УДК 343.9:343.85(477) *DOI: https://doi.org/10.32631/vca.2025.2.38*

Олексій Миколайович ЛИТВИНОВ,
доктор юридичних наук, професор,
заслужений працівник освіти України
(Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», м. Харків)

Кирило Олександрович ЧЕРЕВКО,
кандидат юридичних наук, доцент
(Харківський національний університет внутрішніх справ, м. Харків)

НОВІТНІ ВИКЛИКИ У ВИЯВЛЕННІ ТА ПРОТИДІЇ ПОШИРЕННЮ ДИТЯЧОЇ ПОРНОГРАФІЇ, СТВОРЕНІЙ З ВИКОРИСТАННЯМ ІНСТРУМЕНТІВ ШТУЧНОГО ІНТЕЛЕКТУ

У статті здійснено аналіз зростаючою та багатогранної загрози, яку становить дитяча порнографія, створена з використанням інструментів штучного інтелекту. В статті висвітлюються ключові виклики у виявленні та протидії цьому явищу, а також його глибокі психологічні та соціальні наслідки для жертв, суспільства в цілому. Досліджується, як швидкий розвиток генеративного штучного інтелекту уможливлює створення дедалі реалістичніших зображень, відеоматеріалу, що ускладнює роботу правоохоронних органів та ставить під загрозу безпеку дітей у цифровому просторі. Крім того, окреслюються комплексні підходи до боротьби з дитячою порнографією, створеною з використанням інструментів штучного інтелекту, включаючи посилення законодавства, розробку передових технологічних рішень для виявлення та блокування, впровадження етичних принципів у розробці штучного інтелекту, проведення широкої просвітницької роботи та забезпечення ефективної підтримки жертв.

Ключові слова: *штучний інтелект, генеративні моделі, цифрова*

безпека, дитяча порнографія, кіберзлочинність, виявлення контенту, блокування зображень, жертва, безпека дітей, міжнародна співпраця.

Постановка проблем. Створення та поширення дитячої порнографії – одна з найбільш чутливих для суспільної моралі і кожної не позбавленої відчуття людяності совісті проблема. Це ганебне та небезпечне явища не лише здатне серйозно підривати фізичне та психічне здоров'я дітей, але й розбещувати їх, деформувати параметри (до перверсій) статевого потягу, формуючи підґрунтя до самодетермінації цих злочинів, підживлюючи попит на дитячу порнографію. Тож протидія створенню і поширенню останньої має відводитись пріоритетне значення у системі заходів захисту дітей, суспільної моралі. Водночас інтенсивних розвиток комп'ютерних, мережевий технологій ставить на порядок дня новітні виклики, пов'язані з використанням інструментів штучного інтелекту (далі – ШІ) для генерації порнографічних зображень, відеоматеріалів, в тому числі з образами дітей. На ці виклики має бути адекватна відповідь.

Різні аспекти проблеми протидії злочинності у кіберпросторі досліджували такі вчені-юристи як А. В. Боровик, П. П. Галушко, М. О. Думчиков, О. М. Кравцова, О. В. Манжай, Д. В. Пашнєв, О. С. Передерій, В. В. Сокуренко та інші. Проте проблема використання інструментів ШІ для створення і поширення дитячої порнографії й досі не ставала предметом окремого наукового пошуку.

Мета статті – виявити, описати основні кримінальні та криміногенні загрози, пов'язані з використанням інструментів ШІ для створення і поширення дитячої порнографії, сформувати пропозиції щодо протидії цьому явищу.

Викладення основного матеріалу. Проблема дитячої порнографії, створеної з використанням інструментів штучного інтелекту (далі – ДПсШІ), є однією з найгостріших та найскладніших загроз у сучасному цифровому світі. ДПсШІ визначається як форма синтетичного медіа (*deepfake*)[1], що реалістично зображує сексуальну експлуатацію дітей, створена за допомогою алгоритмів машинного навчання. Важливо, що сам ШІ не створює цей контент; він є лише інструментом, який використовується злочинцями для генерації та поширення таких матеріалів.

Ми входимо в епоху, коли довіряти власним очам стає дедалі складніше. Завдяки генеративному штучному інтелекту та deepfake-технологіям з'являються зображення й відео, настільки реалістичні, що відрізнити їх від справжніх практично неможливо. Це створює серйозні виклики для суспільства, людини та правоохоронних органів, оскільки межі між реальним та синтетичним контентом розмиваються. Масштаби проблеми постійно зростають: звіт *Internet Watch Foundation (IWF)*[2] виявив

[1] Binder M., Newbury E. M. H. Combatting AI-Generated CSAM / Wilson Center. 2025. 5 Feb. URL: https://www.wilsoncenter.org/article/combatting-ai-generated-csam.

