

Gabriela Garcia Vera

Uniwersytet Rzeszowski

ORCID: 0000-0003-3334-4283

**DEEP FAKE – POSTĘP TECHNOLOGICZNY
A PRAWO KARNE****Wprowadzenie**

Rynek przeżywa obecnie zmianę w kierunku czwartej rewolucji przemysłowej – epoki definiowanej przez coraz bardziej zakorzenioną w społeczeństwie cyfryzację, zacierającą granice między tym, co fizyczne i cyfrowe¹. Tę erę definiują przełomowe technologie i innowacje, takie jak mobilny internet, roboty, pojazdy autonomiczne, a także sztuczna inteligencja, która ma szczególne znaczenie dla komunikacji międzyludzkiej. Sztuczna inteligencja, niedawny przełom technologiczny, to szybko rozwijający się obszar rynku globalnego². Sprzęt, oprogramowanie i usługi związane z aplikacjami sztucznej inteligencji zostały wycenione na 93,5 mld USD w 2021 r. i przewiduje się, że w latach 2022–2030 łącznie wzrośnie w tempie złożonej rocznej stopy wzrostu (CAGR) wynoszącej 38,1%³. Na tym rynku technologie komunikacyjne wspomagane przez sztuczną inteligencję cieszą się dużym popytem. Na przykład oczekuje się, że rynek konwersacyjnej sztucznej inteligencji (np. inteligentnych agentów wirtualnych lub chatbotów) odnotuje 20-procentowy wzrost CAGR w latach 2021–2030, osiągając wartość ponad 32 mld dolarów do 2030 r.⁴ Jednym z takich zastosowań komunikacyjnych sztucznej inte-

¹ K. Schwab, *The Fourth Industrial Revolution*. *World Economic Forum 91–93 Route de la Capite CH-1223 Cologny; World Economic Forum: Geneva, Switzerland 2016*, www.weforum.org (21.05.2023).

² U. Paschen, Ch. Pitt, J. Kietzmann, *Artificial intelligence: Building blocks and an innovation typology*, „Business Horizons” 2020, vol. 63, no. 2, s. 147–155.

³ Grand View Research, 2022, <https://www.grandviewresearch.com/industry/technology> (2.06.2023).

⁴ R. Rachita, J. Himanshu, K. Vineet, *Conversational AI Market Outlook – 2030*, <https://www.alliedmarketresearch.com/conversational-ai-market-A13682> (21.05.2023).

ligencji jest szybko rozwijająca się technologia *deep fake*, która umożliwia manipulowanie treściami komunikacyjnymi za pomocą sztucznej inteligencji w celu stworzenia wysokiego stopnia ludzkiego realizmu⁵.

Zjawisko *deep fake*

Deep fakes są często wykorzystywane na platformach komunikacji online, takich jak media społecznościowe, ze względu na ich rosnącą dostępność do tworzenia⁶. Wspierając rozwój tej nowej innowacyjnej formy komunikacji na platformach internetowych, raporty branżowe sugerują, że w 2020 r. w internecie było ponad 100 mln filmów typu *deep fake*, które zgromadziły ponad 5,4 mld wyświetleń w samym YouTube w latach 2017–2020⁷. Uznając rosnące rozpowszechnienie i wykorzystanie tej innowacji komunikacyjnej, praktycy i naukowcy zaczęli dyskutować i przyczyniać się do zrozumienia tworzenia i potencjalnych zastosowań *deep fake*'ów⁸. Pomimo niedawnych dyskusji na temat potencjalnych szans i zagrożeń związanych z *deep fake*'ami dla komunikacji nie ma jeszcze powszechnej zgody co do definicji tego zjawiska. Rzeczywiście, naukowcy zajmujący się *deep fake*'ami przyznali, że są one zdefiniowane luźno, a potrzeba bardziej konkretnej definicji⁹. Zdefiniowanie innowacyjnego zjawiska, takiego jak *deep fake*, ma kluczowe znaczenie i pozwoli na zrozumienie zasad funkcjonowania tej technologii, pomagając złagodzić fragmentację literatury¹⁰, oraz znacząco ułatwi kwestię jej uregulowania.

Ogólnie należy przyjąć, że *deep fake* to technologia wykorzystująca sztuczną inteligencję do tworzenia lub edytowania treści wideo albo obrazu w celu

⁵ J. Kietzmann, L.W. Lee, I.P. McCarthy, T.C. Kietzmann, *Deepfakes: Trick or treat?*, „Business Horizons” 2020, vol. 63, no. 2, s. 135–146.

⁶ *Ibidem*.

⁷ Sentinel, *DEEPAKES 2020: THE TIPPING POINT The Current Threat Landscape, its Impact on the U.S 2020 Elections, and the Coming of AI-Generated Events at Scale*, 2020, <https://thesentinel.ai/media/Deepfakes%202020:%20The%20Tipping%20Point,%20Sentinel.pdf> 92.06.2023).

⁸ C. Campbell, K. Kirk, S. Sands, J. Kietzmann, *Preparing for an Era of Deepfakes and AI-Generated Ads: A Framework for Understanding Responses to Manipulated Advertising*, „Journal of Advertising” 2021, vol. 51, s. 5.

⁹ L. Whittaker, R. Mulcahy, K. Letheren, J. Kietzmann, R. Russell-Bennett, *Mapping the deepfake landscape for innovation: A multidisciplinary systematic review and future research agenda*, „Technovation” 2023, vol. 125, <https://www.sciencedirect.com/science/article/abs/pii/S0166497223000950#preview-section-references> (21.05.2023).

¹⁰ *Ibidem*.

pokazania czegoś, co nigdy się nie wydarzyło¹¹. Postęp techniczny i technologiczny w ostatnich latach sprawił, że tworzenie deep fake'ów jest jeszcze łatwiejsze i szybsze, a otrzymane rezultaty bardziej realistyczne.

Przykładowo platforma Hour One¹² pozwala na stworzenie nagrania wideo z wirtualnymi postaciami¹³. Użytkownicy platformy muszą jedynie wybrać jeden spośród dziesiątek stworzonych przez aktorów awatarów, a także jego głos oraz wpisać monit, aby uzyskać realistyczną „gadającą głowę”¹⁴. Hour One pozwala też użytkownikom na digitalizację własnych twarzy i głosów.

