IWONA BIEŃ-WĘGŁOWSKA

Deepfake w świetle aktu w sprawie sztucznej inteligencji

Deepfakes in the Light of the Artificial Intelligence Act

Abstrakt

The Artificial Intelligence Act is the world's first regulation to establish a comprehensive legal framework specifying harmonized rules for the creation, use, and dissemination of AI systems. One of the areas regulated by the AI Act is deepfake technology, which encompasses AI systems that generate or modify synthetic content in the form of images, sounds, or videos that falsely appear authentic. This technology contributes to numerous abuses and is consistently associated with destructive behavior. EU legislators have recognized the risks associated with deepfake content – the Artificial Intelligence Act provides procedures to protect citizens from manipulation and disinformation, classifying deepfakes as limited risk and introducing transparency obligations for providers of AI systems that generate deepfake content, as well as for those who use them. The aim of this article is to provide a detailed analysis of the provisions of the Artificial Intelligence Act relating to deepfake technology, the scope of their application, and the method of enforcement. Additionally, the assessment will focus on the categorization of risks associated with deepfake content adopted by the EU legislator and its adequacy in the context of fundamental rights protection.

SŁOWA KLUCZOWE: Unia Europejska, system sztucznej inteligencji, *deepfake*, system ograniczonego ryzyka, obowiązki w zakresie przejrzystości, manipulacja, dezinforrmacja

KEYWORDS: European Union, artificial intelligence system, deepfake, limited-risk AI system, transparency obligations, manipulation, disinformation

IWONA BIEŃ-WĘGŁOWSKA – doktor, adiunkt, Katolicki Uniwersytet Lubelski Jana Pawła II, ORCID: 0000-0002-8575-8085, e-mail: iweglowska@kul.pl

HTTPS://DOI.ORG/10.36128/PRIW.VI58.1511

1 Wstęp

Parlament Europejski i Rada (UE) wydały rozporządzenie 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 zwane aktem w sprawie sztucznej inteligencji (dalej: akt w sprawie sztucznej inteligencji lub AI Act)[1]. Regulacja oficjalnie weszła w życie 1 sierpnia 2024 r., przy czym obowiązywanie poszczególnych przepisów zostało rozłożone w czasie i następuje zgodnie z wyznaczonym przez unijnego prawodawcę harmonogramem. Akt w sprawie sztucznej inteligencji stosuje się zasadniczo od 2 sierpnia 2026 r.[2] Przedmiotem unijnego rozporządzenia jest harmonizacja przepisów dotyczących wprowadzania do obrotu, oddawania do użytku i wykorzystania systemów sztucznej inteligencji (dalej również: AI) w Unii Europejskiej. AI Act został ukształtowany w oparciu o zasadę ryzyka i kategoryzuje systemy AI z punktu widzenia potencjału ich zagrożenia dla zdrowia, życia, bezpieczeństwa człowieka oraz jego praw podstawowych[3]. Intencją unijnego prawodawcy jest bowiem promowanie upowszechniania idei regulacji sztucznej inteligencji zorientowanej na człowieka i opartej na zasadzie transparentności. Zakres zastosowania aktu w sprawie sztucznej inteligencji obejmuje dostawców (producentów) systemów AI oraz podmioty stosujące te systemy, bez względu na miejsce ich siedziby. Rozporządzenie ma zastosowanie zarówno do podmiotów unijnych, jak i podmiotów z państw trzecich, jeżeli rezultaty działania systemów AI są wykorzystywane na obszarze Unii Europejskiej[4].

Jednym z obszarów regulowanych przez AI Act jest technologia *deepfake* obejmująca systemy AI, które generują lub modyfikują treści syntetyczne w postaci obrazów, dźwięków lub nagrań wideo – materiałów audiowizualnych, które w rezultacie przedstawiają coś, co faktycznie się nie wydarzyło. W praktyce technologia *deepfake* najczęściej służy publikowaniu w Internecie filmów, które celowo są fałszywe, ale sprawiają wrażenie autentycznych. Termin *deepfake* stanowi połączenie dwóch angielskich słów: *deep* – głęboki

1 Dz.U. UE L 2024.1689 z 12.07.2024 r., LEX/el., 2025.

2 Zgodnie z art. 113 AI Act niektóre z ustanowionych przepisów już obowiązują, np. zakazy określonych praktyk mają zastosowanie od 2 lutego 2025 r., a kary za stosowanie zakazanych praktyk mogą być nakładane od 2 sierpnia 2025 r.

3 Art. 1 AI Act.

4 Art. 2 AI Act.

oraz *fake* – fałszywy. Pierwszy element nazwy *deep* pochodzi od terminu *deep learning*, który oznacza metodę uczenia maszynowego – technologię sztucznej inteligencji wykorzystującą wielowarstwowe sieci neuronowe do samodzielnej analizy i rozpoznawania danych, co pozwala na skuteczne rozwiązywanie złożonych zadań, takich jak rozpoznawanie obrazów, przetwarzanie języka naturalnego, analiza mowy czy generowanie treści. *Deepfake* oznacza więc głębokie fałszerstwo, które jest trudne do rozpoznania przez odbiorcę ze względu na bardzo zaawansowaną technologię, która tworzy lub modyfikuje treści syntetyczne.

Ze względu na to, że technologia *deepfake* od początku zakłada zamierzoną nieprawdziwość (fałszywość) efektów działania (zamiar wprowadzenia nimi w błąd)[5], konsekwentnie kojarzona jest z destrukcyjnym potencjałem[6]. Technologia ta przyczynia się do powstawania licznych nadużyć, w tym prób manipulacji wyborczych, dyskredytowania polityków, tworzenia treści pornograficznych (tzw. *deep porn*), katalizowania schematów oszustw finansowych. Lista szkodliwych sposobów wykorzystania deepfake'ów jest długa i stale rośnie. Deepfake'i uznawane są za poważne narzędzie dezinformacji, a ich pojawienie się w przestrzeni publicznej obniża zaufanie do mediów, zwiększa poczucie braku bezpieczeństwa wśród obywateli i zagraża funkcjonowaniu demokratycznych społeczeństw. Z drugiej strony deepfake'i mogą mieć także pozytywne zastosowanie, np. w dziedzinie edukacji, rozrywki czy turystyki[7].

