УДК 339,004.8
JEL Classification: C45; D83; E65; F01; F20; L86; O33
DOI: https://doi.org/10.32983/2222-4459-2025-2-171-180

ПРОБЛЕМИ ТА ПЕРЕВАГИ ВИКОРИСТАННЯ ДІПФЕЙКІВ У СВІТОВІЙ ЕКОНОМІЦІ В ЕПОХУ ШТУЧНОГО ІНТЕЛЕКТУ

© 2025 **ПОБОЧЕНКО Л. М., ГОРОБЕЦЬ О. Г.**

УДК 339,004.8
JEL Classification: C45; D83; E65; F01; F20; L86; O33

Побоченко Л. М., Горобець О. Г. Проблеми та переваги використання діпфейків у світовій економіці в епоху штучного інтелекту

Дослідження технологій діпфейків у світовій економіці в епоху штучного інтелекту дозволило узагальнити сучасний стан їхнього розвитку, виокремити основні переваги та ризики їхнього використання, а також запропонувати нові підходи до їхньої класифікації та аналізу. У статті досліджено феномен стрімкого розвитку діпфейків в епоху штучного інтелекту. Мета даної наукової роботи полягає в комплексному аналізі технологій діпфейків, їхнього впливу на економічні, соціальні та бізнес-процеси, а також визначення викликів і можливостей, які вони створюють в умовах цифрової трансформації. Проаналізовано сучасний стан створення та використання діпфейків у світі. Обґрунтовано актуальність проблеми використання діпфейків та впливу розвитку штучного інтелекту на їхню якість та кількість. Запропоновано власну типологію діпфейків, яка більше відповідає сучасним реаліям. Так, до діпфейків пропонується віднести генерацію фото, відео, аудіо та тексту. Проведено аналіз позитивних і негативних способів застосування технологій діпфейків. Так, діпфейки можна використати з користю у маркетингу, рекламі, кіноіндустрії, освіті та ігровій сфері, які відкривають нові можливості для бізнесу та персоналізації контенту. Шкідливі наслідки, перш за все, пов'язані з використанням діпфейків з метою шахрайства, маніпуляції та дезінформації, що загрожують репутації компаній чи державі, фінансовим втратам і безпеці суспільства. Проаналізовано статистичні дані, які ілюструють стрімкий розвиток діпфейків у різних сферах, з перевагою на високоприбуткові індустрії. Наголошено на необхідності підвищення інформаційної гігієни, вивченні новітніх технологій та посиленні інтелектуального рівня суспільства для ефективного розпізнавання діпфейків та запобігання їх негативним наслідкам як для населення, так і для бізнесу та країн.

Ключові слова: *діпфейки, штучний інтелект, бізнес, інновації, технології, шахрайство, дезінформація.*

Рис.: *2.* ***Табл.:*** *3.* ***Бібл.:*** *17.*

Побоченко Леся Миколаївна *– кандидат економічних наук, доцент, завідувач кафедри міжнародних економічних відносин, бізнесу та туризму, Державний університет «Київський авіаційний інститут» (просп. Любомира Гузара, 1, Київ, 03058, Україна)*
E-mail: *lesia.pobochenko@npp.nau.edu.ua*
ORCID: *https://orcid.org/0000-0002-3094-6417*
Researcher ID: *https://www.webofscience.com/wos/author/record/32792714*
Scopus Author ID: *https://www.scopus.com/authid/detail.uri?authorId=57148516100*

Горобець Ольга Геннадіївна *– аспірант кафедри міжнародних економічних відносин, бізнесу та туризму, Державний університет «Київський авіаційний інститут» (просп. Любомира Гузара, 1, Київ, 03058, Україна)*
E-mail: *4808638@stud.nau.edu.ua*

UDC 339,004.8
JEL Classification: C45; D83; E65; F01; F20; L86; O33

Pobochenko L. M., Horobets O. H. The Challenges and Advantages of Using Deepfakes in the Global Economy in the Era of Artificial Intelligence

The research on deepfake technologies in the global economy in the era of artificial intelligence has allowed a generalization of the current state of their development, the identification of key advantages and risks associated with their use, and the proposal of new approaches to their classification and analysis. This article examines the phenomenon of the rapid development of deepfakes in the era of artificial intelligence. The aim of this scientific writing is to conduct a comprehensive analysis of deepfake technologies, their impact on economic, social, and business processes, as well as to identify the challenges and opportunities they create in the context of digital transformation. The current state of the creation and use of deepfakes worldwide has been analyzed. The relevance of the problem of deepfake usage and the influence of artificial intelligence development on their quality and quantity is substantiated. An own typology of deepfakes is proposed that better corresponds to modern realities. Thus, it is proposed to refer to deepfakes the generation of photos, videos, audio, and text. An analysis of the positive and negative ways of utilizing deepfake technologies has been conducted. Thus, deepfakes can be beneficially used in marketing, advertising, the film industry, education, and the gaming sector, which open new opportunities for businesses and content personalization. The harmful consequences are primarily associated with the use of deepfakes for the purposes of fraud, manipulation, and misinformation, which pose threats to the reputation of companies or countries, financial losses, and societal safety. Statistical data illustrating the rapid development of deepfakes in various fields, with a focus on high-profit industries, have been analyzed. The necessity of enhancing information hygiene, studying cutting-edge technologies, and strengthening the intellectual level of society for the effective recognition of deepfakes and preventing their negative consequences for both the population and businesses and countries is emphasized.