Jest to jedna z wielu firm tworzących „wirtualnych ludzi”, które dodały do swoich platform narzędzia językowe oparte na sztucznej inteligencji, aby zapewnić swoim awatarom większy zasięg. Ponad 150 firm tworzy obecnie produkty oparte na generatywnej sztucznej inteligencji – jest to ogólne określenie dla systemów wykorzystujących nienadzorowane uczenie się do tworzenia tekstu i multimediów – dla twórców treści, marketerów i firm medialnych¹⁵.

Tym samym *deep fake* stanowi poważne zagrożenie dla wielu dziedzin życia społecznego oraz gospodarczego, w tym bezpieczeństwa narodowego, demokracji, prywatności czy zdrowia. Dezinformacja, również ta kreowana przez sztuczną inteligencję, została uznana za jeden z rodzajów globalnego ryzyka krótkoterminowego w Raporcie Światowego Forum Ekonomicznego¹⁶.

Deep fake a prawo karne

Przechodząc do prawnego punktu widzenia, można wskazać, że *deep fakes* to informacje niezgodne z prawdą, tym samym ich karalność nastąpi wszędzie tam, gdzie prawo karne penalizuje fałszowanie rzeczywistości¹⁷.

¹¹ N. Young, *Deep Fake Technology: Complete Guide to Deepfakes, Politics and Social Media*, New York 2019, s. 14.

¹² Hour One, https://hourone.ai/ln/ai-training-video-generator-v2/?utm_adgroup=150122319748&utm_keyword=hour%20one&utm_matchtype=b&utm_adid=67412536637939&utm_medium=cpc&utm_campaign=18410771213&utm_term=hour%20one&utm_source=google&utm_medium=cpc&utm_campaign=18410771213&utm_term=hour%20one&utm_source=3566637939&hlsa_cam=18410771213&hlsa_grp=150122319748&hlsa_ad=6741253663858&hlsa_src=g&hlsa_tgt=kwd308720758001&hlsa_kw=hour%20one&hlsa_mt=b&hlsa_net=adwords&utm_ver=3&gad_source=1&gclid=Cj0KCQiA3uGqBhDdARIsAFeJ5r0ealrUYOMA28rGpCY4PkvhaHpCIVQQGgDv7SoJjhn79CtK5Fzj3saAkf8EALw_wcB (21.05.2023).

¹³ A. Pasternack, *GPT-powered deepfakes are a 'powder keg'*, 2023, <https://www.fastcompany.com/90853542/deepfakes-getting-smarter-thanks-to-gpt> (21.05.2023).

¹⁴ *Ibidem*.

¹⁵ *Ibidem*.

¹⁶ The Global Risks Report 2023, World Economic Forum 2023, s. 11 i 24, <https://www.weforum.org/reports/global-risks-report-2023> (21.05.2023).

¹⁷ K. Mamak, *Prawnokarne sposoby walki z fake newsami. RAPORT PROJEKTU SPOŁ-TECH*, 2020, s. 34, file:///C:/Users/ASUS/OneDrive/Desktop/phd/DEEPFAKE/Raport_walka-z-fake-newsami%20MAMAK.pdf (2.06.2023).

Jednym z takich przestępstw jest to opisane w art. 132 Kodeksu karnego¹⁸, które odnosi się do przypadków wprowadzenia w błąd organów państwowych¹⁹. Należy jednak zauważyć, iż dobrem chronionym tym przepisem jest bezpieczeństwo zewnętrzne i wewnętrzne Rzeczypospolitej Polskiej, a także prawidłowe funkcjonowanie jej służb wywiadowczych²⁰. W art. 132 k.k. spenalizowano przestępstwo dezinformacji wywiadowczej stanowiące swoiste lustrzane odbicie szpiegostwa²¹. Tym samym do pociągnięcia do odpowiedzialności za to przestępstwo konieczne jest wypełnienie znamienia czynnościowego, które może być zrealizowane jedynie w ramach „oddawania usług wywiadowczych Rzeczypospolitej Polskiej”, a zatem działania na rzecz polskiego wywiadu lub kontrwywiadu. Z uwagi zatem na ten wyjątkowy charakter tego przepisu (przesłanki działania na rzecz polskiego wywiadu lub kontrwywiadu) należy sądzić, iż będzie on miał zastosowanie niezwykle rzadko.

W dalszej kolejności w kontekście penalizacji zjawiska *deep fake* należałoby rozważyć możliwość zastosowania art. 224a k.k. odnoszącego się do fałszywego alarmu²². Trzeba wskazać, że przestępstwo z art. 224a k.k. ma charakter skutkowy polegający na tym, iż w wyniku jego popełnienia instytucja wskazana w treści przedmiotowego przepisu podejmuje działania mające na celu uchylenie nieistniejącego w rzeczywistości zagrożenia²³. W doktrynie podkreśla się ponadto, że przestępstwo z art. 224a k.k. sprawca może popełnić, wyłącznie działając umyślnie z zamiarem bezpośrednim²⁴. Znamię strony podmiotowej wskazanego przepisu nie zostanie zatem wypełnione w przypadku, gdy osoba zawiadamiająca o nieistniejącym w rzeczywistości zagrożeniu nie ma świadomości jego fik-

¹⁸ Ustawa z dnia 6 czerwca 1997 r. – Kodeks karny (Dz.U. 2024, poz. 17), dalej: k.k.

¹⁹ „Kto, oddając usługi wywiadowcze Rzeczypospolitej Polskiej, wprowadza w błąd polski organ państwowy przez dostarczanie podrobionych lub przerobionych dokumentów lub innych przedmiotów albo przez ukrywanie prawdziwych lub udzielanie fałszywych wiadomości mających istotne znaczenie dla Rzeczypospolitej Polskiej, podlega karze pozbawienia wolności od roku do lat 10”.

²⁰ I. Zgoliński, art. 132 [w:] *Kodeks karny. Komentarz*, red. V. Konarska-Wrzošek, LEX 2023, teza 1.

²¹ K. Lipiński, komentarz do art. 132 [w:] *Kodeks karny. Część szczególna. Komentarz*, red. J. Giezek, Warszawa 2021, teza 1.

²² Zgodnie z treścią § 1: „Kto wiedząc, że zagrożenie nie istnieje, zawiadamia o zdarzeniu, które zagraża życiu lub zdrowiu wielu osób lub mieniu w znacznych rozmiarach lub stwarza sytuację, mającą wywołać przekonanie o istnieniu takiego zagrożenia, czym wywołuje czynność instytucji użyteczności publicznej lub organu ochrony bezpieczeństwa, porządku publicznego lub zdrowia mającą na celu uchylenie zagrożenia, podlega karze pozbawienia wolności od 6 miesięcy do lat 8”.