Unijny prawodawca dostrzegł zagrożenia związane z technologią *deepfake* – akt w sprawie sztucznej inteligencji przewiduje procedury mające na celu ochronę obywateli przed manipulacją, kwalifikując deepfake'i do kategorii ograniczonego ryzyka i wprowadzając obowiązki w zakresie przejrzystości dla dostawców systemów AI, które generują treści *deepfake*,

---

5 Zob. Kamil Szpyt, "Sztuczna inteligencja i nowe technologie (nie zawsze) w służbie ludzkości, czyli cywilnoprawna problematyka rozwoju i popularyzacji technologii *deepfake*," w *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane*, red. Kinga Flaga-Gieruszyńska, Jacek Gołaczyński i Dariusz Szostek (Warszawa: C.H. Beck, 2019), 79.

6 „Bezapelacyjnie zjawisko *deepfake* jest nośnikiem przekazywania informacji fałszywej lub domniemanej […]. Należy stwierdzić, że *deepfake* jest nowym sposobem manipulacji, który zatacza nowe obszary w świecie online". Tak Krzysztof Kiełpiński, "Deepfake jako narzędzie do przekazywania informacji fałszywej i domniemanej. Analiza prawnokarna i cybernetyczna," *Kwartalnik Krajowej Szkoły Sądownictwa i Prokuratury*, nr 3 (51) (2023): 97.

7 Zob. Mateusz Łabuz, "Deep Fakes and the Artificial Inteligence Act – An Important Signal or a Missed Opportunity?," *Policy & Internet*, vol. 16, issue 4 (2024): 784.

a także dla podmiotów je stosujących. Celem artykułu jest szczegółowa analiza przepisów aktu w sprawie sztucznej inteligencji odnoszących się do technologii *deepfake*, zakresu ich stosowania oraz sposobu egzekwowania. Dodatkowo przedmiotem oceny będzie przyjęta przez unijnego prawodawcę kategoryzacja ryzyka związanego z treściami *deepfake* oraz jej adekwatność w kontekście ochrony praw podstawowych.

2 Definicja *deepfake* na gruncie przepisów aktu w sprawie sztucznej inteligencji

Unijny prawodawca sformułował definicję *deepfake*, która odgrywa kluczową rolę w regulacjach dotyczących tworzenia treści syntetycznych, a zarazem ma istotne znaczenie dla prowadzenia pogłębionego dyskursu naukowego oraz analizy prawniczej. Zgodnie z art. 3 ust. 60 AI Act „«*deepfake*» oznacza wygenerowane przez AI 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”[8].

Wskazana definicja zawiera trzy istotne elementy: po pierwsze, związek przyczynowy między technologią AI[9] a powstaniem lub modyfikacją treści,

[8] Z treści definicji wynika, że deepfake'iem nie będzie np. prosta modyfikacja obrazu w Photoshopie, która została wykonana ręcznie – o ile nie wykorzystuje systemu AI.

[9] Art. 3 ust. 1 AI Act zawiera następującą definicję systemu AI: „«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”. Definicja systemu AI ma zasadnicze znaczenie dla stosowania aktu w sprawie sztucznej inteligencji, ponieważ tylko systemy tego rodzaju podlegają regulacjom określonym w tym rozporządzeniu. Definicja systemu AI została oparta na kluczowych cechach odróżniających systemy AI od tradycyjnych systemów oprogramowania, takich jak: zdolność do wnioskowania, autonomiczne działanie oraz zdolność adaptacyjna. Por. Damian Flisak, “Akt w sprawie sztucznej inteligencji. Komentarz praktyczny,” LEX/el., 2024, teza 3. Warto podkreślić, że unijny prawodawca ostatecznie i celowo zdecydował

po drugie, charakter treści, który odnosi się do obrazów, dźwięków lub materiałów wideo, oraz ostatni element prawdopodobieństwa wprowadzenia odbiorcy w błąd – istnieje ryzyko, że odbiorca uzna treść za autentyczną lub prawdziwą. Co ważne, definicja *deepfake* koncentruje się głównie na formach audio i wizualnych. Choć nie jest zawężona do konkretnego nośnika, to trzeba podkreślić, że nie dotyczy form tekstowych. Stanowisko prawodawcy w tym zakresie pozostaje w zgodzie z poglądami doktryny, w której przeważa przekonanie, że koncentracja wyłącznie na treściach audio i wizualnych trafniej oddaje istotę zjawiska *deepfake*[10]. Natomiast możliwość błędnego uznania treści *deepfake* za autentyczne lub prawdziwe ma charakter subiektywny i zależy od kontekstu – nie jest bowiem konieczne, by ktoś faktycznie został wprowadzony w błąd, wystarczy, że taka możliwość realnie istnieje (wskazuje na to użyte w definicji sformułowanie: „[...] odbiorca mógłby niesłusznie uznać [...]"[11]). Decydujące znaczenie ma bowiem to, jak obrazy, dźwięki czy nagrania wideo zostaną odebrane przez przeciętnego odbiorcę, którego percepcja może różnić się w zależności od reprezentowanej grupy społecznej, wykształcenia, wieku, stanu zdrowia, użytego środka komunikacji lub konkretnych okoliczności, w jakich doszło do rozpowszechnienia tych treści.

Obok definicji właściwej *deepfake* zawartej w art. 3 ust. 60 AI Act pewną formę wyjaśnienia lub opisu zjawiska *deepfake* zawiera motyw 134 AI Act, który brzmi następująco:

---

się na posługiwanie się pojęciem „system sztucznej inteligencji", a nie „sztuczna inteligencja", co wskazuje na przesunięcie jego uwagi z analizy samej technologii na praktyczne jej zastosowanie w różnych obszarach życia społecznego i gospodarczego, np. w marketingu, edukacji czy służbie zdrowia. Ponadto w nauce wciąż nie ma konsensusu co do definicji „sztucznej inteligencji", co niewątpliwie utrudnia przeniesienie tego pojęcia do aktu prawnego. Zob. Jakub Kozłowski, "Analiza i ocena zmian w Akcie o sztucznej inteligencji (AI Act)," https://law4tech.pl/1881-2/ [dostęp: 1.08.2025]; zob. także Kamil Szpyt i Artur Bilski, "Wybrane wyzwania prawne i organizacyjne związane z wdrażaniem systemów AI w działalności samorządów terytorialnych," *Prawo i Więź*, nr 1 (54) (2025): 568. Na temat konsekwencji prawnych i praktycznych braku definicji „sztucznej inteligencji" w AI Act zob. Jakub Rzymowski, "Definicja prawnicza sztucznej inteligencji na podstawie rozporządzenia PE i Rady (UE) 2024/1689 w sprawie sztucznej inteligencji," *Przegląd Prawa Publicznego*, nr 11 (2024): 56 – 63.