Keywords: *deepfakes, artificial intelligence, business, innovation, technology, fraud, misinformation.*

Fig.: *2.* ***Tabl.:*** *3.* ***Bibl.:*** *17.*

Pobochenko Lesia M. *– Candidate of Sciences (Economics), Associate Professor, Head of the Department of International Economic Relations, Business and Tourism, State University "Kyiv Aviation Institute" (1 Liubomyra Husara Ave., Kyiv, 03058, Ukraine)*

E-mail: *lesia.pobochenko@npp.nau.edu.ua*
ORCID: *https://orcid.org/0000-0002-3094-6417*
Researcher ID: *https://www.webofscience.com/wos/author/record/32792714*
Scopus Author ID: *https://www.scopus.com/authid/detail.uri?authorId=57148516100*

Horobets Olha H. *– Postgraduate Student, Department of International Economic Relations, Business and Tourism, State University "Kyiv Aviation Institute" (1 Liubomyra Husara Ave., Kyiv, 03058, Ukraine)*
E-mail: *4808638@stud.nau.edu.ua*

На сьогоднішній день термін «діпфейк» набирає все більшу популярність. Технологія діпфейк почала змінювати правила гри в економіці, бізнесі та медіа. Практично кожна людина може створити надреалістичний контент за допомогою штучного інтелекту, що створює нові можливості в сферах маркетингу, розваг, комунікацій та інших. Проте також дана технологія стала інструментом для шахрайства та маніпуляцій.

Уже на сьогоднішній день фінансові махінації, репутаційні атаки та політичні дезінформації несуть реальні збитки компаніям, країнам і громадянам. Проте правильне регулювання та стратегічне використання діпфейків можуть зробити з них потужний інструмент для розвитку бізнесу та персоналізованої взаємодії з клієнтами.

Дане дослідження спрямоване на комплексний аналіз впливу діпфейків на світову економіку та бізнес, визначення їхніх ризиків та потенційних переваг, а також на розробку рекомендацій щодо їхнього ефективного й безпечного використання в епоху штучного інтелекту.

Актуальність дослідження полягає в тому, що діпфейки зараз створюються та поширюються надзвичайно швидко, і технології їх генерації постійно покращуються, проте не створена система регулювання такого контенту, яка могла б захистити права людей, а також запобігти негативним наслідкам для населення, бізнесу та країн.

Останнім часом проблема діпфейків привертає все більше уваги науковців. Дослідники вивчають їхній вплив на різні сфери, зокрема економіку, фінанси, маркетинг, політику та інформаційну безпеку. У своїх роботах вони аналізують не лише потенційні загрози, а й позитивні можливості, які відкривають ці технології. Так науковці, як Бенжіо, Рамперсад, Альтіябі, Хао Лі, Мусак, Салмінен, Мянтюмякі, Рахман та Двіведі, Гамаж, Гхасія, Бонагірі, Вітінг і Сасахара зробили вагомий вклад у наукові напрацювання щодо діпфейк-технологій. Науковці сходяться на думці, що поширення діпфейк-технологій вимагає впровадження нових підходів до їхнього регулювання, а також розвитку більш ефективних механізмів розпізнавання підробленого контенту.

Проте багато питань досі лишаються не закритими. Наприклад, немає єдиної типології діпфейків, також етичні та правові аспекти використання діпфейків досі залишаються не до кінця вивченими. Також, у зв'язку з постійним удосконаленням моделей штучного інтелекту та технологій створення діпфейків, зростає актуальність у розробці ефективних методів їхнього виявлення. Крім того, на думку авторів, існує необхідність додаткового висвітлення впливу діпфейків на інформаційну безпеку та довіру до медіа.

Мета даної наукової роботи полягає у комплексному аналізі технологій діпфейків, їхнього впливу на економічні, соціальні та бізнес-процеси, а також визначення викликів і можливостей, які вони створюють в умовах цифрової трансформації.

У сучасних умовах розвиток технологій штучного інтелекту (ШІ) та машинного навчання спричинив появу нових інструментів для створення реалістичних зображень, аудіо та відео, відомих як діпфейки. Генеративний штучний інтелект додає новий вимір до проблеми дезінформації. Безкоштовно доступні та здебільшого нерегульовані інструменти дозволяють будь-кому створювати неправдиву інформацію та фейковий вміст у величезних кількостях. Серед них – імітація голосів реальних людей, створення фотографій і відео, які неможливо відрізнити від справжніх.

Діпфейк (дипфейк), поєднання «глибокого навчання» та «підробки», – це технологія, яка використовує штучний інтелект і глибоке навчання для створення реалістичного фейкового контенту, особливо у відео.

Термін «діпфейк» виник у 2017 році, ознаменувавши початок нової ери в медіа-маніпуляціях. Згідно зі статею «Brand Vision Insights», технологія діпфейк переважно використовується у двох формах: діпфейки, які замінюють або створюють віртуальні обличчя (deepfaces), та діпфейки, які змінюють або імітують голоси у відео (deepvoices). Ці досягнення в синтезі, керованому ШІ, відкрили двері до безпрецедентних можливостей у світі створення медіа [1].

На думку авторів, на сьогодні, межа між типами розмивається, оскільки зображення та голос можна створювати одночасно або поєднувати в одне. Крім того, до діпфейків можна віднести

створення будь-якого роду зображення, не тільки обличчя (наприклад, неіснуюча подія). Також до діпфейків варто віднести написання тексту, який описує неіснуючі події чи спотворює реальність. Дане твердження обґрунтоване тим, що нові моделі, такі як o1-preview від OpenAI, здатні генерувати текст рівня науковця, який буде досить важко розпізнати як неправдивий. Існують різні типології та види діпфейків, проте, на даний момент, жоден з них не відображає повноцінну картину.