[2] How AI is being abused to create child sexual abuse imagery / Internet Watch Foundation. URL: https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-

понад 20 000 ШІ-згенерованих зображень на одному даркнет-форумі за місяць, з яких понад 3 000 були протизаконними. У Великій Британії зафіксовано зростання кількості повідомлень про ШІ-згенерований матеріал сексуального насильства над дітьми (*Child Sexual Abuse Material – CSAM*)[1] на 380% у 2024 році порівняно з 2023 роком. Національний центр зниклих та експлуатованих дітей (*NCMEC*) також повідомив про отримання близько 5000 звітів про ШІ-згенеровані медіафайли експлуатації дітей у 2023 році[2].

Поширення ДПсШІ несе в собі низку серйозних загроз, які глибоко істотно на суспільство, безпеку та благополуччя дітей. До них можна віднести:

1. «Нормалізація» та поширення сексуального насильства над дітьми. Основним ризиком ДПсШІ є те, «що» ця технологія вона створює, а не те, «як» вона може змінити суспільство. «Десенсибілізація насильства»[3] – психологічний феномен, який полягає у тому, що тривалий вплив жорстоких зображень, навіть імітованих, може з часом притупити емоції людини. Частина з нас може спробувати себе заспокоїти, що насильство не існує, коли ми бачимо його не так. Але когнітивне викривлення продовжується в нашому мозку. Постійний потік такого контенту навчає відвертатися, знижувати чутливість і сприймати злочини як статистику, а не як реальну проблему.

2. Стимуляція педофільних нахилів, спонукання до реальних кримінальних правопорушень. Оскільки велика кількість зображень, створених ШІ, легкодоступна, це може підживлювати сексуальні фантазії педофілів і спонукати їх до реальних злочинів проти дітей. Дослідження Д. Фінкельгора (*D. Finkelhor*) підтверджують, що споживання контенту, який демонструє сексуальну експлуатацію неповнолітніх, підвищує ймовірність сексуальної експлуатації дітей і вчинення злочинів проти них[4]. Перегляд CSAM значною мірою пов’язаний із вчиненням злочинів проти статевої недоторканості. Це підкреслює прямий причинно-наслідковий зв’язок між споживанням синтетичного CSAM і підвищенням ризику вчинення справжніх злочинів проти дітей. Таким чином, «несправжній» контент становить реальну загрозу фізичній безпеці дітей[5].

---

abuse-imagery/

[1] Britain's leading the way protecting children from online predators / Home Office ; GOV.UK. 2023. 10 May. URL: https://www.gov.uk/government/news/britains-leading-the-way-protecting-children-from-online-predators.

[2] Annual Report 2023 / National Center for Missing & Exploited Children. 2024. URL: https://www.missingkids.org/footer/about/annual-report.

[3] Quayle E., Taylor M. Child pornography and the Internet: A review of the evidence. *Psychology, Crime & Law*. 2002. Vol. 8. Is. 4. P. 291–319. URL: https://www.tandfonline.com/doi/abs/10.1080/10683160208401824.

[4] Finkelhor D. The Challenge of Creating a Global Indicator of Violence Against Children / Trauma, Violence, & Abuse. 2025. 11 Sept. URL: https://journals.sagepub.com/doi/10.1177/15248380251357613.

[5] Letourneau E. J. et al. Child Sexual Abuse and Boundary-Violating Behaviors in Youth-Serving Organizations: A Scoping Review / Trauma, Violence, & Abuse. 2024. URL: https://pure.johnshopkins.edu/en/publications/child-sexual-abuse-and-boundary-violating-behaviors-in-youth-serv.

3. Ускладнення процесу розслідування та ідентифікації жертв. Використання інструментів ШІ для створення порнографії значно ускладнює роботу поліції та прокуратури. У таких зображеннях практично неможливо знайти реальних жертв. Крім того, надзвичайно важко довести, що зображення були створені штучним інтелектом, а не реальними людьми. Навіть для досвідчених аналітиків і експертів зразки дитячої порнографії, створені штучним інтелектом, візуально не відрізняються від справжнього CSAM[1]. Слідчі органи витрачають час і обмежені ресурси на розслідування правопорушень, яких не було, якщо штучно створений контент стає візуально невідрізним від реального. Ці ресурси могли б допомогти знайти та врятувати реальних жертв. Це збільшує навантаження на судові та правоохоронні органи, які вже перевантажені через зростання онлайн-злочинності. Це також може відволікати значні ресурси від боротьби зі справжнім CSAM.

4. Проблеми, пов'язані з участю малолітніх у створенні порнографічних матеріалів. Існує ризик, що малолітні можуть використовувати ШІ для створення власних порнографічних матеріалів або для створення фейкових зображень інших дітей. Це може призвести до десенсибілізаціних наслідків для них самих та інших дітей. Діти стають не лише пасивними жертвами, а й несвідомими або маніпульованими учасниками створення шкідливого контенту, часто через тиск однолітків, нерозуміння наслідків або навіть як форма ШІ-булінгу (*AI-Bullying*)[2]. Це ускладнює картину віктимізації та вимагає специфічних підходів до профілактики та втручання.