10 Zob. Łabuz, "Deep Fakes and the Artificial Inteligence Act – An Important Signal or a Missed Opportunity?," 787.

11 Dz.U. UE L 2024.1689, LEX/el., 2025.

> Oprócz rozwiązań technicznych wykorzystywanych przez dostawców systemu AI, podmioty stosujące, które wykorzystują system AI do generowania obrazów, treści dźwiękowych lub wideo lub manipulowania nimi, tak by łudząco przypominały istniejące osoby, przedmioty, miejsca, podmioty lub wydarzenia i które to treści mogą niesłusznie zostać uznane przez odbiorcę za autentyczne lub prawdziwe („deepfake") powinny również jasno i wyraźnie ujawnić – poprzez odpowiednie oznakowanie wyniku AI i ujawnienie, że źródłem jest AI – że treści te zostały sztucznie wygenerowane lub zmanipulowane. […][12].

Niewątpliwie celem motywu 134 jest wstępne wyjaśnienie pojęcia *deepfake*[13], zanim zostanie ono ujęte w rozporządzeniu w formie ścisłej definicji, a także zwrócenie uwagi na szkodliwy charakter zjawiska i ogólne uzasadnienie spełnienia obowiązku w zakresie przejrzystości przez dokonanie dwóch czynności: oznakowanie wyniku AI oraz ujawnienie, że treści zostały sztucznie wygenerowane lub zmanipulowane. Należy jednak zwrócić uwagę na brak spójności pomiędzy definicją *deepfake* z art. 3 ust. 60 AI Act a opisem tego zjawiska w motywie 134 AI Act, w tym ostatnim do słowa „przypominały" dodano sformułowanie „tak by łudząco"[14], które może zmieniać sens definicji, nadając podobieństwu treści wygenerowanych

---

12 Dz.U. UE L 2024.1689, LEX/el., 2025.

13 Na opis zjawiska *deepfake* w motywie 134 AI Act zwraca uwagę także M. Łabuz, który twierdzi, że na skutek słabej techniki legislacyjnej ta quasidefinicja zawiera pewne nieścisłości, a przez to odbiega w swojej treści od formalnej definicji *deepfake* zwartej w art. 3 ust. 60 AI Act, wprowadzając zbędne zamieszanie, mimo że motywy rozporządzenia powinny usuwać wątpliwości interpretacyjne i pełnić funkcję doprecyzowującą. Zob. Łabuz, "Deep Fakes and the Artificial Inteligence Act – An Important Signal or a Missed Opportunity?," 787 – 788.

14 Dz.U. UE L 2024.1689, LEX/el., 2025. Sformułowanie „tak by łudząco" w języku angielskim w oryginalnej wersji rozporządzenia AI Act wyrażone zostało w jednym wyrazie „appreciably", a motyw 134 *in principio* brzmi następująco: „Further to the technical solutions employed by the providers of the AI system, deployers who use an AI system to generate or manipulate image, audio or video content that appreciably resembles existing persons, objects, places, entities or events and would falsely appear to a person to be authentic or truthful (deep fakes) […]", Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act), https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng. [dostęp: 1.08.2025].

lub zmanipulowanych przez AI odmienne znaczenie. Wykazana nieprecyzyjność oraz rozbieżności terminologiczne mogą prowadzić do nieuzasadnionych wątpliwości interpretacyjnych oraz skutkować trudnościami w stosowaniu unijnego prawa.

3 *Deepfake* a klasyfikacja ryzyka w akcie w sprawie sztucznej inteligencji

Akt w sprawie sztucznej inteligencji został skonstruowany w oparciu o zasadę ryzyka, zgodnie z którą stopień ingerencji prawodawcy oraz zakres obowiązków dostawców oraz podmiotów stosujących systemy AI uzależnione są od poziomu potencjalnego zagrożenia, jakie dane rozwiązanie może generować dla jednostki lub społeczeństwa. W tym kontekście AI Act dzieli systemy AI na cztery kategorie:

1. zakazane praktyki;
2. systemy wysokiego ryzyka;
3. systemy związane z ograniczonym ryzykiem, wymagające obowiązków w zakresie przejrzystości;
4. systemy minimalnego ryzyka[15].

Na podstawie przepisów AI Act technologia *deepfake* została zakwalifikowana do kategorii systemów ograniczonego ryzyka, której istotą jest realizacja obowiązków związanych z przejrzystością – taka klasyfikacja budzi jednak pewne wątpliwości. Ze względu na wysoką szkodliwość treści *deepfake*, które mogą stanowić zagrożenie dla zdrowia, bezpieczeństwa, a także praw podstawowych, wydaje się wskazane umieszczenie ich w kategorii systemów wysokiego ryzyka, a nawet wyodrębnienie pewnej grupy najbardziej szkodliwych treści *deepfake* (np. medyczne deepfake'i)

15 Por. Flisak, "Akt w sprawie sztucznej inteligencji. Komentarz praktyczny," teza 4.4; por. Piotr Burczaniuk, "Tworzenie prawa sztucznej inteligencji – wyzwania i perspektywy," *Prawo i Więź*, nr 3 (50) (2024): 288.

i zakwalifikowanie ich jako praktyki zakazane[16]. W tym miejscu należy przywołać art. 5 ust. 1(a) AI Act, który zakazuje wprowadzania na rynek systemów AI, które stosują techniki podprogowe, techniki manipulacyjne lub wprowadzające w błąd, których celem jest ograniczenie zdolności do podejmowania świadomych decyzji przez daną osobę lub grupę osób, co prowadzi do znaczącej zmiany ich zachowania i jednocześnie podjęcia decyzji, która wyrządza (lub może wyrządzić) poważną szkodę[17]. Szkodę tę należy rozumieć jako szkodę majątkową, jak również szkodę na osobie (naruszenie praw człowieka oraz naruszenie praw osobistych – prawa do wizerunku lub głosu, dobrego imienia lub czci osoby fizycznej)[18].