²³ M. Kulik, komentarz do art. 224a [w:] *Kodeks karny. Komentarz aktualizowany*, red. M. Mozgawa, Warszawa 2020, teza 5.

²⁴ J. Lachowski, komentarz do art. 224a [w:] *Kodeks karny. Część szczególna. Komentarz*, t. II: *Art. 222–316*, red. M. Królikowski, R. Zawłocki, Warszawa 2017, s. 56.

cyjnego charakteru lub działa z zamiarem ewentualnym, tj. przewidując taką możliwość²⁵. W tym miejscu należy także zwrócić uwagę, że możliwość przypisania osobie zawiadamiającej odpowiedzialności karnej z tytułu popełnienia przedmiotowego przestępstwa jest uzależniona od tego, czy jej świadomość na temat nieistnienia zagrożenia wystąpiła w chwili dokonywania przez nią zawiadomienia²⁶.

Następnym przestępstwem, którego możliwość zastosowania należałoby rozważyć pod kątem zjawiska *deep fake*, jest przestępstwo oszustwa z art. 286 § 1 k.k.²⁷ Jest ono skierowane na osobę, którą sprawca zamierza doprowadzić do niekorzystnego rozporządzenia mieniem, i może przybierać trojką postać: wprowadzenia w błąd tej osoby, wyzyskania błędu tej osoby, wyzyskania niezdolności tej osoby do należytego pojmowania przedsiębranego działania. Warto zauważyć, iż w literaturze przedmiotu wskazuje się, że wprowadzenie w błąd może przejawiać się w najrozmaitszych formach²⁸. Może zostać popełnione słowem, pismem, gestem lub w jakikolwiek inny sposób²⁹. Przykładem takiego zachowania może być stworzenie za pomocą technologii głębokiej syntezy wideo ukazującego porwanie dziecka i wysłanie go do rodziców z żądaniem zapłacenia określonej kwoty pieniędzy lub stworzenie fałszywego nagrania, na którym znana osoba wzywa do wpłat na zbiórkę na szczytny cel, który w rzeczywistości nie istnieje.

W tym drugim przypadku samo rozpowszechnianie takiego nagrania może przyczynić się do popełnienia przestępstwa, bowiem zdaniem Sądu Najwyższego wprowadzenie w błąd pokrzywdzonego lub osoby dokonującej rozporządzenia jego mieniem „może być dokonane także za pośrednictwem innych osób, nieświadomych tego, że tworzą u kogoś mylne wyobrażenie o rzeczywistym stanie rzeczy. W takiej sytuacji sprawca nie musi działać na szkodę zindywidualizowanego pokrzywdzonego. Wystarczające jest, że działa z zamiarem doprowadzenia innej, nieznaney mu jeszcze osoby, do niekorzystnego rozporządzenia mieniem w następstwie wprowadzenia jej w błąd, w tym celu, by samemu osiągnąć korzyść majątkową”³⁰. Analizując przepis, należy jednak wspomnieć, iż aby to prze-

²⁵ M. Niedbała, *Odpowiedzialność karna autorów fake newsów wzbudzających w opinii publicznej poczucie zagrożenia w świetle art. 224A Kodeksu Karnego*, „Konteksty Społeczne” 2020, t. VIII, nr 2(16), s. 164.

²⁶ J. Lachowski, komentarz do art. 224a, s. 55–56.

²⁷ „Kto, w celu osiągnięcia korzyści majątkowej, doprowadza inną osobę do niekorzystnego rozporządzenia własnym lub cudzym mieniem za pomocą wprowadzenia jej w błąd albo wyzyskania błędu lub niezdolności do należytego pojmowania przedsiębranego działania, podlega karze pozbawienia wolności od 6 miesięcy do lat 8”.

²⁸ M. Dąbrowska-Kardas, P. Kardas, art. 286 [w:] *Kodeks karny. Część szczególna*, t. III: *Komentarz do art. 278–363 k.k.*, red. W. Wróbel, A. Zoll, Warszawa 2022.

²⁹ *Ibidem*.

³⁰ Wyrok SN z dnia 13 kwietnia 2006 r., sygn. IV KK 40/06, LEX nr 183008.

stępstwo wystąpiło, musi być popełnione w celu osiągnięcia korzyści majątkowej. O ile więc konkretny *deep fake* nie łączy się z uszczerbkiem na majątku indywidualnej ofiary (choćby w formie usiłowania), nie może być mowy o tym przestępstwie.

Kolejnym przestępstwem zasługującym na uwagę jest kradzież tożsamości penalizowana w art. 190a § 2 k.k.³¹ Czyn ten polega na wykorzystaniu przy podszywaniu się pod inną osobę jej wizerunku, innych danych osobowych lub innych danych, za pomocą których jest publicznie identyfikowana, w celu wyrządzenia jej szkody majątkowej lub osobistej. W doktrynie przyjmuje się, iż wizerunek oznacza „dostřegalne, fizyczne cechy człowieka, tworzące jego wygląd i pozwalające na identyfikację osoby wśród innych ludzi”³². W szerokim znaczeniu „w tym pojęciu mieści się nie tylko obraz fizyczny (obraz wizualny), lecz również głos (wizerunek dźwięczny). Są nimi również objęte wszystkie elementy identyfikujące daną jednostkę jako konkretną osobę fizyczną”³³. Problematyczne w kontekście zastosowania tego przepisu okazuje się jednak znamię czynności wykonawczej „podszywania się” przez sprawcę pod inną osobę. Podszywanie się pod inną osobę to wprowadzenie w błąd osób trzecich co do swojej tożsamości, podawanie się za kogoś innego³⁴. Innymi słowy – można również stwierdzić, że podszywanie się to podejmowanie działań jako inna osoba³⁵. W świetle poglądów doktryny można zatem powiedzieć, że podszywanie się jest bezpośrednio powiązane z osobą sprawcy i wymaga niejako „przyjęcia” przez niego cudzej tożsamości. Tym samym w sytuacji np. fałszywej pornografii (*deep porn*), gdzie mamy do czynienia z „połączeniem wizerunków” dwóch różnych osób bez przyjęcia ich przez sprawcę, z uwagi na brak wypełnienia znamienia podszywania się nie można zgodzić się z propozycją pociągnięcia do odpowiedzialności karnej sprawcy na podstawie art. 190a § 2 k.k.³⁶

³¹ Art. 190a § 1 k.k.: „Kto przez uporczywe nękanie innej osoby lub osoby dla niej najbliższej wzbudza u niej uzasadnione okolicznościami poczucie zagrożenia, poniżenia lub udręczenia lub istotnie narusza jej prywatność, podlega karze pozbawienia wolności od 6 miesięcy do lat 8. § 2. Tej samej karze podlega, kto, podszywając się pod inną osobę, wykorzystuje jej wizerunek, inne jej dane osobowe lub inne dane, za pomocą których jest ona publicznie identyfikowana, przez co wyrządza jej szkodę majątkową lub osobistą”.