5. Розповсюдження неправдивої інформації та дискредитація осіб. Порнографія, згенерована ШІ, може використовуватися для поширення дезінформації та дискредитації конкретних осіб, що може мати серйозні наслідки для їхньої репутації, життя в цілому[3]. Зростання використання зображень, згенерованих ШІ, що містять відомих жертв сексуального насильства над дітьми або дітей, є особливо тривожним. Наслідки – більш значні, ніж просто дискредитація однієї особи. Вони включають руйнування фундаментальної довіри до всього цифрового контенту. Якщо «бачити – значить вірити» перестає бути правдою – основні принципи інформаційного простору будуть зруйновані. Використання фотографій відомих жертв[4] або селебреті[5] не тільки посилює травму, але й відкриває

[1] Europol Spotlight – Generative AI: The dark side looms over the sunny uplands / Europol Innovation Lab. October 2023. URL: https://www.europol.europa.eu/cms/sites/default/files/documents/Transcript-how-crime-is-accelerated-by-ai.pdf.

[2] Generative Artificial Intelligence and Cyber Bullying / Dolly's Dream. 2024. URL: https://www.dollysdream.org.au/blog/generative-artificial-intelligence-and-cyber-bullying

[3] Chesney R., Citron D. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National. *Lawfare*. 2018. 21 Feb. URL: https://www.lawfaremedia.org/article/deep-fakes-looming-challenge-privacy-democracy-and-national-security.

[4] Citron D. K. Hate Crimes in Cyberspace. Cambridge, Massachusetts : Harvard University Press, 2014. P. 118.

[5] Posetti J. et al. The Chilling: A global study of online violence against women journalists / The International Center for Journalists (ICFJ) ; UNESCO. 2022. URL: https://www.icfj.org/sites/default/files/2023-02/ICFJ%20Unesco_TheChilling_OnlineViolence.pdf.

новий аспект кібербулінгу та шантажу, у якому репутації можна завдати шкоди без реального кримінального правопорушення.

Психологічні та соціальні наслідки для жертв та суспільства:

–*психологічний вплив на жертву*. Діти, які стають жертвами візуального сексуального насильства, створеного штучним інтелектом, можуть відчувати сильне приниження, сором, гнів, почуття порушення особистих кордонів і почуття самозвинувачення. Це може призвести до негайного та тривалого емоційного тривожного стану, відсторонення від сім'ї та шкільного життя, а також проблем з підтримкою довірчих стосунків. Це може іноді призвести до самопошкодження, суїцидальних думок[1];

–*соціальний вплив та повторна віктимізація (ревіктимізація)*: якщо *deepfake* поширюються у шкільній спільноті або серед груп однолітків, жертва може стати об'єктом булінгу та переслідувань. Травма посилюється щоразу, коли контент поширюється. Крім того, якщо обличчя дитини використовується у *deepfake*-порнографії, це може завдати шкоди її репутації, знизити успішність у школі та зменшити впевненість у майбутньому, викликати побоювання, що зображення будуть постійно доступні онлайн, навіть якщо вони є фальшивими.

–*бар'єри для звернення по допомогу.* Дослідження серед молоді показали, що один з шести неповнолітніх, які були залучені до потенційно шкідливої онлайн-сексуальної взаємодії, ніколи не розкривають це нікому[2]. Хлопчики менш схильні розповідати іншим, коли вони стають жертвами *deepfake*-порнографії або CSAM. Зображення в *deepfake* може викликати страх, що їм не повірять, посилюючи бар'єри для звернення по допомогу. Це свідчить про специфічні соціальні та психологічні бар'єри для звернення по допомогу серед цієї демографічної групи. Подолання цього «мовчазного» аспекту проблеми вимагає цілеспрямованих програм підтримки та просвітницьких кампаній, адаптованих до потреб дітей;

–*вплив на глядачів та суспільство в цілому.* Споживання такого контенту може призвести до ризиків залежності, втрати контролю над переглядом, зниження інтересу до реальних сексуальних взаємодій[3]. Також спотворюються очікування від реальних сексуальних стосунків, завдається шкода образу тіла. У *deepfake* порнографії непропорційно часто експлуатуються жінки та діти.

Боротьба з ДПсШІ вимагає постійного вдосконалення методів виявлення, оскільки технології генерації ШІ розвиваються з безпрецедентною швидкістю.