Techniki podprogowe określone w art. 5 ust. 1(a) AI Act to sposoby oddziaływania będące poza świadomością danej osoby, a więc chodzi tu o bodźce wizualne lub dźwiękowe, które nie są dostrzegalne przez odbiorcę. Jeśli zatem system AI stworzy treści *deepfake* z użyciem technik manipulacyjnych poniżej progu świadomej percepcji, np. wykorzysta szybkie, ukryte zmiany w obrazie lub dźwięku, które będą wpływały na emocje lub decyzje odbiorcy, wówczas przepis art. 5 ust. 1(a) AI Act może mieć zastosowanie. Należy jednak podkreślić, że praktyki zakazane zostaną zaliczone do treści *deepfake* tylko w ściśle określonych przypadkach.

Unijny prawodawca nie zdecydował się na jednoznaczne zakwalifikowanie systemów generujących treści syntetyczne do systemów zakazanych lub systemów wysokiego ryzyka, choć dostrzega specyficzne zagrożenie, jakie one powodują – ostatecznie ustanowił dla nich normy, których celem

16 Por. Łabuz, "Deep Fakes and the Artificial Inteligence Act – An Important Signal or a Missed Opportunity?," 789.

17 Dz.U. UE L 2024.1689, LEX/el., 2025. Art. 5 AI Act przewiduje bezwzględne zakazy stosowania określonych systemów AI, które ze względu na swój charakter są sprzeczne z prawami człowieka i standardami bezpieczeństwa.

18 Por. Gabriela Bar, "AI Act w działaniu: co się stanie z deepfake'ami?," https://panoptykon.org/ai-act-w-dzialaniu-deepfaki [dostęp: 1.08.2025]; „Katalog dóbr osobistych jest otwarty i podlega stałemu rozwojowi, opierając się na orzecznictwie sądowym, które poszerza ochronę dóbr niematerialnych człowieka, uwzględniając zmiany ocen społecznych dotyczących chronionych wartości wynikających z postępu cywilizacyjnego i technologicznego"; Przemysław Polański, "Nowe i stare wyzwania dotyczące zwalczania bezprawnych treści w Internecie," w *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane*, red. Kinga Flaga-Gieruszyńska, Jacek Gołaczyński, Dariusz Szostek (Warszawa: C.H. Beck, 2019), 186.

jest ograniczenie ryzyka związanego z ich szkodliwym działaniem[19]. M. Łabuz zauważa, że decyzja prawodawcy o zakwalifikowaniu treści *deepfake* do kategorii ograniczonego ryzyka wynika z przekonania, że technologia umożliwiająca tworzenie takich treści nie jest zła sama w sobie, natomiast szkodliwość tych treści jest konsekwencją złego wykorzystania sytemu AI[20]. Należy jednak podkreślić, że sposób zakwalifikowania treści *deepfake* w AI Act stanowi o słabości unijnego rozporządzenia, przede wszystkim ze względu na brak faktycznego odniesienia do wysoce szkodliwych skutków tego zjawiska. W praktyce sytuacja ta obniża poczucie bezpieczeństwa prawnego.

4 Obowiązki w zakresie przejrzystości w kontekście treści *deepfake*

Konsekwencją zakwalifikowania technologii *deepfake* do systemów ograniczonego ryzyka jest związanie jej z obowiązkiem przejrzystości na podstawie art. 50 ust. 2 i 4 AI Act[21]. Celem regulacji jest stworzenie odbiorcom warunków do podejmowania decyzji dotyczących wyboru treści w sposób w pełni świadomy i autonomiczny oraz zapobieganie dezinformacji w przestrzeni cyfrowej. Akt w sprawie sztucznej inteligencji ustanawia obowiązek przejrzystości na dwóch płaszczyznach. Po pierwsze, wymaga od dostawców systemów AI oznaczania treści wygenerowanych i zmanipulowanych przez AI w formacie nadającym się do odczytu maszynowego[22], po drugie, przewiduje konieczność ujawniania przez podmioty stosujące systemy AI wyłącznie treści *deepfake* oraz tekstów publikowanych w celu informowania społeczeństwa o sprawach leżących w interesie publicznym[23].

19 Por. Jakub Kozłowski, "Akt w sprawie sztucznej inteligencji w praktyce – charakterystyka unijnej regulacji opartej na ryzyku," *Europejski Przegląd Sądowy*, nr 12 (2024): 22.

20 Zob. Łabuz, "Deep Fakes and the Artificial Inteligence Act – An Important Signal or a Missed Opportunity?," 789.

21 Dz.U. UE L 2024.1689, LEX/el., 2025.

22 Art. 50 ust. 2 AI Act.

23 Art. 50 ust. 4 AI Act.

Artykuł 50 ust. 2 AI Act *in principio* stanowi:

> Dostawcy systemów AI, w tym systemów AI ogólnego zastosowania, generujących treści w postaci syntetycznych dźwięków, obrazów, wideo lub tekstu zapewniają, aby wyniki systemu AI zostały oznakowane w formacie nadającym się do odczytu maszynowego i były wykrywalne jako sztucznie wygenerowane lub zmanipulowane. Dostawcy zapewniają skuteczność, interoperacyjność, solidność i niezawodność swoich rozwiązań technicznych w zakresie, w jakim jest to technicznie wykonalne, uwzględniając przy tym specyfikę i ograniczenia różnych rodzajów treści, koszty wdrażania oraz powszechnie uznany stan wiedzy technicznej, co może być odzwierciedlone w odpowiednich normach technicznych [...][24].

Powyższa regulacja wyraża ogólny wymóg oznaczenia treści przez dostawców[25], wskazując jedynie, że musi być ono czytelne dla maszyn i pozwalać na rozpoznanie treści syntetycznych, a służące temu rozwiązania techniczne muszą być skuteczne, interoperacyjne, solidne i niezawodne z uwzględnieniem technicznej wykonalności. Analogiczne zastrzeżenia zawiera motyw 133 AI Act, przy czym zostały one dodatkowo rozszerzone o otwarty katalog przykładowych rozwiązań technicznych umożliwiających śledzenie pochodzenia informacji w szczególności poprzez: wykorzystanie znaków wodnych, identyfikację metadanych, metody kryptograficzne służące do potwierdzania pochodzenia i autentyczności treści, metody rejestracji zdarzeń, odciski palców oraz inne skuteczne techniki[26]. W praktyce oznacza to, że dostawcy systemów AI – w tym systemów ogólnego przeznaczenia, które służą generowaniu materiałów graficznych, dźwiękowych, filmowych i tekstowych – mają obowiązek tak projektować swój system, aby automatycznie oznaczał wytwarzane w ten sposób treści. Takie oznaczenia

24 Dz.U. UE L 2024.1689, LEX/el., 2025.

25 Zgodnie z art. 3 pkt 3 AI Act „«dostawca» oznacza osobę fizyczną lub prawną, organ publiczny, agencję lub inny podmiot, które rozwijają system AI lub model AI ogólnego przeznaczenia lub zlecają rozwój systemu AI lub modelu AI ogólnego przeznaczenia oraz które – odpłatnie lub nieodpłatnie – pod własną nazwą lub własnym znakiem towarowym wprowadzają do obrotu lub oddają do użytku system AI".