Таким чином, авторами була розроблена власна типологія діпфейків, яка відображена в *табл. 1.*

Таблиця 1

Види діпфейків

Фото	**Обличчя чи особа**
	Подія/загальна картина; більше ніж одна людина або взагалі без людей
Аудіо	Пряма трансляція/генерація у реальному часі
	Запис
Відео	Повністю згенерована неіснуюча людина
	Імітація існуючої людини (генерація анімованого зображення та голосу)
	Накладення ефекту однієї людини на іншу (наприклад, прямий ефір, коли одна людина говорить, жестикулює, а на відео зображена зовсім інша людина) або видозміна зовнішності людини (наприклад, омолодження чи зміна кольору волосся)
	Генерація події чи групи людей (наприклад, реклама, кліп, фільм)
Текст	Розрахований на велику групу осіб, країну чи групу країн, чи світ загалом
	Розрахований на одну людину чи маленьку групу осіб (висока персоналізація)

Джерело: складено авторами (власна розробка).

Різні сфери бізнесу, так само як і різні верстви населення світу, вже піддалися впливу діпфейк-технологій, а кількість таких випадків продовжує зростати. З одного боку, діпфейки відкривають нові можливості для бізнесу, наприклад у маркетингу та рекламі. Компанії можуть використовувати діпфейки для створення високоякісного контенту з мінімальними витратами, що дозволяє швидко реагувати на зміни на ринку та ефективніше комунікувати зі споживачами. Технології діпфейків можуть бути використані для персоналізації рекламних повідомлень або створення інтерактивного контенту, який залучає аудиторію.

Компанії можуть персоналізувати свої рекламні кампанії та маркетинг інтерактивним способом, пропонуючи свої послуги через налаштовані діпфейк-аватари, які беруть участь у відеоконференціях і презентаціях, а також кастомеризовані відео в емейл-кампаніях. Це дає змогу роздрібним торговцям демонструвати продукти більш креативно та економно, наприклад за допомогою віртуальних покупок.

Контент, створений штучним інтелектом, створює нові можливості для бізнесу. Він може персоналізувати його для своїх клієнтів інтерактивним способом, пропонуючи свої послуги через налаштовані діпфейк-аватари, які беруть участь у відеоконференціях і презентаціях. Це дає змогу менеджерам з продажу демонструвати продукти більш креативно та економно, наприклад за допомогою віртуальних покупок [3].

У кіноіндустрії діпфейки відкривають нові можливості для створення спецефектів, відтворення акторів, які померли, чи для зміни зовнішності персонажів. Наприклад, діпфейк-технологія була використана під час створення фільму «Бунтар-Один. Зоряні Війни: Історія», щоб повернути персонажів, які з'являлися у фільмі 39-річної давнини такими, якими вони були в початковому фільмі [2].

Також діпфейки можуть бути використані для створення інтерактивних навчальних матеріалів. Наприклад, для підвищення зацікавленості учнів чи студентів до історії видатні постаті можуть «оживати» та розповідати про події минулого, що підвищує як зацікавленість до занять, так і полегшує засвоєння матеріалу, оскільки школярі та інші охочі можуть «поспілкуватися» з історичними особистостями.

Технологія «діпфейк» також широко використовується для покращення графіки для відеоігор. Ця технологія є важливою, для створення аватарів в ігровій індустрії. За допомогою штучного інтелекту можна створювати аватари, які синхронізують міміку та рухи. Крім того,

синтетичне мовлення, згенероване з використанням діпфейк-технології, може зробити голос гравців схожим на голос героїв відомих фільмів чи вигаданих героїв [3].

Проте діпфейк-технології мають і негативну сторону. Діпфейки можуть бути використані для шахрайства, маніпуляцій та дезінформації, що ставить під загрозу репутацію бізнесів. Наприклад, фальшиві відео або аудіо можуть спровокувати занепокоєння, втрату довіри клієнтів, підірвати репутацію бренду, старіння технологій тощо. Це, своєю чергою, може призвести до значних фінансових втрат.

Зростаюча хвиля шахрайства з використанням діпфейк-технологій нанесла мільйони доларів збитків компаніям по всьому світу. Фахівці з кібербезпеки зазначають, що ситуація може стати ще гіршою, у зв'язку з удосконаленням моделей ШІ, які використовують для шахрайства [4].

Протягом останніх кількох років зростання створення та розповсюдження діпфейків було експоненційним, причому кількість діпфейків подвоювалася кожні шість місяців.

Згідно з даними DeepMedia, у 2023 р. приблизно 500 тисяч відео та голосових діпфейків були опубліковані в соціальних мережах по всьому світу. До 2025 р. можна очікувати, що в інтернеті поширюватимуться 8 мільйонів діпфейків на рік, що відповідає їх подвоєнню кожні шість місяців. Легкість доступу до потужних інструментів ШІ та велика кількість загальнодоступних даних сприяють поширенню діпфейків [5].

Швидке поширення діпфейків у соціальних мережах погіршує і без того поширену проблему дезінформації. Дослідження, проведене Університетом Балтимора та компанією з кібербезпеки CHEQ, показало, що у 2020 р. фейкові новини коштували світовій економіці 78 мільярдів доларів США [5].

Згідно з даними McKinsey за 2022 р., збитки від кіберзлочинності до 2024 р. досягнуть 5 трильйонів доларів США. Серед основних видів атак виділяють шахрайство та фальшиву маніпуляцію особистими даними за допомогою технології діпфейк [6].

Так, у 2019 р. генеральний директор британського постачальника електроенергії перерахував 220 тис. євро (238 тис. дол. США) шахраю. Зловмисник використав технологію діпфейку для цифрової імітації керівника материнської компанії [4].

