być przekazana w sposób wyróżniający się względem innych informacji i otoczenia, w którym treść jest prezentowana.

Należy pamiętać, że treści mogą trafiać do odbiorców w różnym wieku, zatem należy zadbać, by oznaczenie było czytelne i łatwo zrozumiałe dla takich grup jak dzieci i osoby starsze.

Zgodnie z Wytycznymi informacja o deepfejkach nie będzie uznana za przekazaną w sposób czytelny, jeśli odbiorca treści może łatwo ją przeoczyć lub pominąć podczas normalnej ekspozycji lub interakcji z treścią. Jako przykład podano m.in. zawarcie wzmianki, że treści stanowią deepfejk, jedynie w warunkach użytkowania portalu.

AI GENERATED**AI MODIFIED**

Przykłady ikon, którymi można oznaczać treści wygenerowane lub zmodyfikowane przez AI.

W odniesieniu do treści, do której nie ma zastosowania klauzula artystyczna, oznaczenia mogłyby być prezentowane następująco.

Treści wideo powinny być oznaczane od pierwszej klatki deepfejk, najlepiej przez całą jego długość z racji, że odbiorca może przewinąć materiał (jeśli forma emisji na to pozwala) lub wejść do pomieszczenia w jego trakcie. Należy unikać sytuacji, gdy pierwszy kontakt odbiorcy mógłby nastąpić w momencie, gdy nie jest wyświetlana informacja o tym, że wideo zostało stworzone lub zmanipulowane przez AI.

Kodeks Praktyk w tym zakresie zaleca, aby umieścić na ekranie ikonę z akronimem AI (przygotowaną odgórnie lub własną, jeśli zapewnia odpowiednią czytelność). Zalecane jest – zależnie od sytuacji – dodanie informacji „generated” lub „modified” (po angielsku lub w języku, w którym wyświetlana jest treść). Nie ma narzuconego rozmiaru ikony – musi być ona czytelna.

Na platformach, które umożliwiają ingerowanie w odtwarzacz, zalecane jest dodawanie osobnej warstwy (*second layer*), która pozwoli np. po najechnaniu myszką rozwinąć akronim AI i wyjaśnić, które momenty lub sceny i w jakim zakresie zostały zmodyfikowane.

Przy transmisjach live, które zawierają elementy deepfejk, oznaczenie powinno pojawiać się na ekranie najpóźniej w momencie, w którym pojawia się deepfejk. Dodatkowo warto dodać informację (dźwiękową i w formie obrazu) na początku transmisji z oznaczeniem, że niektóre fragmenty będą zawierać treści stanowiące deepfejk.

Obrazy – analogicznie do treści wideo.

Audio – należy zastosować prosty komunikat przed rozpoczęciem treści stanowiących deepfejk. Komunikat powinien być w języku, w którym są treści audio będące deepfejkiem. Przy dłuższych treściach można dawać wstawki w trakcie trwania deepfejka. Warto również dodać oznaczenie w formie komunikatu po zakończeniu materiału.

Jeśli treści audio stanowiące deepfejk są upubliczniane w medium pozwalającym na dodatkową transmisję obrazu, Kodeks Praktyk zaleca dodanie

oznaczenia wizualnego. Należy pamiętać, by oznaczenie było czytelne – ikona „AI” na okładce albumu będącego deepfejkiem audio może sugerować, że tylko okładka została wygenerowana przez AI. Jeśli dana platforma to umożliwia, można dodawać oznaczenie w nazwie lub przy pomocy mechanizmów konkretnej platformy (jako oznaczenie dodatkowe).

Komisja Europejska zaproponowała ikony do oznaczania treści – ich pełne zestawienie znajduje się pod linkiem: [EU Icons for labelling AI-generated content | Shaping Europe’s digital future](#).

Q 6: Czy można oznaczyć cały materiał jako AI, nawet jeśli tylko część ujęć lub elementów została wygenerowana albo zmodyfikowana?