26 Dz.U. UE L 2024.1689, LEX/el., 2025. W motywie 133 AI Act zasygnalizowano możliwość przyszłej koordynacji działań między UE jako organem regulacyjnym a biznesem na rzecz poświadczenia autentyczności treści cyfrowych (The Coalition for Content Provenance and Authenticity – C2PA); Łabuz, "Deep Fakes and the Artificial Inteligence Act – An Important Signal or a Missed Opportunity?," 792.

będą czytelne dla algorytmów rekomendacyjnych platform internetowych, które dzięki temu będą mogły ostrzegać swoich użytkowników, że mają do czynienia z treściami sztucznie wytworzonymi i potencjalnie zmanipulowanymi. Jeżeli platformy społecznościowe ocenią takie materiały jako mniej wiarygodne lub niosące ryzyko szkodliwości, powinny wprowadzić mechanizmy ograniczające ich wpływ na użytkowników poprzez np. techniczną możliwość ograniczenia zasięgu publikowanych treści określaną w literaturze branżowej jako tzw. *shadowban*[27]. Ponadto należy zwrócić uwagę, że dostawcy bardzo dużych platform internetowych i bardzo dużych wyszukiwarek internetowych mają obowiązek identyfikować i ograniczać ryzyka związane z rozpowszechnianiem treści sztucznie wygenerowanych lub zmanipulowanych, takich jak *deepfake*. Wymóg ten wynika z art. 35 ust. 3 lit. k) aktu o usługach cyfrowych (*Digital Services Act* – DSA)[28]. Niedopełnienie tego obowiązku może skutkować nałożeniem kar finansowych na dostawców platform, sięgających do 6% rocznego światowego obrotu przedsiębiorstwa[29].

AI Act nie jest więc jedynym unijnym aktem regulującym problematykę deepfake'ów, również akt o usługach cyfrowych normuje tę materię, kwalifikując deepfake'i jako element szerszej kategorii zagrożeń dezinformacyjnych, wobec których ustanawia ramy prewencji i nadzoru. Podkreślenia wymaga, że wymóg oznaczania treści sztucznie wygenerowanych lub zmanipulowanych na gruncie AI Act nie narusza obowiązków w tym

---

27 Zob. Bar, "AI Act w działaniu: co się stanie z deepfake'ami?".

28 Rozporządzenie Parlamentu Europejskiego i Rady (UE) 2022/2065 z dnia 19 października 2022 r. w sprawie jednolitego rynku usług cyfrowych oraz zmiany dyrektywy 2000/31/WE (akt o usługach cyfrowych), Dz.U. UE L 2022.277.1 z 27.102022 r., LEX/el., 2025 (dalej: akt o usługach cyfrowych).

Art. 35 ust. 3 lit. k) aktu o usługach cyfrowych stanowi: „Dostawcy bardzo dużych platform internetowych i bardzo dużych wyszukiwarek internetowych wprowadzają rozsądne, proporcjonalne i skuteczne środki zmniejszające ryzyko, dopasowane do konkretnego ryzyka systemowego zidentyfikowanego na podstawie art. 34, ze szczególnym uwzględnieniem wpływu takich środków na prawa podstawowe. W stosownych przypadkach takie środki mogą obejmować: […]

k) zapewnienie, aby informację – bez względu na fakt czy stanowi ona wygenerowany lub zmanipulowany obraz, treść dźwiękową lub treść wideo, które łudząco przypominają istniejące osoby, obiekty, miejsca lub inne podmioty lub zdarzenia, przez co osoba będąca ich odbiorcą niesłusznie uznaje je za autentyczne lub prawdziwe – można było rozróżnić dzięki widocznym oznaczeniom, gdy prezentuje się ją na ich interfejsach internetowych oraz dodatkowo udostępnienie łatwej w użyciu funkcji umożliwiającej odbiorcom usługi wskazanie takiej informacji".

29 Art. 52 ust. 3 aktu o usługach cyfrowych.

zakresie nałożonych na dostawców platform w ramach aktu o usługach cyfrowych, co znajduje wyraźne odzwierciedlenie w motywie 136 AI Act[30]. Należy zatem traktować je jako powinności rozszerzone, intencją unijnego prawodawcy jest bowiem stworzenie spójnego systemu, w którym przepisy obu aktów prawnych pozostają w synergii, wzajemnie się dopełniając i wzmacniając swoje działanie.

Przepis art. 50 ust. 2 AI Act *in fine* wprowadza trzy wyjątki od obowiązku oznaczania treści *deepfake*[31]. Pierwszy wyjątek dotyczy sytuacji, w której systemy AI realizują wyłącznie funkcję wspomagającą w ramach standardowych procesów edycyjnych[32]. Niewątpliwie został on sformułowany w sposób bardzo ogólny, odwołanie się do nieostrego pojęcia „standardowej edycji" może budzić poważne zastrzeżenia interpretacyjne, a nawet stwarzać ryzyko obchodzenia przepisów przez dostawców systemów AI[33]. Po drugie, obowiązek ten nie ma zastosowania w przypadkach, w których systemy AI nie zmieniają w sposób istotny przekazywanych przez podmiot stosujący danych wejściowych ani ich znaczenia. Brakuje jednak jednoznaczności co do tego, czy wyjątek nieistotnej zmiany ma zastosowanie także wtedy, gdy dane wejściowe są dostarczane w ramach osobistej działalności pozazawodowej przez użytkownika nieprofesjonalnego[34]. Trzeci wyjątek odnosi się do sposobu działania organów ścigania, które nie

30 Motyw 136 AI Act *in fine* stanowi: „[...] Wymóg oznakowania treści wygenerowanych przez systemy AI na podstawie niniejszego rozporządzenia pozostaje bez uszczerbku dla określonego w art. 16 ust. 6 rozporządzenia (UE) 2022/2065 obowiązku rozpatrywania przez dostawców usług hostingu zgłoszeń dotyczących nielegalnych treści otrzymanych na podstawie art. 16 ust. 1 tego rozporządzenia i nie powinien mieć wpływu na ocenę i decyzję w sprawie niezgodności z prawem konkretnych treści. Ocena ta powinna być dokonywana wyłącznie w odniesieniu do przepisów regulujących zgodność treści z prawem".