У резонансній справі в січні 2020 р. кіберзлочинці пограбували банк Об'єднаних Арабських Еміратів на суму понад 35 млн дол. США за допомогою голосової технології діпфейку. Технологія була використана для імітації директора компанії, який був знайомий керівника банку. Менеджер дозволив транзакції [7].

У серпні 2023 р. дослідники компанії Mandiant, що займається кібербезпекою та належить компанії Google, задокументували низку випадків використання зловмисниками штучного інтелекту та технології діпфейку для фішингу, дезінформації та інших протиправних цілей. У той час компанія заявила, що використання технології для такої діяльності було обмеженим, але зростаюча доступність нових генеративних інструментів ШІ прискорить її впровадження шахраями [4].

Іншим прикладом стало нещодавнє пограбування на суму 25 млн дол. США, яке стало відоме у лютому 2024 р. Фінансист міжнародної компанії з Гонконгу перерахував шахраям 25 млн дол. США, оскільки він був упевнений, що робить це на запит свого директора, з яким, вважав, що комунікував по відеодзвінку, оскільки шахраї використали діпфейк-технологію для відео [8].

Також китайські державні ЗМІ цього ж року повідомили про подібний випадок у провінції Шаньсі, пов'язаний із фінансовою співробітницею, яку обманом змусили перевести 1,86 мільйона юанів (262 тис. дол. США) на рахунок шахрая після відеодзвінка з діпфейком її боса [4].

Щодо фізичних осіб, то діпфейки можуть бути використані для шантажу, наклепу, переслідувань, крадіжки особистих даних, залякування та порнографії з помсти [9]. Крім того, розмови про діпфейки на таких платформах, як Reddit, вказують на наявність спільноти, яка підтримує створення діпфейків та обмін ними, що викликає стурбованість суспільства, так як кожен може стати наступною жертвою [10].

У звіті Всесвітнього економічного форуму (ВЕФ) про глобальні ризики за 2024 р. місінформація (несвідомі помилки та подання неточних або перекручених фактів) та дезінформація (неправдива, оманлива, маніпулятивна інформація, створена навмисне заради економічних, політичних або інших вигод) є основною загрозою, з якою світ зіткнеться в найближчі два роки [11], (*табл. 2*).

Очікується, що протягом наступних двох років близько трьох мільярдів людей візьмуть участь у виборах у кількох країнах, включно з Бангладеш, Індією, Індонезією, Мексикою, Пакистаном, Сполученим Королівством і Сполученими Штатами. Тому широке використання місінформації, дезінформації та інструментів для їх поширення можуть підірвати легітимність новообраних урядів. Заворушення можуть варіюватися від жорстоких протестів і злочинів на

Таблиця 2

Глобальні ризики, ранжовані за ступенем серйозності в короткостроковій та довгостроковій перспективі

№ з/п	Прогноз на два роки	Прогноз на 10 років
1	Дезінформація та місінформація	Екстремальні погодні явища
2	Екстремальні погодні явища	Критичні зміни в екосистемах Землі
3	Соціальна поляризація	Втрата біорізноманіття та колапс екосистем
4	Кібернезахищеність	Дефіцит природних ресурсів
5	Міждержавний збройний конфлікт	Дезінформація та місінформація
6	Відсутність економічних можливостей	Небажані наслідки розвитку штучного інтелекту
7	Інфляція	Примусова міграція
8	Примусова міграція	Кібернезахищеність
9	Економічний спад	Соціальна поляризація
10	Забруднення	Забруднення

Джерело: складено авторами за даними [11, p. 8].

ґрунті ненависті до громадянського протистояння та тероризму.

Поза виборами сприйняття реальності також може стати більш поляризованим, проникнувши в публічний дискурс з питань, починаючи від охорони здоров'я та закінчуючи соціальною справедливістю. Разом з підривом правди ризик внутрішньої пропаганди та цензури також зростатиме.

У відповідь на місінформацію та дезінформацію уряди можуть отримати більше можливостей контролювати інформацію на основі того, що вони вважають «правдивим». Свободи, пов'язані з інтернетом, пресою та доступом до ширших джерел інформації, які вже скорочуються, ризикують перетворитися на ширші репресії щодо інформаційних потоків у більшій кількості країн [11].

За результатами дослідження компанією «Sumsub» за період з першого кварталу 2023 р. до 1 кварталу 2024 р. було встановлено, що країни з найбільшою кількістю виявлених діпфейків (у першому кварталі 2024 р.) – це Китай, Іспанія, Німеччина, Україна, США, В'єтнам і Велика Британія (*табл. 3*).

Крім того, варто зауважити, що спостерігається помітне зростання інцидентів із діпфейком порівняно з минулим роком у країнах, де вибори заплановані на 2024 рік: Індія (280%), США (303%), Південна Африка (500%), Мексика (500%), Молдова (900%), Індонезія (1550%) і Південна Корея (1625%).

У ЄС (де вибори до Європейського парламенту заплановані на червень) багато країн зіткнулися зі зростанням випадків діпфейків порівняно з минулим роком: це Болгарія (3000%), Португалія (1700%), Бельгія (800%), Іспанія (191%), Німеччина (142%) і Франція (97%).

Навіть у країнах, де виборів у 2024 р. не прогнозується, шахрайство з використанням діпфейк-технологій поширюється безпрецедентними темпами порівняно з минулим роком: Китай (2800%), Туреччина (1533%), Сінгапур (1100%), Гонконг (1000%), Бразилія (822%), В'єтнам (541%), Україна (394%) та Японія (243%) [12].

Хоча шахрайство із застосуванням штучного інтелекту зросло в більшості країн, у деяких країнах, які проводять вибори у 2024 р., кількість інцидентів із використанням діпфейків зменшилася порівняно з минулим роком: Велика Британія (–10%), Хорватія (–33%), Ірландія (–40%) та Литва (–44%) [12].

