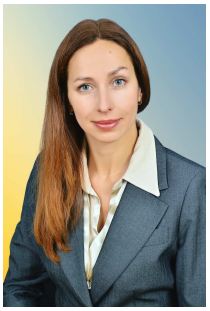
УДК 343.34+ 343.4+343.85



Ксенія Володимирівна ЮРТАЄВА,
кандидат юридичних наук, доцент
(Харківський національний університет внутрішніх справ)

КРИМІНОЛОГІЧНИЙ АНАЛІЗ ВИКОРИСТАННЯ ТЕХНОЛОГІЇ DEEPFAKE: КОЛИ ФЕЙК СТАЄ ЗЛОЧИНОМ

Постановка проблеми. У XXI столітті інформаційні процеси, що ґрунтуються на комп'ютерній обробці великих обсягів даних, набули визначального значення в усіх сферах суспільного життя. Сучасні телекомунікації, які забезпечили безперервний глобальний інформаційний обмін, назавжди змінили ринок сучасних професій і послуг, стали основним джерелом отримання інформації про суспільні події та навколишній світ, виступають центральною платформою міжособистого спілкування, розваг, захоплень, творчого зростання тощо. Відповідно, цифрова інформація, яку людина отримує через телекомунікаційні мережі, є життєво важливим, ессенціальним медіумом та основою розвитку й нормального функціонування сучасного суспільства. Саме тому несанкціоноване втручання в процес маршрутизації інформації, яка передається через телекомунікаційні мережі, спотворення процесу її обробки, а також розповсюдження неправдивої, підробленої інформації (фейків) порушують нормальне функціонування суспільних відносин у різних сферах суспільного життя і виступають загрозою для інформаційної безпеки держави. У зв'язку з вищевикладеним видається актуальним кримінологічне дослідження нової цифрової технології – Deepfake, призначеної для створення високоякісних фото- і відеопідробок, визначення прикладних сфер її застосування, а також пов'язаних із цим потенційних і реальних загроз для інформаційної безпеки держави.

Метою статті є кримінологічний аналіз використання технології Deepfake: визначення правомірних засад і сфер її застосування та окреслення можливостей її використання у вчиненні кримінальних правопорушень.

Аналіз останніх досліджень і публікацій. Наразі питання, пов'язані з оцінюванням криміногенного потенціалу використання технології Deepfake, є недостатньо дослідженими у вітчизняній науковій літературі. Насамперед це пов'язано із доволі незначним часом існування зазначеної комп'ютерної технології. Проте разом із популяризацією діпфейків зазначене питання закономірно отримує дедалі більшу увагу вітчизняних

науковців. Деякі правові аспекти використання технології Deepfake досліджено в роботах О. М. Головко, С. Ф. Денисова, М. В. Дубняк, Т. О. Ісакової, Ю. В. Філей та інших науковців. Серед зарубіжних дослідників слід відзначити праці Н. І. Браун (N. I. Brown), А. В. Вальорскої (A. M. Walorska), Д. В. Мун, В. В. Попети, Н. Ф. Красовської, Б. Дж. Сікірські (B. J. Siekierski) та інших. Водночас постійно зростаюча активність використання технології Deepfake, розширення сфер її застосування та стрімке збільшення протиправних випадків дезінформації, здійснених із застосуванням діпфейків, вимагають комплексного дослідження криміногенних ризиків застосування вказаної цифрової технології.

Виклад основного матеріалу. Технологія Deepfake є абсолютно новим інструментом створення віртуального контенту. Коріння технології Deepfake уходять у 90-ті роки XX ст., проте тоді такими інструментами володіли лише експерти зі спецефектів у кіноіндустрії. Прорив у технології зі створення реалістичних відео відбувся у 2014 р. завдяки алгоритму машинного навчання під назвою «генеративна змагальна мережа» (від англ. *Generative adversarial network*, скорочено GAN), розробленої студентом Стенфордського університету Яном Ґудфеллоу (Ian Goodfellow) та його товаришами. Алгоритм GAN ґрунтується на комбінації двох нейронних мереж, одна з яких (генеративна мережа G) генерує зразки, а інша (дискримінативна мережа D) у процесі антагоністичної гри намагається відрізнити справжні зразки від підроблених. Особливістю функціонування GAN є те, що завдяки постійному вдосконаленню в процесі антагоністичного змагання зазначений алгоритм машинного навчання постійно самовдосконалюється і надає можливість створити реалістичні, проте повністю комп'ютерно генеровані фотореаліастичні зображення інтер'єру, різних предметів і людських облич[1].