31 Dz.U. UE L 2024.1689, LEX/el., 2025.

32 Wyjątek w art. 50 ust. 2 AI Act został sformułowany w następujący sposób: „Obowiązek ten nie ma zastosowania w zakresie, w jakim systemy AI pełnią funkcję wspomagającą w zakresie standardowej edycji [...]".

33 Zob. Łabuz, "Deep Fakes and the Artificial Inteligence Act – An Important Signal or a Missed Opportunity?," 792.

34 Należy zauważyć, że podmiot wykorzystujący system AI w ramach osobistej działalności pozazawodowej został wyłączony z definicji podmiotu stosującego system AI zgodnie z art. 3 ust. 4 AI Act, który brzmi następująco: „«podmiot stosujący» oznacza osobę fizyczną lub prawną, organ publiczny, agencję lub inny podmiot, które wykorzystują system AI, nad którym sprawują kontrolę, z wyjątkiem sytuacji, gdy system AI jest wykorzystywany w ramach osobistej działalności pozazawodowej".

muszą oznaczać treści sztucznie wygenerowanych lub zmanipulowanych, jeśli wykorzystują je w celu wykrywania przestępstw, ich zapobiegania, prowadzenia postępowań przygotowawczych oraz ścigania ich sprawców. Wskazane wyłączenie organów ścigania z obowiązku przejrzystości w tym kontekście nie budzi zastrzeżeń, znajduje bowiem swoje *ratio legis* w specyfice zadań realizowanych przez te podmioty. Dla unijnego prawodawcy nadrzędną wartością jest skuteczność działań operacyjnych i procesowych, a nie transparentność wobec ogółu odbiorców, co można w pełni wytłumaczyć koniecznością zapewnienia bezpieczeństwa publicznego oraz sprawnego funkcjonowania wymiaru sprawiedliwości[35].

Obok dostawców systemów AI – a więc firm technologicznych, które systemy AI opracowują, a następnie wprowadzają na rynek – zobowiązanie identyfikowania treści sztucznie wygenerowanych lub zmanipulowanych mają także podmioty stosujące systemy AI, które generują obrazy, treści audio lub wideo stanowiące deepfake'i lub które takimi treściami manipulują. Obowiązek ten wynika z art. 50 ust. 4 AI Act[36]. Podmioty stosujące konkretny system AI – z wyłączeniem osób korzystających z systemu AI na użytek prywatny, poza działalnością zawodową – muszą poinformować w sposób jasny i wyraźny[37], że treści te zostały sztucznie wytworzone lub zmodyfikowane, aby nie były odbierane jako materiały wiarygodne lub prawdziwe przez użytkowników ich usług. W przypadku treści o charakterze artystycznym, twórczym, satyrycznym, fikcyjnym lub analogicznym obowiązek przejrzystości ogranicza się jedynie do ujawnienia istnienia treści wygenerowanych lub zmanipulowanych w sposób nieutrudniający korzystania z takiego utworu. Natomiast publikowanie tekstów wygenerowanych lub zmodyfikowanych przez AI w celu informowania społeczeństwa o sprawach leżących w interesie publicznym wymaga ich

35 W literaturze podkreśla się jednak, że rozszerzone zastosowania technologii sztucznej inteligencji przez organy ścigania w procesie egzekwowania prawa implikuje prowadzenie ciągłych analiz dotyczących uzyskania kompromisów między prywatnością a bezpieczeństwem publicznym, co wymaga współpracy wszystkich zainteresowanych stron, w tym reprezentantów społeczności i organów wymiaru sprawiedliwości. Zob. szerzej Norbert Malec, "Sztuczna inteligencja a bezpieczeństwo państwa," *Prawo i Bezpieczeństwo – Law & Security*, nr 1 (2) (2024): 29. W sytuacji braku szczególnej ostrożności przy wyborze narzędzi AI ceną walki z przestępczością mogą stać się prawa i wolności obywatelskie. Zob. Monika Szyłkowska, "Sztuczna inteligencja – strategiczne wyzwania i ryzyka w obszarze prewencji kryminalnej – zarys problemu," *The Prison Systems Review*, nr 3 (2024): 288.

36 Dz.U. UE L 2024.1689, LEX/el., 2025.

37 Art. 50 ust. 5 AI Act.

wyraźnego ujawnienia – chyba że treść została zweryfikowana przez człowieka, a odpowiedzialność redakcyjną ponosi osoba fizyczna lub prawna. Artykuł 50 ust. 4 AI Act zawiera także wyjątek – obowiązek ujawniania przez podmioty stosujące system AI nie ma zastosowania w przypadku, gdy wykorzystywanie jest dozwolone na mocy prawa w celu prowadzenia czynności śledczych i procesowych przez organy ścigania[38].

Analiza przepisu art. 50 ust. 2 i 4 AI Act oraz motywu 134 AI Act prowadzi do wniosku, że treści *deepfake* oraz teksty publikowane w celu informowania społeczeństwa o sprawach leżących w interesie publicznym muszą być identyfikowane dwukrotnie, a mianowicie muszą być: oznaczone w formacie nadającym się do odczytu maszynowego przez dostawców systemów AI oraz ujawnione przez podmioty stosujące systemy AI w formacie przeznaczonym dla ludzi. Obowiązki te powinny być spełnione łącznie. Z dużym stopniem pewności można także założyć, że obowiązki w zakresie przejrzystości będą realizowane przez ujawnianie informacji i stosowanie znaków wodnych[39]. Wymóg przekazania odpowiednich informacji zainteresowanym osobom fizycznym musi zostać spełniony przez dostawców oraz podmioty stosujące najpóźniej w momencie pierwszej interakcji lub pierwszego stosowania[40]. Zgodnie z art. 113 AI Act obowiązki w zakresie przejrzystości wchodzą w życie od 2 sierpnia 2026 r.[41], przy czym wskazane niejasności wynikające z interpretacji pojęć zawartych w art. 50 AI Act mogą generować niepewność zarówno podczas wdrażania tej regulacji, jak i na etapie jej późniejszego stosowania.