³² E. Wojnicka, *Prawo do wizerunku w ustawodawstwie polskim*, „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Wynalazczości i Ochrony Własności Intelaktualnej” 1990, z. 56, s. 107.

³³ J. Sieńczyło-Chlabicz, *Przedmiot, podmiot i charakter prawa do wizerunku*, „Przegląd Ustawodawstwa Gospodarczego” 2003, nr 8, s. 20.

³⁴ J. Lachowski, art. 190a [w:] *Kodeks karny. Komentarz*, red. V. Konarska-Wrżosek, LEX 2020, teza 9.

³⁵ A. Zoll, art. 190a [w:] *Kodeks karny. Część szczególna*, t. II, cz. 1: *Komentarz do art. 117–211a*, red. W. Wróbel, A. Zoll, LEX 2017, teza 9.

³⁶ A. Ziobroń, *Deepfake a prawo karne. Uwagi de lege lata i de lege ferenda dotyczące fałszywej pornografii*, „Studenckie Prace Prawnicze, Administratywistyczne i Ekonomiczne” 2021, nr 37, s. 229.

Istotnymi przestępstwami z punktu widzenia analizy zjawiska *deep fake* i możliwości jego karania są: zniesławienie i znieważenie. Należy przyjąć, iż przestępstwo zniesławienia³⁷ wobec braku normatywnych ograniczeń może być popełnione w każdej formie umożliwiającej realizację przekazu informacyjnego drugiej osobie³⁸. Nie chodzi jedynie o zarzut wyrażony w formie ustnej, lecz także przy wykorzystaniu rozmaitych środków przekazu informacji, np. wizerunku, utworu artystycznego³⁹. Przestępstwo opisane w art. 212 k.k. może obejmować więc różne rodzaje *deep fake*'ów, dotyczących nie tylko osób fizycznych, ale także instytucji. *Deep fake* w formie zniesławienia to informacja, która przypisuje ofierze negatywnie oceniane właściwości czy postępowanie, co może ją narazić na poniżenie w opinii publicznej czy utratę zaufania⁴⁰. Przedstawiając konkretną osobę czy instytucję (np. fundację) w złym świetle, *deep fake* może istotnie przyczynić się do zniszczenia jej reputacji i renomy oraz wpłynąć negatywnie na jej funkcjonowanie (w przypadku instytucji) i odbiór społeczny.

Od omówionego powyżej przestępstwa zniesławienia należy odróżnić przestępstwo zniewagi⁴¹, które odnosi się wyłącznie do osoby fizycznej i godzi w wewnętrzną część osoby znieważonej⁴². Przykładem zachowania, które mogłoby wypełnić znamiona omawianego przepisu, jest stworzenie filmu wideo, na którym wykreowana postać z twarzą prawdziwej osoby zachowuje się w sposób obiektywnie uznany społecznie za niewłaściwy. Takim filmem może być też produkcja pornograficzna z twarzą określonej osoby (*deep porn*), w którym zaangażowana jest ona w akt, który w oczach opinii publicznej odebrany będzie jako poniżający.

Omawiając przestępstwo zniewagi, należy też pamiętać, że co do pewnych osób k.k. przewiduje osobną kategorię przestępstw, np. zniewaga prezydenta RP (art. 135 § 2 k.k.)⁴³. Osobnym przestępstwem jest również znieważenie narodu

³⁷ Art. 212 § 1: „Kto pomawia inną osobę, grupę osób, instytucję, osobę prawną lub jednostkę organizacyjną niemającą osobowości prawnej o takie postępowanie lub właściwości, które mogą poniżyć ją w opinii publicznej lub narazić na utratę zaufania potrzebnego dla danego stanowiska, zawodu lub rodzaju działalności, podlega grzywnie albo karze ograniczenia wolności. § 2. Jeżeli sprawca dopuszcza się czynu określonego w § 1 za pomocą środków masowego komunikowania, podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do roku”.

³⁸ J. Raglewski [w:] *Kodeks karny. Część szczególna*, t. II, cz. 2: *Komentarz do art. 212–277d*, red. W. Wróbel, A. Zoll, Warszawa 2017.

³⁹ *Ibidem*.

⁴⁰ K. Mamak, *Prawnokarne sposoby...*, s. 22.

⁴¹ Art. 216 § 1 k.k.: „Kto znieważa inną osobę w jej obecności albo choćby pod jej nieobecność, lecz publicznie lub w zamiarze, aby zniewaga do osoby tej dotarła, podlega grzywnie albo karze ograniczenia wolności”.

⁴² P. Rodak, *Przestępstwo zniesławienia a znieważenia – podobieństwa i różnice*, 2021, <https://www.adwokat-rodak.pl/znieslawienie-zniewaga/> (18.11.2023).

⁴³ Art. 135 § 2 k.k.: „Kto publicznie znieważa Prezydenta Rzeczypospolitej Polskiej, podlega karze pozbawienia wolności do lat 3”.

lub Rzeczypospolitej Polskiej⁴⁴, a także grupy ludności albo poszczególnej osoby z powodu jej przynależności narodowej, etnicznej, rasowej, wyznaniowej albo z powodu jej bezwyznaniowości⁴⁵.

Omówione powyżej przepisy stanowią jedynie kilka wybranych przykładów przestępstw, które mogłyby mieć zastosowanie w przypadku wypełniania ich znamion z użyciem technologii *deep fake*. Jednak już na tym etapie analizy można zauważyć, iż ochrona prawna osób pokrzywdzonych przez zachowania sprawców z użyciem nowych technologii jest jedynie fragmentaryczna i ogranicza się tylko do ściśle określonych sytuacji.