Саме розроблений Я. Ґудфеллоу алгоритм GAN став основою створення технології Deepfake, яку ще називають технологією зміни обличчя (*face swapping*): алгоритм аналізує наявні фото- і відеодані конкретної людини, визначає, як ця людина могла б рухатися, як говорити, а потом генерує новий контент із цим персонажем. Сам термін «Deepfake» виник приблизно в 2017 р. і поєднав у собі два поняття – «*deep learning*» (глибинне навчання – вид машинного навчання, який використовує штучні нейронні мережі для обробки даних) та «*fake*» (підробка). Поняття Deepfake почало використовуватися для позначення технології виготовлення гіперреалістичних підроблених фото- і відеозображень, створених на основі алгоритму глибинного навчання GAN.

Серед найбільш розповсюджених способів створення діпфейків можна виділити такі: 1) заміну обличчя: обличчя однієї людини у відеоматеріалі замінюється на обличчя іншої людини; 2) синхронізацію губ: рухи губ людини на оригінальному відео узгоджуються з неоригінальним

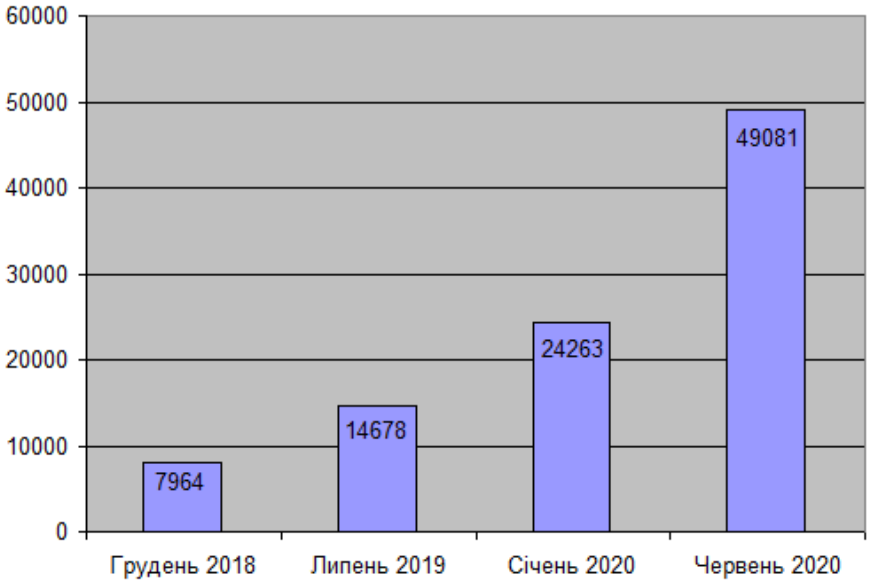
[1] Liu F. The Math Behind Deepfakes. An Overview of GANs. URL: https://towardsdatascience.com/the-math-behind-deepfakes-b9ef8621541f (дата звернення: 01.03.2021).

аудіозаписом; 3) «ляльковод»: відеозображення людини (рухи голови, рухи очима, міміка) анімує виконавець, що сидить перед камерою і виконує дії, які він хоче бачити на відео від створеного персонажу[1]. Отже, застосування технології Deepfake надає безмежні можливості для створення відеоконтенту, в якому можна «примусити» будь-яку особу сказати або зробити будь-що. У зв'язку з вищевикладеним відразу постає декілька очевидних запитань: якими є практичні сфери застосування технології Deepfake і наскільки легальним є її застосування?

Цифрові технології, подібні до Deepfake, відомі протягом декількох десятиріч і використовувалися, зокрема, для зміни зовнішності акторів у кіноіндустрії. Так, найпершим фільмом, який на 70 % був створений шляхом використання комп'ютерної графіки, став «Аватар» режисера Дж. Камерона, який вийшов на екрани у 2009 р. Проте, як зазначає Е. Харст (E. Hurst), відмітною ознакою технології Deepfake є її три найбільш вражаючі характеристики – масштаб їх використання, сфера застосування і високий ступінь довершеності[2]. До цього переліку слід додати ще доступність технології Deepfake для широкого загалу та відсутність вимоги глибоких знань у сфері комп'ютерного програмування для її користувачів. Наразі більшість програм для створення діпфейків можна безкоштовно завантажити з відкритих інтернет-платформ на зразок TensorFlow або GitHub. З докладними інструкціями зі створення діпфейків будь-хто може ознайомитися на різних сайтах мережі Інтернет. Так, наприклад, сайт proglib.io надає докладну доступну усім користувачам інтернет-інструкцію зі створення діпфейків за допомогою нейромережевої програми DeepFaceLab, що змінює обличчя в відеороликах[3], проте застережень щодо застосування вказаної програми автори сайту не зазначають. І таких комп'ютерних програм та інтернет-платформ існує велика кількість. Не дивно, що за перше півріччя 2020 р. кількість діпфейків, виявлених у мережі Інтернет, збільшилася більш, ніж удвічі. З урахуванням тенденції до стрімкого зростання цього виду мережевої активності (див. Табл. 1)[4], кількість підроблених фото і відео, виготовлених з використанням технології Deepfake, збільшуватиметься і в подальшому.