З розвитком генеративного штучного інтелекту обсяг дезінформації потенційно стає нескінченним, що робить перевірку фактів недостатнім інструментом. Оскільки граничні витрати на виробництво дезінформації падають до нуля, витрати на поширення – також майже нульові, завдяки соціальним медіа [11]. У 2023 р. вартість придбання готових діпфейків варіювалася від 300 до 20 тис. дол. США за хвилину залежно від складності, якості та популярності особи, яку імітували [13].

У дослідженні World Economic Forum (WEF), у якому було обрано п'ять ризиків, які найбільш ймовірно створять глобальну кризу у 2024 р., місінформація та дезінформація, створені ШІ, набрали 53% і посіли друге місце в рейтингу. Дана аналітика відображає реальну ситуацію, що склалася у зв'язку з поширенням діпфейків з метою дезінформації [11] (*рис. 1*).

Таблиця 3

Приріст кількості діпфейків у 2024 виборчому році (перший квартал 2023 р. – перший квартал 2024 р.)

Група за наявністю виборів у 2024 році	Країна	Приріст діпфейків порівняно з попереднім періодом
Зростання діпфейків у країнах з виборами у 2024 р.	Південна Корея	+1625%
	Індонезія	+1550%
	Молдова	+900%
	Південна Африка	+500%
	Мексика	+500%
	США	+303%
	Індія	+280%
	Бангладеш	+30%
Зростання діпфейків у країнах Європейського Союзу, де вибори до Європейського парламенту заплановані на червень 2024 р.	Болгарія	+3000%
	Португалія	+1700%
	Бельгія	+800%
	Іспанія	+191%
	Німеччина	+142%
	Франція	+97%
Зростання діпфейків у країнах, де не заплановано вибори на 2024 р.	Китай	+2800%
	Туреччина	+1533%
	Сінгапур	+1100%
	Гонконг	+1000%
	Бразилія	+822%
	В'єтнам	+541%
	Україна	+394%
	Японія	+243%
Зменшення кількості діпфейків у країнах з виборами у 2024 р.	Литва	–44%
	Ірландія	–40%
	Хорватія	–33%
	Велика Британія	–10%

Джерело: складено авторами за даними [12].

Таким чином, населення занепокоєне щодо великої загрози провокацій та спотворення думки з метою налаштування суспільства проти чогось чи когось або схилення на користь чогось чи когось. На думку авторів, дезінформація може нести дуже болючі наслідки включно із провокуванням збройних конфліктів на глобальній арені, які можуть перерости навіть у Третю світову війну. Тому варто ретельно відфільтровувати споживану інформацію та не ігнорувати цифрову гігієну.

Варто відмітити, що зараз можливе використання зручних програм для легкого та швидкого створення складного та переконливого контенту з використанням ШІ, такого як діпфейкові відео та голосові підробки, для створення яких раніше потрібні були цілі команди технічно підкованих людей. Така демократизація технології діпфейку знижує бар'єр для створення та поширення неправдивих наративів і оманливого контенту в інтернеті [11].

Шахраї легко можуть використовувати чатботи для поширення неправди в інтернеті з рекордною швидкістю, незалежно від мови. Чат-боти з перетворенням промпту в текст (такі як ChatGPT та Gemini), або генератори зображень (такі як Midjourney, DALL-E або Stable Diffusion) можна використовувати для створення величезних обсягів тексту, а також високореалістичних фейкових аудіо, зображень і відео для поширення місінформації та дезінформації. Це може призвести до неправдивих наративів, дезінформації щодо конкретної країни, маніпулювання громадською думкою та

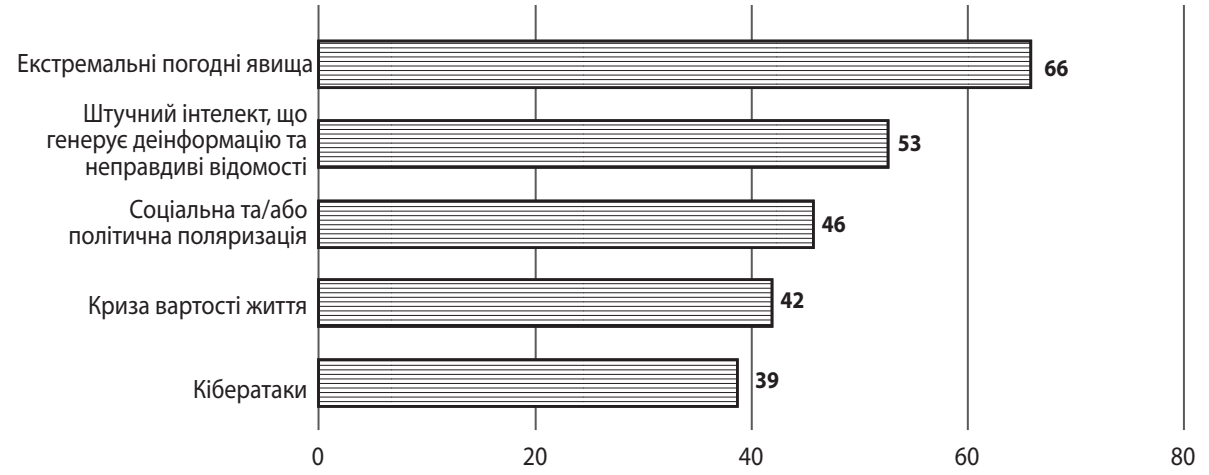

Рис. 1. Поточний ландшафт ризиків, 2024 р., (%)

Джерело: побудовано авторами за даними [11, p. 8].