Pierwsze regulacje prawne technologii *deep fake*

Polski system prawny nie jest jednak jedynym, który może mieć trudności w kwalifikacji przestępstw popełnionych z użyciem środków nowej generacji. Rosnąca liczba *deep fakes* to problem globalny, z którym mierzy się obecnie cały świat. Przykładem kraju, który radykalnie uregulował prawnie kwestie *deep fake*, są Chiny. W styczniu 2023 r. chińska administracja cyberprzestrzeni opublikowała nowe przepisy, których celem jest uregulowanie treści generowanych przez sztuczną inteligencję, w tym – *deep fake*'ów⁴⁶. W art. 6 Przepisów dotyczących administrowania internetowymi usługami informacyjnymi *Deep Synthesis* wskazano, iż usługi głębokiej syntezy nie mogą być wykorzystywane przez żadną organizację lub osobę do tworzenia, powielania, publikowania lub przesyłania informacji zabronionych przez prawo lub przepisy administracyjne lub do angażowania się w działania zabronione przez prawo i przepisy administracyjne, takie jak te, które zagrażają bezpieczeństwu narodowemu i interesom narodowym, szkodzą wizerunkowi narodu i społecznemu interesowi publicznemu, zakłócają porządek gospodarczy lub społeczny lub szkodzą prawowitym prawom i interesom innych osób⁴⁷. Dostawcy usług głębokiej syntezy i użytkownicy nie mogą korzystać z usług głębokiej syntezy do tworzenia, powielania, publikowa-

⁴⁴ Art. 133 k.k.: „Kto publicznie znieważa Naród lub Rzeczpospolitą Polską, podlega karze pozbawienia wolności do lat 3”.

⁴⁵ Art. 257 k.k.: „Kto publicznie znieważa grupę ludności albo poszczególłą osobę z powodu jej przynależności narodowej, etnicznej, rasowej, wyznaniowej albo z powodu jej bezwyznaniowości lub z takich powodów narusza nietykalność cielesną innej osoby, podlega karze pozbawienia wolności do lat 3”.

⁴⁶ M. Fraser, *Chińskie regulacje deepfake. Przykład dla świata czy nadgorliwość?*, 2023, <https://cyberdefence24.pl/cybermagazyn/cybermagazyn-chinskie-regulacje-deepfake-przyklad-dla-swiatea-czy-nadgorliwosc> (2.06.2023).

⁴⁷ CHINA LAW TRANSLATE, Provisions on the Administration of Deep Synthesis Internet Information Services, 2022, <https://www.chinalawtranslate.com/en/deep-synthesis> (2.06.2023).

nia lub przesyłania fałszywych informacji⁴⁸. Dostawcy głębokiej syntezy muszą też zapewnić dodawanie etykiet do treści generowanych przy użyciu technologii głębokiej syntezy oraz przeprowadzać przeglądy swoich algorytmów i procedur bezpieczeństwa⁴⁹.

Podobne prawo w ubiegłym roku wprowadzone zostało w Malezji, gdzie za szerzenie nieprawdziwych informacji groziło nawet 6 lat więzienia⁵⁰. Jednak władze po kilku miesiącach zdecydowały się na wycofanie tych nowych przepisów⁵¹. W 2018 r. malezyjskie społeczeństwo obywatelskie i międzynarodowe organizacje praw człowieka wyraziły obawy, że ustawa jest niezgodna z międzynarodowym prawem dotyczącym praw człowieka i zostanie wykorzystana do cenzurowania wypowiedzi i karania krytyków rządu⁵². Krajem, w którym obecnie obowiązują takie przepisy, jest Singapur⁵³. Tam osoby, które tworzą lub udostępniają nieprawdziwe informacje, muszą się liczyć z karami dochodzącymi nawet do 720 tys. dolarów grzywny lub 10 lat pozbawienia wolności⁵⁴.

Również Unia Europejska podejmuje próby uregulowania kwestii szerzącego się zjawiska *deep fake* w projekcie rozporządzenia Parlamentu Europejskiego i Rady ustanawiającego zharmonizowane przepisy dotyczące sztucznej inteligencji⁵⁵. Przyjęte 14 czerwca 2023 r. przez Parlament Europejski poprawki do tego projektu, konkretnie poprawka 203, wprowadzają definicję *deep fake*, zgodnie z którą oznacza on „zmanipulowane lub syntetyczne treści audio, obrazy lub wideo, które fałszywie wydają się autentyczne lub prawdziwe i które zawierają wizerunki osób, które wydają się mówić lub robić rzeczy, których nie powiedziały ani nie zrobiły, wyprodukowane przy użyciu technik sztucznej inteligencji, w tym uczenia maszynowego i głębokiego uczenia”⁵⁶.

⁴⁸ *Ibidem*.

⁴⁹ *Ibidem*.

⁵⁰ W. Kulik, *Więzenie za deepfake'i i fake newsy – czy w Polsce też tak powinno być?*, 2019, <https://www.benchmark.pl/aktualnosci/wiezienie-za-deepfake-fake-newsy-czy-w-polsce-tez-tak-powinno-byc.html> (2.06.2023).

⁵¹ *Ibidem*. W 2021 r. w Malezji ponownie wprowadzono przepisy określające szereg przestępstw związanych z „fałszywymi wiadomościami”. Szerzej na ten temat por. m.in. L.Schuldt, *The rebirth of Malaysia's fake news law – and what the NetzDG has to do with it*, 2021, <https://verfassungsblog.de/malaysia-fake-news/> (2.06.2023).

⁵² ARTICLE 19, *Malaysia: Repeal 'fake news' emergency ordinance*, 2021, <https://www.article19.org/resources/malaysia-fake-news-ordinance/> (2.06.2023).

⁵³ W. Kulik, *Więzenie za deepfake'a...*

⁵⁴ *Ibidem*.

⁵⁵ Projekt rozporządzenia Parlamentu Europejskiego i Rady ustanawiającego zharmonizowane przepisy dotyczące sztucznej inteligencji (akt w sprawie sztucznej inteligencji) Bruksela, 21.04.2021, COM(2021) 206 *final*, 2021/0106(COD), dalej: projekt rozporządzenia.