[1] Алексіюк Є. Медійна реальність у стилі deep fake. URL: https://cedem.org.ua/articles/deep-fake/ (дата звернення: 01.03.2021).
[2] Hurst E. How can the law deal with Deepfake? URL: https://www.allaboutlaw. co. uk/commercial-awareness/legal-spotlight/how-can-the-law-deal-with-deepfake- (дата звернення: 01.03.2021).
[3] DeepFake-туториал: создаем собственный дипфейк в DeepFaceLab. URL: https://proglib. io/p/deepfake-tutorial-sozdaem-sobstvennyy-dipfeyk-v-deepfacelab-2019-11-16 (дата звернення: 01.03.2021).
[4] Ajder H. Deepfake Threat Intelligence: a statistics snapshot from June 2020. URL: https://sensity. ai/deepfake-threat-intelligence-a-statistics-snapshot-from-june-2020/ (дата звернення: 01.03.2021).

Таблиця 1. Кількість діпфейків, виявлених в мережі Інтернет



Технологія Deepfake вже знайшла своє застосування у низці сфер індустрії, освіти, розваг, соціальних медіа тощо. Окрім кіноіндустрії, технологію Deepfake використовують для локалізації реклами, коли відеоролик зі світовою знаменитістю поєднують із відео з місцевим актором, який декламує рекламні слогани рідною мовою. Це надає можливість створити видимість, буцім знаменитість говорить мовою країни, в якій здійснюється продаж рекламованого товару чи послуги. До речі, такий само підхід можна застосовувати і для спрощення процесу кінодубляжу. Цифрові двійники у відеоіграх, створені за допомогою технології Deepfake, використовують для більшого емоційного залучення гравців. Самостійно створених цифрових двійників (реплік) пропонується використовувати для віртуальних примірок в інтернет-крамницях. Застосування технології Deepfake також може бути корисним для подолання мовного бар'єру під час перекладу доповідей на відеоконференціях з одночасною зміною міміки і руху губ доповідача[1].

Безперечно, широкі можливості існують і для застосування технології Deepfake у гуманітарній сфері, зокрема для відтворення образів відомих постатей з мистецькою або навчальною метою. Так, зокрема, у травні 2019 р. співробітники музею Сальвадора Далі у м. Санкт-Пітербург, штат Флорида (США) у співробітництві з компанією Goodby Silverstein & Partners відкрили відеоінсталяцію the Dalí Lives до 115-річчя з дня народження митця. Для відтворення образу С. Далі комп'ютерний алгоритм на кшталт Deepfake проаналізував його прижиттєві відеозаписи, організатори доповнили

[1] Westerlund M. The Emergence of Deepfake Technology: A Review. Technology Innovation Management Review. November 2019 (Volume 9, Issue 11). P. 41.

реальні відеоматеріали новими діалогами, створеними на основі інтерв'ю, книг та особистої переписки митця, а потім за допомогою технології «зміни обличчя» було створено інтерактивний образ митця шляхом накладання вихідних даних на гру актора, подібного до С. Далі за фізичними параметрами. Інсталяція the Dalí Lives складається з 125 різних відеороликів, на основі яких реалізовується один з 190 512 сценаріїв залежно від реакції відвідувачів, серед яких навіть передбачається можливість зробити селфі з митцем[1]. Водночас слід зазначити, що використання відеоінсталяцій і голограм осіб, які пішли з життя, з навчальною або мистецькою метою стає предметом публічних дискусій, зокрема щодо його етичності і можливого порушення посттанативних немайнових прав померлої особи. Так, зокрема, Секція 50-F Законодавства про цивільні права штату Нью-Йорк установлює заборону використання цифрових реплік померлих відомих особистостей або митців, що були мешканцями Нью-Йорка, з комерційною метою. Під цифровою реплікою (digital replica) закон розуміє новостворений оригінальний комп'ютерно генерований електронний перформанс особи з окремими і новоствореними звуком або аудіовізуальною роботою, в якому особа участі не брала та який є настільки реалістичним, що пересічний глядач міг б повірити, що цей перформанс здійснює особа, у ньому відображена, а не будь-яка інша особа[2]. Закон має набути чинності 29 травня 2021 р., і його дію буде розповсюджено на осіб, що пішли з життя після набуття законом чинності. Охорона вказаних немайнових прав померлого відповідно до цього закону триває протягом 40 років зі смерті особи. Водночас закон не обмежує використання цифрової репліки померлої особи в навчальних, творчих або інших суспільно корисних цілях.