Co do zasady powinno oznaczać się treści, które stanowią deepfejk. Jednak przy krótkich i dynamicznych treściach takie oznaczenie byłoby nieczytelne dla odbiorcy. Zatem w zakresie, w jakim nie będzie to wprowadzać odbiorcy w błąd, można oznaczać całe materiały lub większe sekcje.

Jeśli istnieje możliwość wykorzystania osobnej warstwy (interaktywnej, zintegrowanej z odtwarzaczem), należałoby dokonać w niej oznaczeń szczegółowych, które sceny i elementy są wygenerowane przez AI, lub na odwrót: które sceny nie są wygenerowane ani modyfikowane, jeśli są one w mniejszości. Gdy zastosowanie osobnej warstwy nie jest możliwe, można tworzyć tablice na końcu materiału, które wyjaśniają tę kwestię.

Q 7: Kto odpowiada za oznaczenie i zgodność z prawem: klient, agencja, studio, producent, influencer czy publikujący?

Za oznaczenie odpowiada podmiot stosujący, czyli podmiot, który przy generowaniu treści wykorzystał system AI, nad którym sprawuje kontrolę. W standardowym schemacie, w którym klient zleca przygotowanie materiału reklamowego agencji, a ta wykonuje zlecenie przy pomocy swoich pracowników lub współpracowników, podmiotem stosującym będzie zwykle agencja. To ona koordynuje przygotowanie materiału i odpowiada za niego przed klientem. Podejmuje również decyzję o wykorzystaniu systemu AI oraz kontroluje sposób jego użycia.

Znacznie trudniejsze jest ustalenie, czy w analizowanym schemacie rola podmiotu stosującego może być przypisana również pracownikowi/

podwykonawcy agencji oraz klientowi. W przypadku pracownika agencji z pomocą przychodzi Wytyczne, zgodnie z którymi pracownik faktycznie pracujący nad przygotowaniem materiału i wykorzystujący do tego celu system AI nie będzie podmiotem stosującym, jeżeli działa na rzecz agencji oraz pod jej nadzorem i kierownictwem. Niestety Wytyczne nie odnoszą się wprost do przypadku twórcy związanego z agencją umową inną niż umowa o pracę. Znajdujemy w tym przypadku szereg analogii z przypadkiem pracownika, niemniej jednak z uwagi na brak jednoznacznych wytycznych w przypadku takich twórców ryzyko zakwalifikowania jako podmiot stosujący będzie wyższe.

Z kolei klient, na pierwszy rzut oka, nie wykorzystuje systemu AI ani nie sprawuje nad nim kontroli. Patrząc z tej perspektywy nie powinien zostać uznany za podmiot stosujący, co potwierdzają też kierunkowo Wytyczne. Niemniej jednak w tym przypadku zalecamy daleko idącą ostrożność. Według nas nie można całkowicie wykluczyć, że w określonych przypadkach to również klient będzie mógł zostać uznany za podmiot stosujący, w szczególności, jeżeli jest bezpośrednio zaangażowany w proces tworzenia materiału, w tym podejmuje na bieżąco decyzje na temat jego kształtu oraz narzędzi wykorzystywanych do jego tworzenia.

Podmiot stosujący musi odpowiednio oznaczyć materiał i to on może być adresatem kary administracyjnej za jego nieprawidłowe oznaczenie. Materiał niezgodny z oznaczeniem staje się też automatycznie materiałem niezgodnym z prawem, co może mieć szereg innych implikacji (np. może skutkować zablokowaniem materiału przez platformy online; emitowanie błędnie oznaczonej reklamy może być też czynem nieuczciwej konkurencji).

Dodatkowo podmioty stosujące, które uczestniczą w złożonych łańcuchach produkcji i dystrybucji treści, powinny podjąć odpowiednie działania, aby oznaczenia treści, wymagane na podstawie art. 50 ust. 4 AI Act, były wyraźnie widoczne i łatwo rozpoznawalne dla odbiorców już przy pierwszym kontakcie z treścią. W praktyce może to wymagać dodatkowo np. wprowadzenia odpowiednich postanowień w umowach z partnerami dystrybucyjnymi.