На ранніх етапах розвитку ШІ, ДПсШІ можна було відрізнити за

[1] Kobriger K. et al. Out of Our Depth with Deep Fakes: How the Law Fails Victims of Deep Fake Nonconsensual Pornography. *Richmond Journal of Law and Technology*. 2021. Vol. 28. Is. 2. URL: https://scholarship.richmond.edu/jolt/vol28/iss2/1/

[2] State of the Issue: Annual Youth Survey on Online Sexual Interaction / Thorn. 2024. URL: https://www.thorn.org/research/state-of-the-issue/

[3] Баловсяк Н. Швидкий контент, повільний розум? Як короткі відео викликають залежність та впливають на мозок // DSnews.ua. 2025. 22 лют. URL: https://www.dsnews.ua/ukr/society/shvidkiy-kontent-povilniy-rozum-yak-korotki-video-viklikayut-zalezhnist-ta-vplivayut-na-mozok-22022025-517271.

низкою візуальних аномалій: «занадто ідеальні об'єкти», без природних недоліків, «неправдоподібні текстури», «проблеми з фоном» (розмитість, артефакти), «неправильні тіні та відблиски», «неправильна анатомія» (наприклад, зайві пальці, аномальні кінцівки), «нелогічні композиції», «повторювані елементи» та «неправильні відображення». Однак, Д. В. Хамула[1] прямо зазначає, що відрізнити зображення, створене штучним інтелектом, від реального фото може бути дедалі складніше з кожним днем, оскільки технології ШІ постійно вдосконалюються. Це є чітким, тривожним трендом, який показує, що традиційні візуальні ознаки як надійні індикатори швидко зникають, а методи виявлення, які були ефективними вчора, стають застарілими сьогодні.

Деякі генератори зображень додають водяні знаки, які вказують на те, що зображення було створено ШІ[2]. Крім того, інформація, прихована в зображенні (метадані), може містити підказки про його створення. Проте, хоча водяні знаки та метадані є потенційними індикаторами походження, їхня ефективність у боротьбі зі злочинністю значно обмежується тим, що зловмисники можуть легко їх видаляти, змінювати або підробляти. Це робить їх ненадійними в контексті кримінальних розслідувань, де метою є приховати походження контенту, а не маркувати його.

Існують різні спеціалізовані програми та онлайн-сервіси, які можуть допомогти визначити, чи було зображення згенеровано ШІ. Штучний інтелект є не лише джерелом загрози, але й потужним інструментом протидії злочинності, що дозволяє перейти від реактивного виявлення контенту до активного аналізу поведінки та фінансових потоків[3]. ШІ може бути використаний не тільки для виявлення зображень, але й для ідентифікації підозрілої поведінки, наприклад, аналізу комунікаційних тенденцій та лінгвістичних маркерів, що вказують на спонукання до сексуальних дій. Також ШІ може допомагати у виявленні фінансових транзакцій, пов'язаних з придбанням CSAM. Деякі організації, такі як, наприклад, *Thorn*[4], розробили інструменти машинного навчання для автоматичного виявлення, перегляду та звітування про CSAM у великих масштабах, що є завданням, занадто великим для виключно людської модерації. Департамент внутрішньої безпеки США (DHS)[5] використовує інструменти ШІ, такі як *Speech and Language Tool*, *PinPoint Tool* та *HORUS*

[1] Хамула Д. В. До питання ознак «закамуфльованої» дитячої порнографії у фотомистецтві при мистецтвознавчому дослідженні продукції порнографічного характеру // Криміналістичні та експертні аспекти протидії злочинності : матеріали Всеукр. наук.-практ. конф. (м. Одеса, 24 листоп. 2023 р.). Одеса, 2023. С. 273–276.

[2] Коваленко А. В. Метадані як джерело криміналістично значущої інформації // Використання цифрових технологій в криміналістиці і судовій експертизі : матеріали міжнар. наук.-практ. круглого столу (м. Харків, 11 груд. 2023 р.). Харків : Право, 2024. С. 82–85.

[3] Crimes against children / INTERPOL. URL: https://www.interpol.int/Crimes/Crimes-against-children

[4] How CSAM Classifiers Use AI to Build a Safer Internet / Thorn. 2025. URL: https://www.thorn.org/blog/how-thorns-csam-classifier-uses-artificial-intelligence-to-build-a-safer-internet/

[5] Artificial Intelligence and Combatting Online Child Sexual Exploitation and Abuse / U.S. Department of Homeland Security. 2024. URL: https://www.dhs.gov/sites/default/files/2024-09/24_0920_k2p_genai-bulletin.pdf.

(Face Recognition), для підтримки ідентифікації жертв та виявлення, ліквідації транснаціональної онлайн-сексуальної експлуатації дітей.