⁵⁶ Thum. własne z języka angielskiego: „*deep fake*» means manipulated or synthetic audio, image or video content that would falsely appear to be authentic or truthful, and which features depictions of persons appearing to say or do things they did not say or do, produced using AI

Ponadto zgodnie z art. 52 § 3 projektu rozporządzenia użytkownicy systemu sztucznej inteligencji *deep fake* są zobowiązani do ujawniania w odpowiedni, terminowy, jasny i widoczny sposób, że treść została sztucznie wygenerowana lub zmanipulowana, a także w miarę możliwości imię i nazwisko osoby fizycznej lub prawnej, która ją wygenerowała lub zmanipulowała. Ujawnianie oznacza umieszczenie w takim wytworze typu *deep fake* wyraźnej i widocznej dla odbiorcy informacji, że treść jest nieautentyczna⁵⁷. Powyższe nie ma zastosowania, jeżeli korzystanie z systemu sztucznej inteligencji *deep fake* jest dozwolone przez prawo lub jeżeli jest to konieczne do korzystania z prawa do wolności wypowiedzi oraz prawa do wolności sztuki i nauk ścisłych gwarantowanych w Karcie praw podstawowych Unii Europejskiej i pod warunkiem odpowiednich gwarancji praw i wolności osób trzecich. Jeżeli wytwór typu *deep fake* stanowi część wyraźnie twórczego, satyrycznego, artystycznego lub fikcyjnego filmu kinematograficznego, materiałów wizualnych do gier wideo lub analogicznego utworu lub programu, obowiązki w zakresie przejrzystości dotyczące stosowania wyraźnego oznaczenia, iż odbiorca ma do czynienia z wytworem typu *deep fake*, również mają zastosowanie, jednak oznaczenie to powinno być umieszczone w sposób, który nie utrudnia wyświetlania utworu i ujawniania obowiązujących praw autorskich, tam gdzie ma to zastosowanie. Nie uniemożliwia to także organom egzekwowania prawa korzystania z systemów sztucznej inteligencji przeznaczonych do wykrywania głębokich podróbek oraz zapobiegania przestępstwom związanym z ich stosowaniem, prowadzenia dochodzeń w ich sprawie i ścigania ich⁵⁸.

Należy też wspomnieć, iż 16 czerwca 2022 r. Komisja Europejska przedstawiła zmieniony Kodeks postępowania w zakresie zwalczania dezinformacji⁵⁹. Kodeks wraz z aktem o usługach cyfrowych tworzy unijny zestaw narzędzi do przeciwdziałania dezinformacji w internecie⁶⁰. Przystąpienie do Kodeksu jest dobrowolne, a sam Kodeks określa mechanizm dla podmiotów do walki z dezinformacją w internecie⁶¹. Obecnie do przestrzegania Kodeksu zobowiązały się

techniques, including machine learning and deep learning”. Parlament Europejski, *Amendments adopted by the European Parliament on 14 June 2023 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD))*, https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html (2.06.2023).

⁵⁷ *Ibidem*.

⁵⁸ *Ibidem*.

⁵⁹ Komisja Europejska, Kodeks postępowania w zakresie dezinformacji z 2022 r., <https://digital-strategy.ec.europa.eu/pl/policies/code-practice-disinformation> (2.06.2023).

⁶⁰ *Ibidem*.

⁶¹ P. Zegarow, *Nowy kodeks w zakresie zwalczania dezinformacji*, 2022, <https://cyberpolicy.nask.pl/nowy-kodeks-postepowania-w-zakresie-zwalczania-dezinformacji/> (2.06.2023).

34 firmy, m.in. Adobe, Meta, Microsoft, Seznam, TikTok, Twitter (obecnie Platforma X)⁶². Kodeks zawiera 44 zobowiązania i 128 konkretnych środków⁶³ w następujących obszarach: demonetyzacja; ograniczenie zachęt finansowych dla dostawców dezinformacji; przejrzystość reklam politycznych; integralność usług; rola użytkowników i naukowców; rola społeczności weryfikującej fakty⁶⁴.

Wnioski *de lege lata* oraz postulaty *de lege ferenda*

Mając powyższe na uwadze, należy stwierdzić, iż problemy związane z rosnącą liczbą nowych zachowań, a potencjalnie przestępstw popełnianych przy użyciu systemów nowoczesnych technologii typu *deep fake* rośnie, a obecnie obowiązujące prawo nie zawsze jest w stanie zapewnić obywatelom odpowiednią ochronę prawną. Tym samym właściwe byłoby rozpoczęcie dyskusji dotyczącej uaktualnienia przepisów również polskiego k.k. lub dodania nowych. Wprowadzenie rozwiązań prawnokarnych może być jednak trudne, gdyż z uwagi na wielość rodzajów *deep fake*ów oraz ciągły dynamiczny rozwój technologiczny nie ma powszechnie uznanej definicji tego pojęcia. Ponadto zmiana prawa nie powinna polegać na wprowadzeniu karalności wszystkich *deep fake*ów, ponieważ regulacja ograniczałaby wolność słowa, a zgodnie z art. 54 Konstytucji RP⁶⁵ „każdemu zapewnia się wolność wyrażania swoich poglądów oraz pozyskiwania i rozpowszechniania informacji”. Tym samym w doktrynie słusznie wskazuje się na możliwość wprowadzenia zakazu jedynie wybranych rodzajów *deep fake*⁶⁶. Wobec tego można zaproponować wprowadzenie do k.k. przepisu penalizującego zachowanie polegające na tworzeniu i rozpowszechnianiu treści typu *deep fake* mogących wywołać niepokój publiczny. Z uwagi na dobro prawnie chronione, jakim jest porządek publiczny, stosowne byłoby umieszczenie takiego przepisu np. jako art. 255b w rozdziale XXXII k.k. Celem wprowadzenia takiego przepisu byłaby penalizacja tworzenia i rozpowszechniania treści, które mogłyby doprowadzić do zachwiania zaufania członków społeczeństwa do niezakłóconego stałego porządku, rozbudzenia niepokoju wewnętrznego wielu osób, wywołania u nich poczucia zagrożenia, które w rzeczywistości nie istnieje. Należy w tym miejscu zauważyć, iż podobny przepis funkcjonował już jako art. 170 rozporządzenia Prezydenta Rzeczypospolitej z dnia 11 lipca 1932 r. – Kodeks

⁶² *Ibidem.*

⁶³ *Ibidem.*

⁶⁴ Komisja Europejska, Kodeks postępowania w zakresie dezinformacji z 2022 r.

⁶⁵ Konstytucja Rzeczypospolitej Polskiej z dnia 2 kwietnia 1997 r. (Dz.U. 1997, nr 78, poz. 483 ze zm.).