Під час аналізу правомірності створення діпфейків за українським законодавством слід урахувати, що вказані гіперреалістичні фото- й відеопідробки можуть створюватися як за згодою особи, щодо образу якої робиться новий відеоконтент, так і без згоди «персонажу». За наявності згоди особи, образ якої використовується для підробки, створення діпфейкого відео за зальним правилом не суперечить нормам чинного законодавства. Виняток стосується лише випадків, якщо створене відео порушує етичні норми або чинні правові приписи, особа поширює дезінформацію, діпфейк виступає способом вчинення певного кримінального правопорушення тощо. Водночас деякі випадки створення діпфейків хоча безпосередньо і не порушують чинних норм законодавства, але породжують занепокоєння щодо можливості їх використання для маніпуляції свідомістю інших осіб. Так, у лютому 2020 р. президент Індійської народної партії (Bharatiya Janata Party) Маной Тіварі (Manoj Tiwari) за один день до парламентських виборів використав технологію Deepfake у передвиборчій кампанії, створивши відео, в якому він звернувся

[1] Aouf R. S. Museum creates deepfake Salvador Dalí to greet visitors. URL: https://www.dezeen.com/2019/05/24/salvador-dali-deepfake-dali-musuem-florida (дата звернення: 01.03.2021).
[2] Civil Rights. Consolidated Laws of New York. URL: https://www.nysenate.gov/legislation/laws/CVR (дата звернення: 01.03.2021).

до своїх виборців з промовою на індоарійському діалекті, яким він у дійсності не володіє. Відео, яке стало відразу вірусним, переглянуло 15 млн осіб у 5 800 групах у додатку WhatsApp[1]. Хоча М. Тіварі й не порушував індійських законів, використання технології Deepfake у виборчій кампанії спричинило занепокоєння далеко за межами країни.

Окремо слід зупинитися на особливостях створення діпфейків без згоди особи, щодо образу якої здійснюється підробка. У вітчизняному законодавстві не існує спеціальних норм щодо вказаного питання, що вимагає звернутися до загальних принципів захисту інтересів фізичної особи під час проведення фото-, кіно-, теле- та відеозйомок та під час використання фотографій та інших художніх творів, що зображують особу, закріплених у статтях 307 і 308 Цивільного кодексу України. Не зосереджуючись на детальному розкритті особливостей охорони вказаних немайнових прав особи, у контексті аналізованого нами питання слід зазначити, що вільне використання художніх творів із зображенням фізичної особи (до яких можна віднести і діпфейки) без згоди останньої може дозволятись за наявності однією з таких умов:

– зйомки проводилися відкрито на вулиці, на зборах, конференціях, мітингах та інших заходах публічного характеру (згода особи зрезюмується);

– у разі поширення інформації про історичних постатей (за відсутності мети комерціалізації вказаної діяльності);

– з навчальною чи науковою метою або під час здійснення заходів, які мають історичну, культурну або іншу цінність для суспільства;

– під час створення карикатур, пародій та інших творів саркастичного жанру[2].

Розробники і користувачі програм, що використовують технологію Deepfake, найчастіше застосовують останнє із вищевказаних застережень для аргументації правомірності використання образу інших людей, особливо відомих постатей, для створення діпфейкових «персонажів». Так, наприклад, Р. Могильний, розробник популярного українського додатку Reface App, що надає можливість замінювати обличчя у зображеннях у форматі GIF і відео, чітко визначає мету використання цього додатку як «креатив і фан». Як указує Р. Могильний, «ми не підтримуємо створення deepfake та заохочуємо інших використовувати технологію заміни обличчя лише в хороших цілях», зокрема з метою створення пародійного контенту[3].

Відповідно до статті 1 Закону України «Про авторське право і суміжні права» пародія є твором, який є творчою переробкою іншого правомірно оприлюдненого твору або його частини, що за своїм змістом має комічний,

[1] Deepfakes inIndian politics: BJP use the tech to reach out to voters in Delhi. URL: https://in.mashable.com/tech/11562/deepfakes-in-indian-politics-bjp-uses-the-tech-to-reach-out-to-voters-in-delhi (дата звернення: 01.03.2021).

[2] Кулініч О. Охорона інтересів фізичної особи при створенні й використанні фотографій та інших художніх творів з її зображенням. *Теорія і практика інтелектуальної власності*. 2015. № 6. С. 37.

[3] Міняйло Н. Розробник додатка Reface App: «Мета використання нашої програми – це креатив і фан». URL: https://ms.detector.media/it-kompanii/post/25408/2020-09-03-rozrobnyk-dodatka-reface-app-meta-vykorystannya-nashoi-programy-tse-kreatyv-i-fan/ (дата звернення: 01.03.2021).