---

38 Dz.U. UE L 2024.1689, LEX/el., 2025.

39 Znak wodny z pewnością może pomóc w identyfikacji deepfake'ów, ale nie rozwiązuje problemu złośliwych intencji stojących za ich tworzeniem i rozpowszechnianiem, zwłaszcza w kontekście politycznym. W sferze politycznej deepfake'i mogą być wykorzystywane do fałszywego przedstawiania osób publicznych mówiących lub robiących rzeczy, których nigdy nie zrobiły, w celu wpłynięcia na opinię publiczną lub wybory. Nawet jeśli znak wodny identyfikuje deepfake'i, to takie treści nadal mogą być rozpowszechniane, aby wprowadzać odbiorców w błąd, licząc na to, że wielu z nich przeoczy lub źle zrozumie znak wodny. Zob. Justyna Lisinska i Daniel Castro, "Why AI-Generated Content Labeling Mandates Fall Short," https://www2.datainnovation.org/2024-ai-watermarking.pdf [dostęp: 01.08.2025]; o wyzwaniach technicznych związanych z realizacją obowiązku przejrzystości zob. także Mateusz Łabuz, "Regulating Deep Fakes in the Artificial Intelligence Act," *Applied Cybersecurity & Internet Governance*, nr 2 (1) (2023): 23 – 25.

40 Art. 50 ust. 5 A IAct.

41 Dz.U. UE L 2024.1689, LEX/el., 2025.

5 Odpowiedzialność finansowa za naruszenie obowiązków w zakresie przejrzystości

Niewypełnienie obowiązków nakazanych przepisami aktu w sprawie sztucznej inteligencji zagrożone jest administracyjnymi karami pieniężnymi określonymi w art. 99 AI Act[42]. Artykuł 99 AI Act określa ogólne zasady nakładania kar za naruszenia przepisów rozporządzenia. Wysokość kar uzależniona jest od rodzaju naruszenia. Nieprzestrzeganie przez dostawców systemów AI i podmiotów je stosujących obowiązków w zakresie przejrzystości określonych w art. 50 AI Act zagrożone jest karą w wysokości do 15 mln euro lub – jeśli sprawcą naruszenia jest przedsiębiorstwo – w wysokości do 3% jego całkowitego rocznego światowego obrotu z poprzedniego roku obrotowego w zależności od tego, która z tych kwot jest wyższa[43]. Dla małych i średnich przedsiębiorstw, w tym start-upów, kary są ograniczone do niższych z wymienionych kwot lub procentów[44]. Zgodnie z przepisami AI Act administracyjne kary pieniężne nakładane są przez krajowe organy nadzoru rynku[45].

Niewątpliwie ustalenie sankcji finansowych na tak wysokim poziomie wskazuje, że prawodawca unijny traktuje zjawisko *deepfake* jako istotne zagrożenie dla transparentności informacyjnej, ochrony odbiorców treści syntetycznych oraz stabilności systemu komunikacji społecznej. Przyjęta wysokość administracyjnych kar pieniężnych stanowi także element polityki prewencji i odstraszania, natomiast ocena skuteczności mechanizmu egzekwowania obowiązku przejrzystości w kształcie przewidzianym w art. 50 AI Act będzie możliwa dopiero po praktycznym wdrożeniu przepisów rozporządzenia przez państwa członkowskie UE.

42 Dz.U. UE L 2024.1689, LEX/el., 2025.
43 Art. 99 ust. 4 lit. g) AI Act.
44 Art. 99 ust. 6 AI Act.
45 Art. 99 ust. 8–9 AI Act.

6 Podsumowanie

Akt w sprawie sztucznej inteligencji jest pierwszym na skalę światową aktem prawnym, który ustanawia jednolite ramy prawne w zakresie tworzenia i korzystania z systemów sztucznej inteligencji, jednak jego wejście w życie stwarza szereg wyzwań o charakterze prawnym i technicznym. W odniesieniu do technologii *deepfake* zasadnicze wyzwania koncentrują się wokół interpretacji przepisów i precyzyjnego wyznaczenia zakresu ich stosowania. Pierwsze trudności interpretacyjne dotyczą pojęć stosowanych przez unijnego prawodawcę. Przykładem jest definicja *deepfake*, która w AI Act pojawia się dwukrotnie, najpierw jako opis zjawiska, następnie jako legalna definicja. Niestety treść obu wersji nie jest tożsama, co może wskazywać na słabą technikę legislacyjną i skutkować trudnościami w stosowaniu unijnego prawa, szczególnie że definicja *deepfake* została sformułowana w taki sposób, aby służyć jako kryterium uruchamiające obowiązek w zakresie przejrzystości. Kolejne wyzwania interpretacyjne dotyczą wyjątków od wymogu oznaczania treści *deepfake* określonych w art. 50 ust. 2 AI Act. Znaczenie takich pojęć jak: „standardowa edycja" oraz „nieznaczna zmiana danych wejściowych lub ich semantyki" pozostaje niejasna, a przecież od ich interpretacji zależy zakres i efektywność stosowania art. 50 ust. 2 AI Act.

Zakwalifikowanie treści *deepfake* do kategorii ograniczonego ryzyka związanego z obowiązkiem przejrzystości również budzi wątpliwości, gdyż niektóre materiały wytworzone lub zmanipulowane przez AI stanowią poważne zagrożenie dla wielu obszarów życia społecznego. Jednak unijny prawodawca nie zdecydował jednoznacznie o przypisaniu treści audio i wizualnych do systemów zakazanych lub systemów wysokiego ryzyka. Stwarza to pewne problemy praktyczne, a mianowicie, jak kwalifikować treści syntetyczne do poszczególnych ryzyk ustanowionych przez AI Act w sytuacji, gdy zagrażają one dobrom chronionym przez to rozporządzenie. Należy także zauważyć, że choć AI Act został ukształtowany na zasadzie ryzyka, to kwalifikując technologię *deepfake* do systemów ograniczonego ryzyka, prawodawca koncentruje się na procesie tworzenia treści, a nie na skutkach, jakie wynikają z publikacji tych materiałów – obowiązek przejrzystości polega bowiem na informowaniu, czy treści zostały sztucznie wygenerowane lub zmanipulowane. Można założyć, że wynika

to z przekonania unijnego prawodawcy, że sama technologia *deepfake* nie jest zła, natomiast to jej wykorzystanie może być niewłaściwe[46].