⁶⁶ K. Mamak, *Prawnokarne sposoby...*

karny⁶⁷ i miał na celu penalizację rozpowszechniania fałszywych wiadomości mogących wywołać niepokój publiczny⁶⁸.

Innym godnym rozważenia rozwiązaniem byłoby uregulowanie sposobu tworzenia i rozpowszechniania treści typu *deep fake* wraz z jednoczesnym przyjęciem definicji w osobnym akcie administracyjnym. Akt ten zawierałby też przepisy przewidujące odpowiedzialność karną za brak przestrzegania zasad tworzenia wyżej wspomnianych treści.

Zdaniem autorki, mając jednak na uwadze nieuniknioną wejścia w życie w najbliższym czasie projektu rozporządzenia⁶⁹, który wprowadza zarówno definicję, jak i reguluje zjawisko *deep fake* i tym samym obowiązku stosowania tego aktu w całości na całym obszarze Unii Europejskiej, należy się opowiedzieć za pierwszym z zaproponowanych rozwiązań.

Podsumowując, jak wynika z przytoczonych rozważań, technologia *deep fake* rozwija się bardzo dynamicznie, znacząco przyczyniając się do szerzenia niebezpiecznej dezinformacji – tak państwowej, jak i światowej. Świat, dostrzegając siłę technologii *deep fake*, która może zostać użyta w celach destabilizacyjnych, podjął już działania w celu uregulowania zasad zmierzających do ograniczenia oraz przeciwdziałania skutkom tej technologii i je kontynuuje.

Zaprezentowane rozważania dotyczące polskiego stanu prawnego wskazują na niewystarczające regulacje i istnienie luki prawnej wynikającej z postępu technologicznego, którego polski ustawodawca nie przewidział. Choć jak wykazano w ramach wniosków *de lege lata*, *deep fake* może podlegać penalizacji na gruncie przytoczonych polskich przepisów, to jednak ochrona ta jest fragmentaryczna i niewystarczająca. W celu jej uporządkowania warto skorzystać ze wskazanych wyżej doświadczeń zagranicznych unormowań prawnych, co można traktować jako postulaty *de lege ferenda* w celu rozwiązania problemu ogólnoswiatowego, za jaki należy uznać zjawisko *deep fake*.

Bibliografia

ARTICLE 19, *Malaysia: Repeal 'fake news' emergency ordinance*, 2021, <https://www-article19.org/resources/malaysia-fake-news-ordinance/> (2.06.2023).

⁶⁷ Rozporządzenie Prezydenta Rzeczypospolitej z dnia 11 lipca 1932 r. – Kodeks karny (Dz.U. 1932, nr 60, poz. 571).

⁶⁸ M. Niedbała, *Odpowiedzialność karna...*, s. 174.

⁶⁹ 9 grudnia 2023 r., po trzech dniach rozmów, Rada i Parlament Unii Europejskiej osiągnęły porozumienie w sprawie projektu rozporządzenia. Uzgodniony tekst będzie teraz musiał zostać formalnie przyjęty zarówno przez Parlament, jak i Radę, aby stał się prawem Unii. Zob. szerzej: Rada UE, *Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world*, <https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/> (5.01.2024).

- Campbell C., Kirk K., Sands S., Kietzmann J., *Preparing for an Era of Deepfakes and AI-Generated Ads: A Framework for Understanding Responses to Manipulated Advertising*, „Journal of Advertising” 2021, vol. 51.
- CHINA LAW TRANSLATE, Provisions on the Administration of Deep Synthesis Internet Information Services, 2022, <https://www.chinalawtranslate.com/en/deep-synthesis> (2.06.2023).
- Dąbrowska-Kardas M., Kardas P. [w:] *Kodeks karny. Część szczególna*, t. III: *Komentarz do art. 278–363 k.k.*, red. W. Wróbel, A. Zoll, Warszawa 2022.
- Fraser M., *Chińskie regulacje deepfake. Przykład dla świata czy nadgorliwość?*, 2023, <https://cyberdefence24.pl/cybermagazyn/cybermagazyn-chinskie-regulacje-deep-fake-przyklad-dla-swiata-czy-nadgorliwosc> (2.06.2023).
- Grand View Research, 2022, <https://www.grandviewresearch.com/industry/technology> (2.06.2023).
- Hour One*, https://hourone.ai/Ind/ai-training-video-generator-v2/?utm_adgroup=150122319748&utm_keyword=hour%20one&utm_matchtype=b&utm_adid=674125363858&utm_source=google&utm_medium=cpc&utm_campaign=18410771213&utm_term=hour%20one&hsa_acc=3566637939&hsa_cam=18410771213&hsa_grp=150122319748&hsa_ad=674125363858&hsa_src=g&hsa_tgt=kwd308720758001&hsa_kw=hour%20one&hsa_mt=b&hsa_net=adwords&hsa_ver=3&gad_source=1&gclid=Cj0KCQiA3uGqBhDdARIsAFeJ5r0ealrUYOMA28rGpCY4PkvhaHpCCIVQQGgDv7SoJhn79CtK5Fzj3saAkf8EALw_wcB (21. 05.2023).
- Kietzmann J., Lee L.W., McCarthy I.P., Kietzmann T.C., *Deepfakes: Trick or treat?*, „Business Horizons” 2020, vol. 63, no. 2.
- Kulik M., komentarz do art. 224a [w:] *Kodeks karny. Komentarz aktualizowany*, red. M. Mozgawa, Warszawa 2020.
- Kulik W., *Więzienie za deepfake'i i fake newsy – czy w Polsce też tak powinno być?*, 2019, <https://www.benchmark.pl/aktualnosci/wiezienie-za-deepfake-fake-newsy-czy-w-polsce-tez-tak-powinno-byc.html> (2.06.2023).
- Lachowski J., art. 190a [w:] *Kodeks karny. Komentarz*, red. V. Konarska-Wrzošek, LEX 2020.
- Lachowski J., komentarz do art. 224a [w:] *Kodeks karny. Część szczególna. Komentarz*, t. II: *Art. 222–316*, red. M. Królikowski, R. Zawłocki, Warszawa 2017.
- Lipiński K., komentarz do art. 132 [w:] *Kodeks karny. Część szczególna. Komentarz*, red. J. Giezek, Warszawa 2021.
- Mamak K., *Prawnokarne sposoby walki z fake newsami. RAPORT PROJEKTU SPOLTECH*, 2020, file:///C:/Users/ASUS/OneDrive/Desktop/phd/DEEPFAKE/Raport_walka-z-fake-newsami%20MAMAK.pdf (2.06.2023).
- Niedbała M., *Odpowiedzialność karna autorów fake newsów wzbudzających w opinii publicznej poczucie zagrożenia w świetle art. 224A Kodeksu Karnego*, „Konteksty Społeczne” 2020, t. VIII, nr 2 (16).
- Paschen U., Pitt Ch., Kietzmann J., *Artificial intelligence: Building blocks and an innovation typology*, „Business Horizons” 2020, vol. 63, no. 2.
- Pasternack A., *GPT-powered deepfakes are a 'powder keg'*, 2023, <https://www.fastcompany.com/90853542/deepfakes-getting-smarter-thanks-to-gpt> (21.05.2023).
- Rachita R., Himanshu J., Vineet K., *Conversational AI Market Outlook – 2030*, <https://www.alliedmarketresearch.com/conversational-ai-market-A13682> (21.05. 2023).
- Rada UE, *Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world*, <https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/> (5.01.2024).