Незважаючи на значні досягнення, існуючі інструменти виявлення стикаються з численними проблемами. Так, проблеми узагальнення вирішуються за допомогою детекторів, які добре працюють на одному наборі даних, але часто погано працюють на новому, незнайомому, без додаткового налаштування. Це свідчить про те, що універсального детектора немає. Багато *deepfake*-детекторів більше орієнтовані на певні характеристики генеративного методу, який використовується для створення фейків, ніж на загальні характеристики, які відрізняють фейк від реального. Прикладами таких характеристик є незначні спотворення від техніки заміни обличчя мережі GAN.[1] Це обмежує їхню ефективність у порівнянні з новими або модифікованими методами створення *deepfake*. Модель, налаштована на один сценарій, часто працює погано при застосуванні до іншого, оскільки в різних організаціях є відмінності в якості камер, освітленні, поведінці користувачів та демографічних даних.

Методи створення *deepfake* безперервно розвиваються, часто випереджаючи розробку інструментів виявлення та регуляторних рамок. Це створює постійну «гру в наздоганялки» для правоохоронців[2]. Існує значна розбіжність між публічними академічними наборами даних (наприклад, FaceForensics++ або DFDC) та реальними атаками *deepfake*, які часто використовують інше програмне забезпечення, методи стиснення та стратегії ухилення, що ускладнює ефективне узагальнення моделей.

Існують високі показники помилкових спрацьовувань (*false positives*) та пропусків (*false negatives*) у реальних сценаріях, що може призвести до хибних звинувачень або пропуску шкідливого контенту[3]. Ці обмеження засвідчують, що поточні методи є переважно швидкими та сфокусованими на відомих артефактах. Пропоновані в наукових працях рішення, такі як використання різноманітних даних для навчання, симуляція реальних умов, безперервне навчання, багатошарова безпека, свідчать про те, що майбутнє виявлення має бути активним, адаптивним і включати кілька рівнів захисту, що виходять за рамки простого аналізу контенту[4]. Це значний стратегічний поворот, який підкреслює необхідність переходу від швидкого, артефакт-орієнтованого підходу до активного та багатошарового захисту у виявленні ДПсШІ.

Правова відповідь на загрозу ДПсШІ є дуже важливою, хоча й стикається зі значними викликами через швидкість технологічного

[1] Farid H. Creating, Using, Misusing, and Detecting Deep Fakes. *Journal of Online Trust and Safety*. 2022. Vol. 1. Is. 4. DOI: https://doi.org/10.54501/jots.v1i4.56.

[2] Іванченко А. Діпфейки: розбираємося з технологією та її наслідками. *Acer Corner*. 2024. 15 січ. URL: https://blog.acer.com/ua/discussion/2686/dipfeyki-rozbirayemosya-z-tehnologiyeyu-ta-yiyi-naslidkami.

[3] False Positives and False Negatives. *GeeksforGeeks*. 2025. 14 May. URL: https://www.geeksforgeeks.org/machine-learning/false-positives-and-false-negatives/

[4] Pratt D., Huff E. A New Approach to Proactive Deepfake Detection & Remediation / Kyndryl. 2024. URL: https://www.kyndryl.com/content/dam/kyndrylprogram/doc/en/2024/proactive-deepfake-detection-remediation.pdf.

розвитку та глобальний характер проблеми. В Україні спостерігається швидка, але фрагментована реакція законодавства, що створює правові лазівки та ускладнює міжнародну співпрацю. Для цього потрібно подивитися на законодавчі новели з приводу ДПсШІ в інших країнах та міжнародних організаціях.

У Сполучених Штатах Америки спостерігається швидка реакція на проблему ДПШІ на рівні штатів. 38 штатів США вже криміналізували ДПсШІ або комп'ютерно-відредаговані матеріали. Причому більше половини цих законів були прийняті лише у 2024 році. Деякі штати, як Каліфорнія, мають дуже деталізовані закони, які вимагають (у межах ознак складу злочину) використання зображення реальної, ідентифікованої дитини, тоді як інші криміналізують будь-яке зображення, що схоже на неповнолітнього, залученим до сексуальної активності. Наприклад, Сенат Техасу одноголосно прийняв закон SB20, який набирає чинності з 01.09.2025 р. та криміналізує нове діяння – володіння, перегляд або поширення візуального матеріалу, що зображує дитину у сексуальному контексті, навіть якщо зображення створене ШІ та пов'язане з образом, що видається дитиною до 18 років, незалежно від того, чи це реальна дитина[1]. Значні варіації між штатами, де деякі закони є дуже детальними, а інші не включають комп'ютерно-генеровані зображення, створюють частковий правовий ландшафт. Кримінальна протиправність дії може залежати від конкретної юрисдикції, що є викликом для міжнародної боротьби, оскільки злочинці можуть використовувати ці лазівки, переміщуючи свою діяльність між юрисдикціями.

Велика Британія стала першою країною у світі, яка криміналізувала нові злочини, пов'язані з сексуальним насильством за допомогою ШІ. Це включає незаконне володіння, створення або розповсюдження ШІ-інструментів, спеціально розроблених для генерації CSAM, а також «посібників для педофілів», які навчають використовувати ШІ для сексуального насильства над дітьми. Крім того, надано розширені повноваження прикордонній службі для перевірки цифрових пристроїв з метою запобігання поширенню CSAM[2].