сатиричний характер або спрямовується на висміювання певних подій або осіб. Згідно зі статтею 21 вказаного закону створення пародій і карикатур допускається без згоди автора (чи іншої особи, яка має авторське право), але з обов'язковим зазначення імені автора і джерела запозичення[1]. Отже, створення діпфейків з метою надати одній особі подібність до іншої особи, відображеної на певному фото або відео, без згоди останньої може здійснюватись лише за наявності сукупності таких умов:

– як вихідний матеріал використовується правомірно оприлюднене фото, відео або художній твір;

– у разі, якщо вихідний матеріал є предметом охорони авторського права, обов'язковим є зазначення імені автора і джерела запозичення або отримання згоди автора відповідного твору (чи іншої особи, яка має авторське право);

– метою створення діпфейку є виготовлення продукту пародійного жанру.

Окрім вищезазначених умов, існує і низка загальних застережень щодо створення творів з використанням зображення іншої особи без її згоди. Як зазначає О. О. Кулініч, найбільш важливими серед таких є:

– дотримання права фізичної особи на гідне, достовірне й автентичне власне зображення (ст. ст. 297, 299, 308 ЦК, ст. 3 Закону України «Про захист суспільної моралі»);

– дотримання конфіденційності інформації під час поширення художнього твору із зображенням фізичної особи (ст. ст. 301, 442 ЦК);

– дотримання права на честь, гідність, ділову репутацію під час поширення художнього твору із зображенням фізичної особи (ст. ст. 297, 299, 442 ЦК);

– дотримання норм публічної моралі під час поширення художнього твору із зображенням фізичної особи (ч. 3 ст. 13, ст. 2 ст. 319 ЦК, ст. 3 Закону України «Про захист суспільної моралі»)[2].

Серед інших застережень, що можуть стати підставою для визначення створеного діпфейку порушенням прав особи в деяких іноземних юрисдикціях, є мета умисного заподіяння емоційних страждань. Проте зазначена підстава може бути використана лише у тому разі, якщо позивач зможе довести, що діпфейк емоційно травмував його такою мірою, що перейшла розумні межі суспільної пристойності[3].

Отже, в разі створення діпфейкового пародійного продукту або діпфейку з навчальною, науковою чи іншою суспільно корисною метою особа не має порушувати вищевказані немайнові права особи, а також загальноприйняті норми суспільної моралі. У протилежному випадку

[1] Про авторське право і суміжні права : Закон України від 23.12.1993 № 3792-XII // БД «Законодавство України» / ВР України. URL: https://zakon.rada.gov.ua/laws/show/3792-12#Text (дата звернення: 01.03.2021).

[2] Кулініч О. Охорона інтересів фізичної особи при створенні й використанні фотографій та інших художніх творів з її зображенням. *Теорія і практика інтелектуальної власності*.2015. № 6. С. 38.

[3] Fink D., Diamond S. Deepfakes: 2020 and Beyond. URL: https://www.law.com/therecorder/2020/09/03/deepfakes-2020-and-beyond/?slreturn=20210124103514 (дата звернення: 01.03.2021).

потерпіла особа, образ якої був використаний для створення діпфейку, має право звернутися до суду. Проте знаменитості, пам'ятаючи про так званий ефект Стрейзанд, який полягає в тому, що спроби видалити або приховати якусь інформацію призводять до вірусного розповсюдження забороненого контенту[1], найчастіше такі випадки ігнорують, щоб не привертати зайвої уваги користувачів інтернет-мережі до такої інформації. У разі порушення діпфейком норм суспільної моралі заборона такого відео може бути здійснена представником влади або власниками інтернет-ресурсу. Наразі найбільш відомі інтернет-платформи такі, як Facebook, Twitter і Google, розробляють зміни до правил розміщення на них діпфейків, однак при цьому це не зупиняє зазначені сервіси від реклами додатків, призначених для створення діпфейкових фото і відео.

Після аналізу правомірного використання технології Deepfake або форм її використання, що спричинюють занепокоєння громадськості, проте прямо не порушують правових приписів, хотілося б приділити окрему увагу використанню діпфейків для вчинення кримінальних правопорушень.