Q 8: Kto będzie kontrolował użycie AI, jak będzie można sprawdzić sposób wykonania materiału i kto nakłada kary?

W Polsce będzie to Komisja Rozwoju i Bezpieczeństwa Sztucznej Inteligencji – nowy organ administracji, który ma zostać powołany do nadzoru nad systemami sztucznej inteligencji (powołanie tego organu przewiduje ustawa z dnia 3 lipca 2026 r. o systemach sztucznej inteligencji).

Q 9: Czy awatary lub postaci wygenerowane przez AI mogą polecać produkty, których realnie nie mogły przetestować?

AI Act nie zakazuje awatarom i postaciom generowanym przez AI polecenia produktów. Nakłada jednak obowiązek oznaczenia treści. Kwestię prawdziwości samej rekomendacji reguluje odrębnie prawo ochrony konsumentów.

Kiedy awatar AI to deepfejk wymagający oznaczenia?

AI Act obejmuje definicję deepfejka nie tylko cyfrowe repliki konkretnych, istniejących osób (np. celebrytów), ale także realistycznie wygenerowane ludzkie awatary, nawet jeśli nie przedstawiają osób faktycznie istniejących. Decydujące jest to, czy odbiorca mógłby fałszywie uznać taką osobę za istniejącą.

Niemal na pewno mamy do czynienia z deepfejkiem, jeśli awatar:

- wygląda jak prawdziwy człowiek (ma realistyczną twarz, głos, ciało),
- pojawia się w kontekście reklamowym lub promocyjnym – np. mówi do kamery o zaletach produktu, recenzuje produkt lub zachęca do zakupu.

Jako przykład deepfejka objętego obowiązkiem oznaczenia Wytyczne wskazują wideo z wygenerowanym przez AI wizerunkiem celebryty-influencera w kontekście reklamowym lub promocyjnym. Materiał musi być wtedy wyraźnie i czytelnie oznaczony najpóźniej w chwili, gdy odbiorca po raz pierwszy się z nim zetknie.



Wygenerowana influencerka reklamująca kombuchę (opracowanie własne)

AI Act i prawo ochrony konsumentów – dwa niezależne reżimy

AI Act nie reguluje tego, czy to, co awatar mówi o produkcie, jest zgodne z prawdą – ta kwestia pozostaje w gestii innych regulacji, w tym prawa ochrony konsumentów.

Problematyczne jest przede wszystkim sugerowanie osobistego doświadczenia, którego AI nie mogło mieć (np. „testowałam ten krem”, „to moje ulubione słuchawki”). Takie stwierdzenia wypowiedziane przez fikcyjną postać mogą wprowadzać odbiorców w błąd. AI-influencer może rekomendować produkty na podstawie danych, ale nie powinien przedstawiać nieistniejących osobistych doświadczeń jako faktów.

Q 10: Jak klauzula artystyczna wpływa na obowiązek oznaczania treści AI?

Jeśli deepfejk stanowi część dzieła o wyraźnie artystycznym, twórczym, satyrycznym, fikcyjnym lub analogicznym charakterze, obowiązek oznaczenia zostaje złagodzony, ale nie zlikwidowany.

Wytyczne wskazują, że:

- **Dzieła artystyczne** to dzieła tworzone dla celów sztuki (muzyka, dzieła filmowe, sztuki wizualne).
- **Dzieła twórcze** to dzieła wykazujące twórcze wybory (wykluczone są dzieła motywowane wyłącznie względami funkcjonalnymi lub technicznymi).
- **Dzieła satyryczne** to te krytykujące społeczeństwo, politykę lub osoby publiczne przy użyciu technik humorystycznych (ironia, sarkazm, pastisz).
- **Dzieła fikcyjne** to dzieła osadzone w wyobraźniowym otoczeniu.
- **Dzieła analogiczne** to takie, które dzielą kluczowe cechy z powyższymi kategoriami, choć nie przynależą do żadnej z nich wprost.