навіть завдати шкоди окремим особам чи організаціям [11].

У дослідженні 2023 р. науковці Цюріхського університету в Швейцарії виявили, що генеративний ШІ може створювати точну інформацію, яку легше зрозуміти, але він також може створювати досить переконливу дезінформацію. Учасники також не змогли розрізнити дописи на X (колишній Twitter), написані GPT-3 (на даний момент вже далеко не найпотужніша модель) і написані реальними людьми [11].

Програми на основі ШІ можна поєднувати для автоматизації всього процесу створення та розповсюдження діпфейків. Повністю синтетичний візуальний матеріал можна створювати з текстової підказки, а веб-сайти можна програмувати автоматично [11]. Наприклад, DeepFaceLab використовується для понад 95% усіх відеопідробок. Програма, доступна у вигляді відкритого вихідного коду на GitHub, використовує штучні нейронні мережі для копіювання візуальних і звукових функцій з оригінального відео на цільове [14].

За допомогою правильного швидкого налаштування кожен може створити реалістичні зображення або змусити голосом видатних постатей говорити все, що йому заманеться. Хоча створення діпфейків саме по собі не є кримінальним злочином, багато урядів, тим не менш, рухаються до жорсткішого регулювання використання штучного інтелекту, щоб запобігти шкоді залученим сторонам.

Крім основного напряму діпфейків, створення порнографічного вмісту без згоди (за участю переважно жінок-знаменитостей) також може використовуватися для шахрайства з особистими даними шляхом виготовлення підроблених фото чи відео або видавання себе за інших по телефону. Згідно з графіком, заснованим на останньому щорічному звіті Sumsub, кількість випадків шахрайства, пов'язаних із діпфейками, стрімко зросла між 2022 і 2023 роками в багатьох країнах світу [15] (*рис. 2*).

Наприклад, кількість спроб шахрайства на Філіппінах зросла на 4500% за рік, за ними йдуть такі країни, як В'єтнам, Сполучені Штати та Бельгія [16]. Варто зауважити, що очікується, що у США збитки від шахрайства, спричинені генеративними технологіями штучного інтелекту, зростуть до 40 млрд дол. США до 2027 р. Цей прогноз відображає загальний річний темп зростання у США на 32% з 12,3 млрд дол. США у 2023 р. [14].

У 2023 р. кількість виявлених діпфейків у всіх галузях промисловості зросла в 10 разів у світі. Криптовалюта стала головною цільовою галуззю для шахрайства з використанням діпфейків. Так, на криптосферу припадає 88% усіх випадків діпфейк-шахрайств, виявлених у 2023 р. Ця статистика підкреслює особливу вразливість індустрії криптовалют до сучасних методів шахрайства, ймовірно, через її цифрову природу та потенційно високі фінансові ставки. Також у 2023 р. кількість інцидентів діпфейків у фінтех зросла на 700%, що також підтверджує факт вразливості галузей, де задіяні великі гроші [14].

За останніми даними, кількість випадків використання технологій діпфейків зросла на 1520% у сфері iGaming у період з першого кварталу 2023 р. до першого кварталу 2024 р. Така динаміка є показовою для тенденцій в інших галузях, зокрема на ринках електронної комерції, фінтеху та криптовалют, де зростання також де-

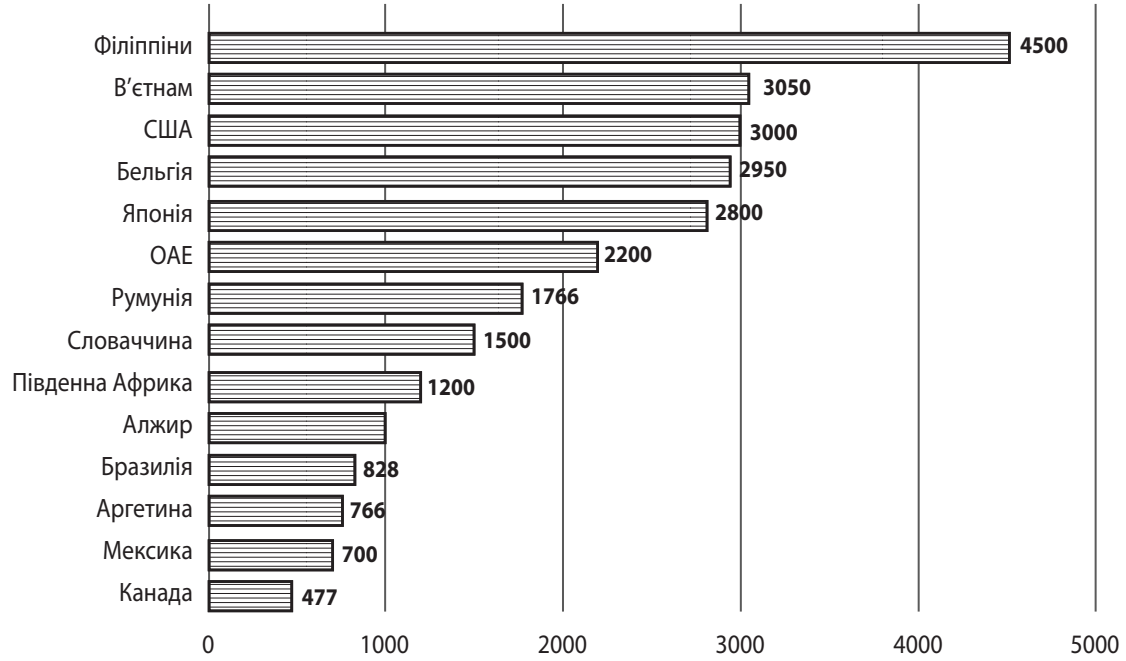

Рис. 2. Зростання випадків шахрайства з використанням діпфейк-технологій у 2022–2023 рр., (%)

Джерело: побудовано авторами за даними [15].