Як зазначають дослідники, у 96 % випадків діпфейки використовуються для створення порнографічних відео з відомими жінками[2]. Одним із перших таких порнографічних діпфейків став відеоролик, розміщений у мережі Інтернет у грудні 2017 р., в якому обличчя героїні було змінено на обличчя ізраїльської акторки Галь Гадот. Творець зазначеного фейкового відео – звичайний програміст, який цікавився технологіями машинного навчання, майже відразу розкрив інформацію, що для створення порноролику з відомою акторкою він використав комп'ютерну програму типу TensorFlow, яку Google надавав у вільному доступі своїм користувачам[3]. Менш ніж за місяць створення порнороликів за допомогою технології Deepfake стало надзвичайно популярним. При цьому як зазначає М. А. Вальорска, жертвами діпфейкерів ставали в переважній більшості жінки: програми на кшталт DeepNude за лічені секунди перетворювали фото або відео, на якому зображена жінка, на компрометуючі еротичні матеріали; проте в разі спроби використати вказану програму для зображення чоловіка програма просто генерувала жіночі геніталії[4].

Іншою небезпекою створення підроблених порнороликів є можливість їх використання з метою «порнопомсти» (*Revenge Porn*), погрози використанням інтимних візуальних матеріалів для шантажу, спричинення шкоди репутації особи або її дискредитації. Наразі в Україні заподіяння дифамаційної шкоди фізичній особі не переслідується в кримінально-правовому порядку, проте створення порнографічної продукції, зокрема з підробленими персонажами, може бути кваліфіковано за статтею 301 КК

[1] Jansen S. C., Martin B. The Streisand Effect and Censorship Backfir. *International Journal of Communication.* 2015. Vol. 9. P. 656–657.

[2] Вальорска М. А. Діпфейк та дезінформація : практ. посіб. / пер. з нім. В. Олійника. Київ, 2020. С. 9.

[3] Flint J. A fake porn video of Gal Gadot highlights social ramifications of machine learning. URL: https://www.modmy.com/fake-porn-video-gal-gadot-highlights-dangers-machine-learning (дата звернення: 01.03.2021).

[4] Вальорска М. А. Діпфейк та дезінформація : практ. посіб. / пер. з нім. В. Олійника. Київ, 2020. С. 20.

України «Ввезення, виготовлення, збут і розповсюдження порнографічних предметів». Цікавим є той факт, що в деяких країнах, зокрема в Законодавстві про цивільні права штату Нью-Йорк, США (Секція 50-F), вже встановлено спеціалізовану норму щодо заборони створення реалістично зображених порнографічних матеріалів фізичних осіб, виготовлених шляхом діджилалізації персонажа[1], якими фактично є аналізовані вище порнографічні діпфейки. Такі підроблені матеріали стають настільки популярними, що навіть найбільші порносайти підтримують їх розміщення на своїх хостінгах. На думку експертів, у подальшому розвиток Deepfake-порно скоріш за все призведе до створення спеціалізованих майданчиків[2].

Іншою сферою використання зловмисниками технології Deepfake стали кібершахрайства. Для досягнення злочинної мети шахраї використовують незаконне використання чужих особистих даних (*identity theft*), зокрема голосу головних виконавчих директорів компаній. Один з найперших і водночас найбільш відомий випадок шахрайського використання технології Deepfake відбувся у 2019 р., коли невідомі шахраї шляхом використання алгоритму глибинного навчання GAN створили високоякісну імітацію голосу директора німецької компанії і за допомогою телефонного зв'язку від його імені наказали Генеральному директору дочірньої енергетичної компанії з Великобританії відправити кошти на суму 220 тис. євро угорському постачальнику. Повернути зазначені гроші не вдалося: з угорського банку гроші були миттєво переведені до Мексики, а потім розділені й переведені до інших локацій[3].

Ще одним способом використання технології Deepfake для вчинення шахрайства в Інтернеті є створення фейкових відео з відомими особами для заманювання потерпілих на фішингові сайти. Хоча такі шахрайські схеми наразі є порівняно нечисленними, подальше вдосконалення технології Deepfake може збільшити її використання у фішингових кібератаках.

Проте найбільш небезпечними наразі вважають зростаючі можливості використання діпфейків для маніпуляції свідомістю широких мас населення. Фейковий контент може бути використаний з метою підбурювання громадського незадоволення, підриву довіри громадськості до інформації, що надається представниками влади, розпалювання ненависті до певних груп населення або конкретних особистостей тощо. Політичні діячі вже зараз називають діпфейки новою зброєю дезінформації, яка може застосовуватися під час проведення інформаційних і гібридних війн, і закликають негайно вступити у діалог щодо обмеження розроблення та використання такої «зброї»[4]. Деякі країни вже почали активну протидію

[1] Civil Rights. Consolidated Laws of New York. URL: https://www.nysenate.gov/legislation/laws/CVR (дата звернення: 01.03.2021).

[2] Агутин Н. Крупнейшие порно-сайты НЕ против Deepfake-видео. URL: https://dtf.ru/aboutporn/201107-krupneyshie-porno-sayty-ne-protiv-deepfake-video (дата звернення: 01.03.2021).