Tworzony контент może kwalifikować się do kilku z ww. kategorii. Dlatego każdorazowo konieczna będzie indywidualna analiza.

Co ważne, aby klauzula artystyczna miała zastosowanie, przynależność do którejś z ww. kategorii musi być oczywista dla odbiorcy.

Przykładowo za dzieło artystyczne można uznać wygenerowane przez AI efekty specjalne w scenach filmowych, które stanowią deepfejk, takie jak:

- odmladzanie istniejących aktorów,
- cyfrowe repliki osób zmarłych,

- zmodyfikowany za pomocą AI obraz istniejącego polityka, przedstawiający go w scenie, której celem jest wyraźnie skrytykowanie pewnych decyzji politycznych podjętych lub popieranych przez tę osobę.

Natomiast obraz znanych osób wygenerowany przez AI, sugerujący ich udział w określonych działaniach, pozbawiony jakiegokolwiek charakteru fikcyjnego, satyrycznego lub analogicznego, nie kwalifikuje się do żadnej z ww. kategorii.

Jeżeli odbiorca nie ma pewności, czy głównym celem danego dzieła było przekazanie treści artystycznych, satyrycznych itp., nie możemy skorzystać z klauzuli artystycznej.

Jeśli natomiast nie ma wątpliwości, że do dzieła można zastosować klauzulę artystyczną, to wtedy zamiast stosować pełne, widoczne oznaczenie, wystarczy ujawnić **istnienie** wygenerowanej lub zmanipulowanej treści w odpowiedni sposób, który **nie przeszkadza w odbiorze dzieła**.

Sposób i miejsce ujawnienia informacji o treści wygenerowanej przez AI zależy od medium i formatu dzieła – warunkiem jest jedynie to, by oznaczenie było możliwe do zauważenia przez odbiorcę i nie zakłócało odbioru samego dzieła. Ocena tego, co jest odpowiednim sposobem ujawnienia w danym przypadku, jest każdorazowo dokonywana z uwzględnieniem charakteru dzieła, kanału dystrybucji oraz oczekiwań odbiorców.

W każdym przypadku odbiorca musi mieć realną możliwość zapoznania się z tą informacją, zanim lub w chwili, gdy zetknie się z dziełem. Na przykład:

- w filmie kinowym – w napisach końcowych i towarzyszących materiałach dostępnych dla widza (np. program kinowy, strona dystrybutora),
- na wystawie w galerii sztuki – na tabliczce informacyjnej przy wejściu lub przy samym eksponacie, o ile jest ona wyraźnie widoczna dla zwiedzających,
- w grze wideo – w opisie produktu w sklepie, sekcji kredytów lub ekranie informacyjnym wyświetlanym przy uruchomieniu.

Dla branży reklamowej kluczowe znaczenie ma informacja, że Wytyczne **wyłączają stosowanie klauzuli artystycznej dla treści**, które mają charakter wyłącznie informacyjny lub **komercyjny**. Większość reklam nie będzie zatem mogła skorzystać z tej klauzuli.

Autorzy



Julia Padlewska-Oleksy
radca prawny
własność intelektualna



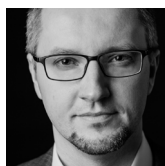
Monika A. Górka
radca prawny, partner
własność intelektualna



Lena Marcinowska-Boulangé
adwokat, partner
własność intelektualna



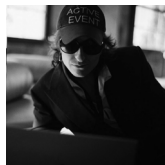
Bartosz Kurzec
młodszy prawnik
prawo nowych technologii



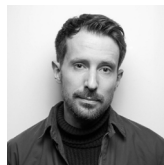
Krzysztof Wojdyło
adwokat, wspólnik
prawo nowych technologii



Marysia Surzycka
CEO & Head of Production
@surzycka



Szymon Furi Brzozowski
AI Lead
@furi.user



Maciej Miliszkiewicz
AI Artist & educator
@mmiliszkiewicz