Австралія впровадила норму, що прямо забороняє «*synthetic child sexual abuse material*» (штучно згенеровані зображення чи відео з ознаками сексуального насильства над дітьми), в рамках *Online Safety Act 2021* та спеціального *Industry Code* під назвою *Protecting Children from Sexual Abuse Material Industry Code. Online Safety Act 2021*[3], що встановлює обов'язкові вимоги до платформ і провайдерів онлайн-послуг щодо виявлення та

[1] Senate Bill 20: Relating to the creation of the criminal offense of possession, promotion, or production of certain obscene visual material appearing to depict a child / Texas Senate. 2025. URL: https://legiscan.com/TX/text/SB20/id/3156303.

[2] Crime and Policing Bill: child sexual abuse material factsheet / GOV.UK. 2025. URL: https://www.gov.uk/government/publications/crime-and-policing-bill-2025-factsheets/crime-and-policing-bill-child-sexual-abuse-material-factsheet.

[3] Online Safety Act 2021 / Parliament of Australia. 2021. URL: https://www.legislation.gov.au/Details/C2021A00055.

видалення CSAM, в тому числі й *deepfake*.

У Європейському Союзі існує EU AI Act. Цей акт є знаковим кроком у регулюванні ШІ. Він визначає deepfake як ШІ-згенерований або маніпульований контент, що реалістично імітує існуючих осіб, об'єкти або події, і вимагає чіткого та помітного розкриття того, що контент є штучно згенерованим або маніпульованим. Постачальники також повинні забезпечити маркування вихідних даних ШІ-системи у машинному форматі. За недотримання передбачаються значні штрафи, що можуть сягати до 35 мільйонів євро або 7 % від загального світового річного обороту компанії[1]. Законодавство ЄС також визнає, що *deepfake* становлять більші ризики для дітей, оскільки їхні когнітивні здібності ще розвиваються, і їм складніше ідентифікувати *deepfake*.

Інтерпол активно реагує на нові загрози. Було видано «*Purple Notice*» для попередження країн-членів про використання *deepfake*-технологій для різних видів шахрайства та вимагання, включаючи онлайн-сексуальний шантаж. DevOps Group Інтерполу, фінансована Safe Online, зосереджена на розробці інструментів для протидії онлайн-сексуальній експлуатації дітей, сприяючи обміну передовим досвідом та інноваційними стратегіями між правоохоронними органами, неурядовими організаціями та академічними установами[2].

ЮНІСЕФ[3] та ЮНЕСКО[4] активно працюють над захистом прав дітей у цифрову епоху. ЮНІСЕФ визнає як потенціал, так і ризики ШІ для дітей, виступаючи за розробку політики та систем ШІ, орієнтованих на дітей. ЮНЕСКО розробляє рамки компетентності ШІ для студентів та вчителів, наголошуючи на необхідності розуміння потенціалу та ризиків ШІ для безпечної, етичної та відповідальної взаємодії з ним в освіті та за її межами. Дії ЄС, Великої Британії, Інтерполу, ЮНІСЕФ та ЮНЕСКО демонструють, що міжнародна спільнота поступово переходить до активного регулювання самої технології.

З огляду на те, наскільки швидко розвивається ШІ, законодавство має постійно слідувати технологічним досягненням, щоб запобігти будь-яким лазівкам і забезпечити ефективний захист. Це динамічний процес. Законодавчі рамки, які зараз діють, можуть застаріти завтра, вимагає створення гнучких, адаптивних правових систем, які можуть швидко реагувати на нові загрози та постійно відстежувати технологічні тенденції, можливі зловживання.

Висновки. Дитяча порнографія, створена штучним інтелектом, є

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 12 July 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*. 2024. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

[2] Purple Notices / INTERPOL. 2025. URL: https://www.interpol.int/en/How-we-work/Notices/About-Notices.

[3] Generative AI: Risks and Opportunities for Children / UNICEF Innocenti – Global Office of Research and Foresight. 2024. URL: https://www.unicef.org/innocenti/innocenti/generative-ai-risks-and-opportunities-children.

[4] Recommendation on the Ethics of Artificial Intelligence / UNESCO. 2021. URL: https://unesdoc.unesco.org/ark:/48223/pf0000377897.