[3] Damiani J. A Voice Deepfake Was Used To Scam A CEO Out Of $243,000. URL: https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/?sh=75a961842241 (дата звернення: 01.03.2021).

[4] Корсунський С. «Оборонний суверенітет» Європи в епоху DeepFake. URL: https://zn.ua/ukr/international/oboronniy-suverenitet-yevropi-v-epohu-deepfake-323491_. html (дата звернення:

зростаючим потокам дезінформації. Наприклад, у КНР з 1 січня 2020 р. набуває чинності закон, який установлює кримінальну відповідальність за публікацію діпфейків або фейкових нових без указівки, що такий контент був створений з використанням технологій штучного інтелекту або віртуально реальності. При цьому влада КНР залишає за собою право притягти до відповідальності як користувача, що створив або розповсюдив діпфейк, так і хостинг, який не забезпечив дотримання вказаного закону[1]. Такий законопроєкт з назвою-акронімом DEEP FAKES Accountability Act (Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019) був презентований Конгресу США 6 червня 2019 р. Він передбачав обов'язок маркувати діпфейки цифровим водяним знаком (*Digital Watermark*) та аудіовізуальним ідентифікатором (*Audiovisual Disclosure*)[2]. Проте він так і не був ухвалений через невідповідність деяким положенням Першої поправки до Конституції США, яка гарантує свободу совісті, слова, преси, зборів і петицій, а також неефективності передбаченого механізму протидії діпфейкам, що походять з іноземних юрисдикцій.

Окремим аспектом неправомірного використання технології Deepfake є можливість здійснення протиправного впливу або маніпуляцій під час проведення передвиборчих кампаній. Як доводить практика, застосування навіть дешевих, або поверхневих фейків (*Cheap* або *Shallow Fakes*) – підробок, створених простими технічними засобами, здатне спричинити гучні політичні скандали та вплинути на політичне життя країни. Так, створений у травні 2019 р. відеоролик зі спікером Палати представників Конгресу США Ненсі Пелосі, в якому оригінальна швидкість запису була знижена приблизно до рівня 75 %, а висота звуку була підвищена для підтримки природного тембру голосу, створив враження, що вона читає промову в стані сп'яніння. Фейкове відео набуло вірусного поширення в мережі Інтернет, лише в Facebook його було переглянуто 1,4 млн разів. Визнавши відео лише «частково підробленим», керівництво Facebook не видалило відеоролик, а лише позначило його як підроблений, що за словами менеджера в зв'язках з громадськістю Facebook Енді Стоуна (Andy Stone) значно знизило кількість його подальших переглядів. Відсутність належної реакції з боку Facebook призвело до повтору створення такого фейкового відео стосовно Н. Пелосі у серпні 2020 р., в результаті чого його було переглянуто вже більше 2 млн разів[3]. Протягом 2018–2020 років в інтернет-мережі отримала поширення значна кількість інших відео, створених з

01.03.2021).

[1] Statt N. China makes it a criminal offense to publish deepfakes or fake news without disclosure. URL: https://www.theverge.com/2019/11/29/20988363/china-deepfakes-ban-internet-rules-fake-news-disclosure-virtual-reality (дата звернення: 01.03.2021).

[2] Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to Accountability Act of 2019. URL: https://www.congress.gov/bill/116th-congress/house-bill/3230/text (дата звернення: 01.03.2021).

[3] Denham H. Another fake video of Pelosi goes viral on Facebook. URL: https://www.washingtonpost.com/technology/2020/08/03/nancy-pelosi-fake-video-facebook (дата звернення: 01.03.2021).

використанням технології Deepfake, на яких політики вчиняють дії, які в реальному житті вони не вчиняли, або на яких вони роблять компрометуючі фейкові заяви[1]. Видається, що поширення таких фейкових або «частково підроблених» відео, спрямованих на створення «чорного піару» певним політичних силам або громадським діячам, має нарешті стати приводом для ґрунтовного перегляду своїх внутрішніх політик найбільш популярними соціальними мережами, а також для передбачення відповідних заборон на законодавчому рівні. У зв'язку з цим слід зазначити, що першими штатами США, які законодавчо заборонили використання діпфейків у передвиборчій період, стали Каліфорнія і Техас. У вересні 2019 р. у Техасі був ухвалений закон (Senate Bill 751) щодо криміналізації розміщення або публікування діпфейкових відеороликів з метою негативного впливу на кандидата або на виборчий процес у період 30 днів до проведення виборів[2]. У жовтні 2019 р. у Каліфорнії був ухвалений закон (Assembly Bill 730), який передбачив цивільно-правову відповідальність за створення і розповсюдження фейкових аудіо- і відеозаписів кандидатів у період 60 днів до проведення виборів[3]. Проте за відсутності подальшого розширення державного регулювання у вказаній сфері та через досягнення домовленостей з власниками найбільш популярних інтернет-ресурсів ефективність зазначених законодавчих новел може залишитися доволі низькою.