монструє значні показники – на 900%, 533% і 217% відповідно за даний період [12].

Аналіз ринків свідчить про збільшення кількості випадків використання діпфейків не тільки в B2B і B2C сегментах, а й у більш широкому цифровому середовищі, що стає загрозою для інформаційної безпеки. Технологічний прогрес у сфері штучного інтелекту несе із собою нові виклики у вигляді поширення дезінформації, шахрайських схем і фальсифікацій, які ставлять під загрозу стабільність фінансових ринків та кібербезпеку суспільства в цілому [12].

Дійсно, діпфейки несуть загрозу та занепокоєння як бізнесу, так і громадськості. Наприклад, у 2023 р. 60% американців висловили значну стурбованість діпфейками, більше, ніж будь-яким іншим ризиком, пов'язаним зі штучним інтелектом. Також, згідно з опитуванням, у 2024 р. 65% американців висловлюють занепокоєння щодо можливих порушень конфіденційності через технології ШІ. Це побоювання відображає зростаюче занепокоєння щодо здатності ШІ використовувати особисті дані та порушувати права на конфіденційність. Технології синтетичного контенту, такі як діпфейки, зменшують довіру суспільства до автентичності медіа та надійності використання ШІ [14].

За допомогою штучного інтелекту зловмисники можуть здійснювати масштабне шахрайство, націлюючись на кількох жертв одночасно, використовуючи схожі підходи. Лише у 2022 р. ФБР нарахувало 21 832 випадки шахрайства з корпоративною електронною поштою зі збитками приблизно у 2,7 млрд дол. США. За оцінками Центру фінансових послуг Deloitte, збитки від шахрайства на електронній пошті за допомогою штучного інтелекту можуть скласти близько 11,5 млрд дол. США до 2027 р. за сценарієм «агресивного» впровадження [17].

ВИСНОВКИ

Отже, використання діпфейків і розвиток самої технології швидко розвивається. Основна причина такого стрімкого вдосконалення та розповсюдження є успіхи в розробці моделей ШІ та програм на його основі, можливість його використання безкоштовно або за мінімальну суму, а також його недостатньо контрольоване використання.

Діпфейки мають як переваги, так і недоліки. До *основних переваг* можна віднести: персоналізацію, автоматизацію, інтерактивність, скорочення витрат і створення нових можливостей. До *основних недоліків* варто віднести: розвиток кібератак і шахрайства, розповсюдження місінформації та дезінформації, підрив довіри суспільства, загострення чи розв'язання конфліктів, підрив репутації людини, країни чи компанії, суперечливість права, фінансові збитки.

Оскільки можливості штучного інтелекту продовжують збільшуватися, про що свідчать такі продукти, як відеогенератор Sora від компанії

OpenAI та нова версія ChatGPT – o1-preview від тієї ж компанії, шахрайство з використанням діпфейків продовжуватиме зростати. Тому важливо виробляти інформаційну гігієну в суспільстві, вивчати новітні технології та підвищувати інтелектуальний рівень суспільства загалом з метою розпізнавання діпфейків та уникнення провокацій та злочинів.

Наукова новизна дослідження полягає в розробці нової типології діпфейків, яка охоплює не лише відео та аудіо, а й текстові та графічні фальсифікації, що відповідає сучасним реаліям розвитку штучного інтелекту. Здійснено комплексний аналіз економічного впливу діпфейків, включаючи їхнє застосування в бізнесі, рекламі та кіноіндустрії, а також пов'язані з ними загрози, такі як шахрайство та маніпуляція громадською думкою. У ході дослідження також було виявлено взаємозв'язок між швидким розвитком генеративного штучного інтелекту та зростанням масштабів діпфейк-шахрайств, що підтверджує необхідність посилення цифрової гігієни та механізмів регулювання. ■