Ponadto stosowanie zabezpieczeń technicznych przed nadużywaniem deepfake'ów, które polegają na oznaczaniu treści syntetycznych np. za pomocą znaków wodnych, jest rozwiązaniem o ograniczonej skuteczności – może ono pomóc w identyfikacji deepfake'ów, ale nie rozwiązuje problemu złośliwych intencji stojących za ich tworzeniem. Nawet jeśli znak wodny identyfikuje deepfake'i, to takie treści nadal mogą być rozpowszechniane, a więc unijna regulacja nie eliminuje szkodliwych deepfake'ów z przestrzeni wirtualnej. Zasada przejrzystości nie zapewnia także ochrony indywidualnym ofiarom technologii *deepfake*, ponieważ akt w sprawie sztucznej inteligencji podmiotem ochrony prawnej ustanawia jedynie odbiorców treści wytworzonych lub zmanipulowanych przez AI, nie odnosi się w tym zakresie do osób, którym technologia *deepfake* mogła wyrządzić poważną szkodę.

Analiza regulacji AI Act wskazuje jednak, że unijny prawodawca traktuje zjawisko *deepfake* jako istotne zagrożenie dla transparentności informacyjnej oraz stabilności systemu komunikacji społecznej, świadczy o tym m.in. przyjęta wysokość administracyjnych kar pieniężnych za niewypełnienie obowiązków w zakresie przejrzystości. Ponadto akt w sprawie sztucznej inteligencji opowiada się za synergią z przepisami aktu o usługach cyfrowych w zakresie obowiązków nakładanych na dostawców bardzo dużych platform internetowych lub bardzo dużych wyszukiwarek internetowych do identyfikowania i ograniczania ryzyka systemowego, które może wynikać z rozpowszechniania treści sztucznie wygenerowanych lub zmanipulowanych. Przepisy obu aktów prawnych wzajemnie się uzupełniają, co zwiększa skuteczność ich regulacji.

Podsumowując, należy stwierdzić, że akt w sprawie sztucznej inteligencji łączy cele wspierania innowacji technologicznych z potrzebą ochrony społeczeństwa przed szkodliwymi skutkami treści *deepfake*, które są nośnikiem dezinformacji i współczesnym narzędziem manipulacji w świecie cyfrowym. AI Act odgrywa istotną rolę w realizacji strategii cyfrowej Unii Europejskiej, jednak o jego powodzeniu zdecyduje efektywne wdrożenie jego przepisów przez poszczególne państwa członkowskie, choć biorąc pod uwagę wyzwania prawne i techniczne związane z realizacją obowiązku przejrzystości, można wnioskować, że nie będzie to łatwe zadanie.

---

46 Zob. Łabuz, "Deep Fakes and the Artificial Inteligence Act – An Important Signal or a Missed Opportunity?," 784.

Bibliografia

Bar, Gabriela. "AI Act w działaniu: co się stanie z deepfake'ami?" https://panoptykon.org/ai-act-w-dzialaniu-deepfaki.

Burczaniuk, Piotr. "Tworzenie prawa sztucznej inteligencji – wyzwania i perspektywy." *Prawo i Więź*, nr 3 (50) (2024): 283 – 300. https://doi.org/10.36128/PRIW.VI51.707.

Flisak, Damian. "Akt w sprawie sztucznej inteligencji. Komentarz praktyczny." LEX/el., 2024.

Kiełpiński, Krzysztof. "Deepfake jako narzędzie do przekazywania informacji fałszywej i domniemanej. Analiza prawnokarna i cybernetyczna." *Kwartalnik Krajowej Szkoły Sądownictwa i Prokuratury*, nr 3 (51) (2023): 83 – 99. https://doi.org/10.53024/5.3.51.2023.

Kozłowski, Jakub. "Akt w sprawie sztucznej inteligencji w praktyce – charakterystyka unijnej regulacji opartej na ryzyku." *Europejski Przegląd Sądowy*, nr 12 (2024): 20 – 25.

Kozłowski, Jakub. "Analiza i ocena zmian w Akcie o sztucznej inteligencji (AI Act)." https://law4tech.pl/1881-2/.

Lisinska, Justyna i Daniel Castro. "Why AI-Generated Content Labeling Mandates Fall Short." https://www2.datainnovation.org/2024-ai-watermarking.pdf.

Łabuz, Mateusz. "Deep Fakes and the Artificial Inteligence Act – An important Signal or a Missed Ppportunity?" *Policy & Internet*, vol. 16, issue 4 (2024): 783 – 800. https://doi.org/10.1002/poi3.406.

Łabuz, Mateusz. "Regulating Deep Fakes in the Artificial Intelligence Act." *Applied Cybersecurity & Internet Governance*, nr 2 (1) (2023): 1 – 42. https://doi.org/10.60097/ACIG/162856.

Malec, Norbert. "Sztuczna inteligencja a bezpieczeństwo państwa." *Prawo i Bezpieczeństwo – Law & Security*, nr 1 (2) (2024): 20 – 42. https://doi.org/10.4467/29567610PIB.24.002.19838.

Polański, Przemysław. "Nowe i stare wyzwania dotyczące zwalczania bezprawnych treści w Internecie." W *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane*, red. Kinga Flaga-Gieruszyńska, Jacek Gołaczyński i Dariusz Szostek, 183 – 191. Warszawa: C.H. Beck, 2019.

Rzymowski, Jakub. "Definicja prawnicza sztucznej inteligencji na podstawie rozporządzenia PE i Rady (UE) 2024/1689 w sprawie sztucznej inteligencji." *Przegląd Prawa Publicznego*, nr 11 (2024): 56 – 63.

Szpyt Kamil i Artur Bilski. "Wybrane wyzwania prawne i organizacyjne związane z wdrażaniem systemów AI w działalności samorządów terytorialnych." *Prawo i Więź*, nr 1 (54) (2025): 565 – 596.

Szpyt, Kamil. "Sztuczna inteligencja i nowe technologie (nie zawsze) w służbie ludzkości, czyli cywilnoprawna problematyka rozwoju i popularyzacji technologii *deepfake*." W *Sztuczna inteligencja, blockchain, cyberbezpieczeństwo oraz dane osobowe. Zagadnienia wybrane*, red. Kinga Flaga-Gieruszyńska, Jacek Gołaczyński i Dariusz Szostek, 75 – 94. Warszawa: C.H. Beck, 2019.

Szyłkowska, Monika. "Sztuczna inteligencja – strategiczne wyzwania i ryzyka w obszarze prewencji kryminalnej – zarys problemu." *The Prison Systems Review*, nr 3 (2024): 273 – 290. https://doi.org/ 10.52694/ThPSR.2024.3.14.