серйозною та швидко еволюціонуючою загрозою, яка тягне за собою глибокі негативні психологічні, соціальні наслідки. Вона підриває безпеку дітей, сприяє злочинності та дезінформації, створює безпрецедентні виклики для правоохоронних органів через дедалі більшу невідповідність синтетичного контенту реальному. Для того, щоб протистояти цьому явищу, потрібна комплексна, багатогранна та координована глобальна стратегія. Це передбачає посилення та гармонізацію законодавства, розробку та впровадження законів, які оперативно реагують на технологічні зміни, закривають правові лазівки. Існує потреба у окремій криміналізації ДПсШІ, а також у розробка сучасних і адаптивних технологічних рішень, які включають перехід від простого аналізу артефактів до поведінкового та контекстуального аналізу, а також багатошарові механізми перевірки джерела, цілісності даних. Розробники повинні включати стандарти безпеки за замовчанням у весь життєвий цикл створення ШІ, оскільки вони несуть відповідальність за те, щоб їхні інструменти не використовували для створення шкідливого контенту.

Необхідно розширити доступ до соціальної, психологічної та юридичної допомоги, а також до послуг із видалення шкідливого контенту, щоб забезпечити тривале відновлення та реабілітацію для тих, хто постраждав від цього злочину. Забезпечуючи безпечніше цифрове майбутнє для дітей у всьому світі, лише спільні, міждисциплінарні та постійно еволюціонуючи зусилля науковців і практиків можуть ефективно протистояти цій складній і нагальній загрозі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Баловсяк Н. Швидкий контент, повільний розум? Як короткі відео викликають залежність та впливають на мозок // DSnews.ua. 2025. 22 лют. URL: https://www.dsnews.ua/ukr/society/shvidkiy-kontent-povilniy-rozum-yak-korotki-video-viklikayut-zalezhnist-ta-vplivayut-na-mozok-22022025-517271.

2. Іванченко А. Діпфейки: розбираємося з технологією та її наслідками. *Acer Corner*. 2024. 15 січ. URL: https://blog.acer.com/ua/discussion/2686/dipfeyki-rozbirayemosya-z-tehnologiyeyu-ta-yiyi-naslidkami.

3. Коваленко А. В. Метадані як джерело криміналістично значущої інформації // Використання цифрових технологій в криміналістиці і судовій експертизі : матеріали міжнар. наук.-практ. круглого столу (м. Харків, 11 груд. 2023 р.). Харків : Право, 2024. С. 82–85.

4. Хамула Д. В. До питання ознак «закамуфльованої» дитячої порнографії у фотомистецтві при мистецтвознавчому дослідженні продукції порнографічного характеру // Криміналістичні та експертні аспекти протидії злочинності : матеріали Всеукр. наук.-практ. конф. (м. Одеса, 24 листоп. 2023 р.). Одеса, 2023. С. 273–276.

5. Annual Report 2023 / National Center for Missing & Exploited Children. 2024. URL: https://www.missingkids.org/footer/about/annual-report.

6. Artificial Intelligence and Combatting Online Child Sexual Exploitation and Abuse / U.S. Department of Homeland Security. 2024. URL: https://www.dhs.gov/sites/default/files/2024-09/24_0920_k2p_genai-bulletin.pdf.

7. Binder M., Newbury E. M. H. Combatting AI-Generated CSAM / Wilson Center. 2025. 5 Feb. URL: https://www.wilsoncenter.org/article/combatting-ai-generated-csam.

8. Britain's leading the way protecting children from online predators / Home Office ; GOV.UK. 2023. 10 May. URL: https://www.gov.uk/government/news/britains-leading-the-way-protecting-children-from-online-predators.

9. Chesney R., Citron D. Deep Fakes: A Looming Challenge for Privacy, Democracy, and National. *Lawfare*. 2018. 21 Feb. URL: https://www.lawfaremedia.org/article/deep-fakes-looming-challenge-privacy-democracy-and-national-security.

10. Citron D. K. Hate Crimes in Cyberspace. Cambridge, Massachusetts : Harvard University Press, 2014. 304 p.

11. Crime and Policing Bill: child sexual abuse material factsheet / GOV.UK. 2025. URL: https://www.gov.uk/government/publications/crime-and-policing-bill-2025-factsheets/crime-and-policing-bill-child-sexual-abuse-material-factsheet.

12. Crimes against children / INTERPOL. URL: https://www.interpol.int/Crimes/Crimes-against-children

13. Europol Spotlight – Generative AI: The dark side looms over the sunny uplands / Europol Innovation Lab. October 2023. URL: https://www.europol.europa.eu/cms/sites/default/files/documents/Transcript-how-crime-is-accelerated-by-ai.pdf.

14. False Positives and False Negatives. *GeeksforGeeks*. 2025. 14 May. URL: https://www.geeksforgeeks.org/machine-learning/false-positives-and-false-negatives/

15. Farid H. Creating, Using, Misusing, and Detecting Deep Fakes. *Journal of Online Trust and Safety*. 2022. Vol. 1. Is. 4. DOI: https://doi.org/10.54501/jots.v1i4.56.