- Raglewski J. [w:] *Kodeks karny. Część szczególna*, t. II, cz. II: *Komentarz do art. 212–277d*, red. W. Wróbel, A. Zoll, Warszawa 2017.
- Schuldts L., *The rebirth of Malaysia's fake news law – and what the NetzDG has to do with it*, 2021, <https://verfassungsblog.de/malaysia-fake-news/> (2.06.2023).
- Schwab K., *The Fourth Industrial Revolution. World Economic Forum 91–93 Route de la Capite CH-1223 Cologny; World Economic Forum: Geneva, Switzerland* 2016, www.weforum.org (21.05.2023).
- Sentinel, *DEEPFAKES 2020: THE TIPPING POINT The Current Threat Landscape, its Impact on the U.S 2020 Elections, and the Coming of AI-Generated Events at Scale*, 2020, <https://thesentinel.ai/media/Deepfakes%202020:%20The%20Tipping%20Point,%20Sentinel.pdf> (2.06.2023).
- Sieńczyło-Chlabicz J., *Przedmiot, podmiot i charakter prawa do wizerunku*, „Przegląd Ustawodawstwa Gospodarczego” 2003, nr 8.
- The Global Risks Report 2023*, World Economic Forum 2023, <https://www.weforum.org/reports/global-risks-report-2023> (21.05.2023).
- Whittaker L., Mulcahy R., Letheren K., Kietzmann J., Russell-Bennett R., *Mapping the deepfake landscape for innovation: A multidisciplinary systematic review and future research agenda*, „Technovation” 2023, vol. 125, <https://www.sciencedirect.com/science/article/abs/pii/S0166497223000950#preview-section-references> (21.05.2023).
- Wojnicka E., *Prawo do wizerunku w ustawodawstwie polskim*, „Zeszyty Naukowe Uniwersytetu Jagiellońskiego. Prace z Wynalazczości i Ochrony Własności Intelektualnej” 1990, z. 56.
- Zegarow P., *Nowy kodeks w zakresie zwalczania dezinformacji*, 2022, <https://cyberpolicynask.pl/nowy-kodeks-postepowania-w-zakresie-zwalczania-dezinformacji/> (2.06.2023).
- Zgoliński I., art. 132 [w:] *Kodeks karny. Komentarz*, red. V. Konarska-Wrzošek, LEX 2023.
- Ziobroń A., *Deepfake a prawo karne. Uwagi de lege lata i de lege ferenda dotyczące fałszywej pornografii*, „Studenckie Prace Prawnicze, Administratywistyczne i Ekonomiczne” 2021, nr 37.
- Zoll A., art. 190a [w:] *Kodeks karny. Część szczególna*, t. II, cz. 1: *Komentarz do art. 117– 211a*, red. W. Wróbel, A. Zoll, LEX 2017.
- Young W., *DeepFake Technology: Complete Guide to Deepfakes, Politics and Social Media*, New York 2019.

Streszczenie

Pojęcie *deep fake* odnosi się do technologii, która dzięki użyciu sztucznej inteligencji generuje obrazy, filmy i nagrania symulujące rzeczywistość. Postęp techniczny i technologiczny w ostatnich latach sprawił, że tworzenie *deep fake* jest jeszcze łatwiejsze i szybsze do wykonania, a otrzymane rezultaty bardziej realistyczne. Tym samym *deep fake* stanowi poważne zagrożenie dla wielu dziedzin życia społecznego oraz gospodarczego. Mając powyższe na uwadze, tematem artykułu jest przybliżenie zagadnienia *deep faków* z perspektywy polskiego prawa karnego. Wskazane zostały zagrożenia wynikających z ich funkcjonowania oraz problemy prawnekarne.

Wynikiem przeprowadzonych rozważań jest wskazanie, że w obecnym stanie prawnym *deep fake* może podlegać penalizacji na gruncie prawa karnego, jednak ochrona ta jest jedynie fragmentaryczna. Ponadto wprowadzenie nowych regulacji prawnych jest też kwestią problematyczną, ponieważ zakaz wszystkich *deep fake’ów* wkraczałoby w zakres konstytucyjnie chronionej wolności wypowiedzi.

Słowa kluczowe: *deep fake*, polskie prawo karne, odpowiedzialność karna, międzynarodowe prawo karne, sztuczna inteligencja, prawo nowych technologii

DEEP FAKE – TECHNOLOGICAL PROGRESS AND CRIMINAL LAW

Summary

The concept of deep fake refers to technology that, thanks to the use of artificial intelligence, generates images, videos and recordings that simulate reality. Technical and technological progress in recent years has made creating deep fakes even easier and faster, and the results obtained are more realistic. Therefore, deep fake poses a serious threat to many areas of social and economic life. With the above in mind, the topic of the article is to present the issue of deep fakes from the perspective of Polish criminal law. The threats resulting from their operation and criminal law problems will be indicated.

The result of the considerations carried out is to indicate that, in the current legal situation, deep fakes may be subject to penalization under criminal law, but this protection is only fragmentary. Moreover, the introduction of new legal regulations is also a problematic issue, because the introduction of regulations banning all deep fakes would violate the scope of constitutionally protected freedom of expression.

Keywords: deep fake, Polish criminal law, criminal liability, international criminal law, artificial intelligence, new technology law