БІБЛІОГРАФІЯ

1. The Rise of Deepfake Marketing – What Are the Cons and Pros? // *BrandVM*. November 22, 2024. URL: https://www.brandvm.com/post/deepfake-marketing
2. Sarkar S. Rogue One filmmakers explain how they digitally recreated two characters. *Polygon*. December 27, 2016. URL: https://www.polygon.com/2016/12/27/14092060/rogue-one-star-wars-grand-moff-tarkin-princess-leia
3. Milenković A. The Positive Impact of Deepfakes. *BlueGridio*. August 13, 2024. URL: https://bluegrid.io/blog/the-positive-impact-of-deepfakes/
4. Butts D. Deepfake scams have robbed companies of millions. Experts warn it could get worse. *CNBC*. May 27, 2024. URL: https://www.cnbc.com/2024/05/28/deepfake-scams-have-looted-millions-experts-warn-it-could-get-worse.html
5. Jacobson N. Deepfakes and Their Impact on Society. *OpenFox*. February 26, 2024. URL: https://www.openfox.com/deepfakes-and-their-impact-on-society/
6. What is Deepfake, Its Losses, and How to Secure Data from Deepfake Attacks. *VIDA*. April 05, 2024. URL: https://vida.id/en/blog/what-is-deepfake-its-losses-and-how-to-secure-data-from-deepfake-attacks
7. Kshetri N. The Economics of Deepfakes. *IEEE Computer Society*, 2023. Vol. 56. P. 89–94. DOI: 10.1109/MC.2023.3276068
8. Chen H., Magramo K. Finance worker pays out $25 million after video call with deepfake 'chief financial officer'. *CNN*. February 4, 2024. URL: https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/
9. Mustak M., Salminen J., Mäntymäki M. at al. Deepfakes: Deceptions, mitigations, and opportunities. *Journal of Business Research*. 2023. Vol. 154. DOI: https://doi.org/10.1016/j.jbusres.2022.113368
10. Gamage D., Ghasiya P., Bonagiri V. at al. Are Deepfakes Concerning? Analyzing Conversations of Deepfakes on Reddit and Exploring Societal Implications. *CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. DOI: https://doi.org/10.1145/3491102.3517446
11. The Global Risks Report 2024. *World Economic Forum*. 19th Edition. URL: https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf
12. Deepfake Cases Surge in Countries Holding 2024 Elections, Sumsub Research Shows. *Sumsub*. URL: https://sumsub.com/newsroom/deepfake-cases-surge-in-countries-holding-2024-elections-sumsub-research-shows/
13. Simonchik K. Deepfakes in 2024 – a Summary of Trends in KYC. *LinkedIn*. 16.01.2024. URL: https://www.linkedin.com/pulse/deepfakes-2024-summary-trends-kyc-konstantin-simonchik-s25ae
14. Chipeta C. Deepfake statistics (2024): 25 new facts for CFOs. *Eftsure*. 12 July 2024. URL: https://eftsure.com/statistics/deepfake-statistics/#source-wrapper
15. Zandt F. How Dangerous are Deepfakes and Other AI-Powered Fraud? *statista*. Mar 13, 2024. URL: https://www.statista.com/chart/31901/countries-per-region-with-biggest-increases-in-deepfake-specific-fraud-cases/
16. Sumsub Research: Global Deepfake Incidents Surge Tenfold from 2022 to 2023. *Sumsub*. URL: https://sumsub.com/newsroom/sumsub-research-global-deepfake-incidents-surge-tenfold-from-2022-to-2023/
17. Lalchand S. et al. Generative AI is expected to magnify the risk of deepfakes and other fraud in banking. *Deloitte*. May 29, 2024. URL: https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk-on-the-rise.html

REFERENCES

Butts, D. "Deepfake scams have robbed companies of millions. Experts warn it could get worse". CNBC. May 27, 2024. https://www.cnbc.com/2024/05/28/deepfake-scams-have-looted-millions-experts-warn-it-could-get-worse.html

Chen, H., and Magramo, K. "Finance worker pays out $25 million after video call with deepfake 'chief financial officer'". CNN. February 4, 2024. https://edition.cnn.com/2024/02/04/asia/deepfake-cfo-scam-hong-kong-intl-hnk/

Chipeta, C. "Deepfake statistics (2024): 25 new facts for CFOs". Eftsure. July 12, 2024. https://eftsure.com/statistics/deepfake-statistics/#source-wrapper

"Deepfake Cases Surge in Countries Holding 2024 Elections, Sumsub Research Shows". Sumsub. https://sumsub.com/newsroom/deepfake-cases-surge-in-countries-holding-2024-elections-sumsub-research-shows/

Gamage, D. "Are Deepfakes Concerning? Analyzing Conversations of Deepfakes on Reddit and Exploring Societal Implications". *CHI '22: Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*.
DOI: https://doi.org/10.1145/3491102.3517446

Jacobson, N. "Deepfakes and Their Impact on Society". OpenFox. February 26, 2024. https://www.openfox.com/deepfakes-and-their-impact-on-society/

Kshetri, N. "The Economics of Deepfakes". *IEEE Computer Society*, vol. 56 (2023): 89-94.
DOI: 10.1109/MC.2023.3276068

Lalchand, S. et al. "Generative AI is expected to magnify the risk of deepfakes and other fraud in banking". Deloitte. May 29, 2024. https://www2.deloitte.com/us/en/insights/industry/financial-services/financial-services-industry-predictions/2024/deepfake-banking-fraud-risk-on-the-rise.html

Milenkovic, A. "The Positive Impact of Deepfakes". BlueGridio. August 13, 2024. https://bluegrid.io/blog/the-positive-impact-of-deepfakes/

Mustak, M. et al. "Deepfakes: Deceptions, mitigations, and opportunities". *Journal of Business Research*, vol. 154 (2023).
DOI: https://doi.org/10.1016/j.jbusres.2022.113368

"Sumsub Research: Global Deepfake Incidents Surge Tenfold from 2022 to 2023". Sumsub. https://sumsub.com/newsroom/sumsub-research-global-deepfake-incidents-surge-tenfold-from-2022-to-2023/

Sarkar, S. "Rogue One filmmakers explain how they digitally recreated two characters". Polygon. December 27, 2016. https://www.polygon.com/2016/12/27/14092060/rogue-one-star-wars-grand-moff-tarkin-princess-leia

Simonchik, K. "Deepfakes in 2024 - a Summary of Trends in KYC". LinkedIn. January 16, 2024. https://www.linkedin.com/pulse/deepfakes-2024-summary-trends-kyc-konstantin-simonchik-s25ae

"The Global Risks Report 2024". World Economic Forum. https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf

"The Rise of Deepfake Marketing - What Are the Cons and Pros?" BrandVM. November 22, 2024. https://www.brandvm.com/post/deepfake-marketing

"What is Deepfake, Its Losses, and How to Secure Data from Deepfake Attacks". VIDA. April 05, 2024. https://vida.id/en/blog/what-is-deepfake-its-losses-and-how-to-secure-data-from-deepfake-attacks

Zandt, F. "How Dangerous are Deepfakes and Other AI-Powered Fraud?" statista. March 13, 2024. https://www.statista.com/chart/31901/countries-per-region-with-biggest-increases-in-deepfake-specific-fraud-cases/