16. Finkelhor D. The Challenge of Creating a Global Indicator of Violence Against Children / Trauma, Violence, & Abuse. 2025. 11 Sept. URL: https://journals.sagepub.com/doi/10.1177/15248380251357613.

17. Generative AI: Risks and Opportunities for Children / UNICEF Innocenti – Global Office of Research and Foresight. 2024. URL: https://www.unicef.org/innocenti/innocenti/generative-ai-risks-and-opportunities-children.

18. Generative Artificial Intelligence and Cyber Bullying / Dolly's Dream. 2024. URL: https://www.dollysdream.org.au/blog/generative-artificial-intelligence-and-cyber-bullying

19. How AI is being abused to create child sexual abuse imagery / Internet Watch Foundation. URL: https://www.iwf.org.uk/about-us/why-we-exist/our-research/how-ai-is-being-abused-to-create-child-sexual-abuse-imagery/

20. How CSAM Classifiers Use AI to Build a Safer Internet / Thorn. 2025. URL: https://www.thorn.org/blog/how-thorns-csam-classifier-uses-artificial-intelligence-to-build-a-safer-internet/

21. Kobriger K. et al. Out of Our Depth with Deep Fakes: How the Law Fails Victims of Deep Fake Nonconsensual Pornography. *Richmond Journal of Law and Technology*. 2021. Vol. 28. Is. 2. URL: https://scholarship.richmond.edu/jolt/vol28/iss2/1/

22. Letourneau E. J. et al. Child Sexual Abuse and Boundary-Violating Behaviors in Youth-Serving Organizations: A Scoping Review / Trauma, Violence, & Abuse. 2024. URL: https://pure.johnshopkins.edu/en/publications/child-sexual-abuse-and-boundary-violating-behaviors-in-youth-serv.

23. Online Safety Act 2021 / Parliament of Australia. 2021. URL: https://www.legislation.gov.au/Details/C2021A00055.

24. Posetti J. et al. The Chilling: A global study of online violence against women journalists / The International Center for Journalists (ICFJ) ; UNESCO. 2022. URL: https://www.icfj.org/sites/default/files/2023-02/ICFJ%20Unesco_TheChilling_OnlineViolence.pdf.

25. Pratt D., Huff E. A New Approach to Proactive Deepfake Detection & Remediation / Kyndryl. 2024. URL: https://www.kyndryl.com/content/dam/kyndrylprogram/doc/en/2024/proactive-deepfake-detection-remediation.pdf.

26. Purple Notices / INTERPOL. 2025. URL: https://www.interpol.int/en/How-we-work/Notices/About-Notices.

27. Quayle E., Taylor M. Child pornography and the Internet: A review of the evidence. *Psychology, Crime & Law*. 2002. Vol. 8. Is. 4. P. 291–319. URL: https://www.tandfonline.com/doi/abs/10.1080/10683160208401824.

28. Recommendation on the Ethics of Artificial Intelligence / UNESCO. 2021. URL: https://unesdoc.unesco.org/ark:/48223/pf0000377897.

29. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 12 July 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*. 2024. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

30. Senate Bill 20: Relating to the creation of the criminal offense of possession, promotion, or production of certain obscene visual material appearing to depict a child / Texas Senate. 2025. URL: https://legiscan.com/TX/text/SB20/id/3156303.

31. State of the Issue: Annual Youth Survey on Online Sexual Interaction / Thorn. 2024. URL: https://www.thorn.org/research/state-of-the-issue/

Стаття надійшла до редакції 30.07.2025

Oleksii M. LYTVYNOV,
Doctor in Law, Professor,
Honored Worker of Education of Ukraine
(M. Ye. Zhukovsky National Aerospace University «Kharkiv Aviation Institute», Kharkiv, Ukraine)

Kyrylo O. CHEREVKO,
PhD in Law, Associated Professor
(Kharkiv National University of Internal Affairs, Kharkiv, Ukraine)

MODERN CHALLENGES IN DETECTING AND COUNTERING THE SPREAD OF CHILD PORNOGRAPHY CREATED WITH THE USE OF ARTIFICIAL INTELLIGENCE TOOLS

The article analyzes the growing and multifaceted threat posed by child pornography created with the use of artificial intelligence tools. The article highlights the key challenges in detecting and countering this phenomenon, as well as its profound psychological and social consequences for victims and society as a whole. It examines how the rapid development of generative artificial intelligence enables the creation of increasingly realistic images and video material, which complicates the work of law enforcement agencies and endangers the safety of children in the digital space. In addition, comprehensive approaches to combat child pornography created using artificial intelligence tools are outlined, including strengthening legislation, developing advanced technological solutions for detection and blocking, implementing ethical principles in the development of artificial intelligence, conducting extensive educational work and ensuring effective support for victims.

Keywords: *artificial intelligence, generative models, digital security, child pornography, cybercrime, content detection, image blocking, victim, child safety, international cooperation.*