Висновки. Кримінологічний аналіз використання технології Deepfake демонструє різноманітність і водночас суперечливість способів її використання. Найбільш вузьким питанням у цьому контексті видається використання особистих даних іншої особи, яке лише за наявності низки умов можна визнати правомірним. Водночас очевидним є те, що зараз потенціал використання технології Deepfake далеко не повною мірою оцінено як суспільством загалом, так і злочинним світом зокрема. Це дає змогу зробити обґрунтоване припущення, що кількість та види кримінальних правопорушень, що вчиняються з використанням технології Deepfake, в подальшому лише зростатимуть. Видається, що найбільш небезпечними з них можуть стати правопорушення, пов'язані зі спробою маніпуляції значними групами населення та вчиненням гібридних диверсійних атак. Вважаємо, що протидіяти протиправному використанню технології Deepfake можна лише за умови об'єднання зусиль урядів різних держав і власників провідних інтернет-ресурсів. Методологічною основою такої протидії має стати комплексне оцінювання наявних механізмів протидії цифровим злочинам і можливостей їх використання у протидії правопорушенням, вчиненим з використанням технології Deepfake, а також визначення необхідності встановлення нових додаткових форм та організаційних засад протидії їм. Обсяг цієї статті не надає

[1] Див., наприклад: Vincent J. Watch Jordan Peele use AI to make Barack Obama deliver a PSA about fake news. URL: https://www.theverge.com/tldr/2018/4/17/17247334/ai-fake-news-video-barack-obama-jordan-peele-buzzfeed (дата звернення: 01.03.2021).

[2] Senate Bill 751. Texas Legislature Online. URL: https://capitol. texas.gov/BillLookup/History. aspx?LegSess=86R&Bill=SB751 (дата звернення: 01.03.2021).

[3] Assembly Bill No.730 Elections: deceptive audio or visual media. California legislative information. URL: https://leginfo. legislature. ca.gov/faces/billNavClient. xhtml?bill_id=201920200AB730 (дата звернення: 01.03.2021).

можливості висвітлити весь арсенал наявних наразі правових, організаційних і технічних засобів, що вже застосовуються у протидії несумлінному використанню технології Deepfake, так само як ефективність і недоліки їх застосування, тому це має стати предметом подальших кримінологічних досліджень у вказаній сфері.

Стаття надійшла до редакції 03. 03.2021

Ксения Владимировна ЮРТАЕВА,
кандидат юридических наук, доцент
(Харьковский национальный университет внутренних дел)

КРИМИНОЛОГИЧЕСКИЙ АНАЛИЗ ИСПОЛЬЗОВАНИЯ ТЕХНОЛОГИИ DEEPFAKE: КОГДА ФЕЙК СТАНОВИТСЯ ПРЕСТУПЛЕНИЕМ

В статье осуществлен комплексный криминологический анализ использования технологии Deepfake, предназначенной для создания гиперреалистических поддельный фото- и видеоизображений на основе алгоритма глубокого обучения GAN. Определены сферы применения технологии Deepfake в современном обществе, охарактеризованы правовые основы ее применения. Охарактеризованы возможности противоправного применения технологии Deepfake и ее использования для совершения правонарушений. Дан криминологический прогноз относительно дальнейшего расширения использования технологии Deepfake в преступной деятельности. Определены основные направления противодействия криминогенным рискам, которые несет технология Deepfake для современного общества.

Ключевые слова: *глубокое обучение, искусственный интеллект, цифровая технология, Deepfake, информационная безопасность, криминогенные риски, криминологический анализ, правонарушение, противодействие преступности.*

Kseniya V. YURTAYEVA,
PhD in Law, associate professor
(Kharkiv National University of Internal Affairs)

CRIMINOLOGICAL ANALYSIS OF APPLYING DEEPFAKE TECHNOLOGY: WHEN A FAKE EVOLVES INTO A CRIME

The author of the article conducts comprehensive criminological analysis of applying Deepfake technology, designed for creating hyperrealistic photo and video images based on GAN deep learning algorithm. Spheres of applying Deepfake technology in modern society are defined; legal bases for its application are characterized. The author of the article characterizes illegal methods of applying Deepfake technology and its facilitation of criminal undertakings. Criminological prognosis for future expansion of applying Deepfake technology in criminal activities is provided. The main directions for preventing criminogenic risks posed by emergence of Deepfake technology to modern society are outlined in the article.

Key words*: deep learning, artificial intelligence, digital technology, Deepfake, informational safety, criminogenic risks, criminological analysis, offence, crime prevention.*