

SECTION 7. MODERN MEDIA CULTURE

DOI <https://doi.org/10.30525/978-9934-26-488-7-12>

THE PECULIARITIES OF USING DEEPFAKE IN THE INTERNET CULTURE

ОСОБЛИВОСТІ ЗАСТОСУВАННЯ ДІПФЕЙКУ В СЕРЕДОВИЩІ ІНТЕРНЕТ-КУЛЬТУРИ

Okhrymovych V. O.

*Postgraduate Student
National Academy of Culture and Arts
Management
Kyiv, Ukraine*

Охримович В. О.

*аспірантка
Національна академія керівних кадрів
культури і мистецтв
м. Київ, Україна*

Штучний інтелект (ШІ) та його інструменти активно інтегруються в медіакультурний простір, відкриваючи нові можливості для створення різних форм креативного вираження. Разом з тим, використання інструментів та нових технологій ШІ в практиках Інтернет-культури призводить до виникнення етико-комунікативних, соціокультурних проблем, появи питань автентичності культурного продукту, що обумовлює необхідність дослідження подібних інновацій.

При цьому особливі перестороги викликає технологія створення дїпфейків. Дїпфейки (англ. «Deepfake» від deep learning + fake «підробка») – це сформовані ШІ цифрові фейки, здебільшого, підроблені фото та відео, які характеризуються настільки високим ступенем реалістичності, що їх важко відрізнити від справжніх. Технологія дїпфейку ґрунтується на тому, що спеціальні алгоритми комп'ютерних програм (наприклад, DeepFaceLab – 95% ринку, Reface, Zao, FaceApp, GauGAN) використовують інформацію з різних джерел (фотографії, відео), аналізуючи вихідні дані – найдрібніші фізіологічні особливості якоїсь людини (наприклад, її обличчя у різних ракурсах та освітленні). Синхронізуючи міміку, рухи, жести тощо, програма об'єднує все це в один матеріал, створюючи т.зв. «маску», що накладається на фото або відео іншої людини. У такий спосіб з різних реальних даних у результаті утворюється несправжнє – фейк.

Як зазначають фахівці з маркетингу, політологи, психологи, культурологи, відеокреатив, а саме дїпфейки, сприяють ефективнішому досягненню високого рівня емоційної залученості (наприклад,

потенційного споживача товару чи послуги) та взаємодії з контентом. Відтак, легше викликати емоції, які стимулюють сприйняття і запам'ятовування інформації. Відео дозволяє комбінувати зображення, звук і текст, що створює багатошарову комунікацію, яку складніше отримати через текстові або статичні зображення.

Розглянемо як інструменти ШІ, зокрема, діпфейки змінили медіа. Якщо звернути увагу на соціальні медіа, то очевидно, що діпфейки стали важливим елементом сучасної культури мемів, які активно поширюються у таких соціальних мережах, як TikTok, YouTube, Instagram тощо. Вірусні діпфейк-відео на популярні фільми та серіали, а також відео-фрагменти, що імітують виступи знаменитостей або відомих політиків, стали популярними інструментами для створення пародій та сатиричних роликів. Останнє явище ілюструє, як технології можуть змінювати сприйняття реальності та впливати на масову свідомість. Окрім того, діпфейки часто використовуються як розважальний контент. Наприклад, нейромережа Viggle представила новий інструмент ШІ, за допомогою якого користувач може змусити будь-кого грати в футбол, танцювати та ходити подіумом.

Вказані тенденції демонструють здатність діпфейків створювати нові форми мережевого контенту, водночас, піднімаючи питання про довіру до контенту, відповідальність та етику їх використання в медіакультурних практиках. На сьогодні деяку пересторогу викликає стимулювання довіри до цієї технології в Інтернет-просторі. В 2019 році у передмові до Європейської Білої книги зі штучного інтелекту (AI) було задекларовано наступне твердження: «Оскільки цифрові технології стають усе більш важливою частиною кожного аспекту життя, люди повинні мати можливість довіряти їм. Достовірність також є умовою його реалізації [3, с. 4]. Водночас, згідно з дослідженням 2020 року, опублікованим в журналі Crime Science, ризики застосування діпфейку вважалися такими, що несуть найбільшу загрозу.

Висвітлюючи цю проблематику варто згадати цільовий фішинг. Фішинг (від англ. «phishing» – видобування) – це метод шахрайства в Інтернеті, метою якого є збір конфіденційної інформації або встановлення зловмисного програмного забезпечення за допомогою цифрового повідомлення, яке нібито відправлено довіреною особою. При цільовому фішингу – т.зв. «діпфейк фішингу» – зловмисник використовує існуючу довіру до іншої особи, щоб переконати користувача виконати певні дії (йдеться про розкриття паролів або перехід за сумнівними посиланнями). Щодо подібного використання діпфейк-технологій, то очевидно, що вони виходять за межі правового поля, коли актуалізується фактор «прихованої довіри людей до засобів

масової інформації, включаючи видачу себе за дітей старих батьків за допомогою відео-дзвінків для отримання доступу до засобів; використання телефону для запиту доступу до захищених систем; фальшиві відео публічних діячів, які говорять або здійснюють негідні дії з метою маніпулювання підтримкою [1, с. 64]. Тож, як слушно зауважує Г. Андрощук «можливості, пропоновані deepfake, безмежні» [там само].

Крім того, експерти з політтехнологій зазначають, що дідфейки розповсюджуються задля втручань у вибори, підсилення політичної напруги, маніпулювання громадською думкою тощо. В США одним з найвідоміших прикладів використання дідфейку вважають «промову» Барака Обами в 2018 році. Йдеться про виступ цифрового аватара колишнього президента США, що стосувався світових загроз, які можуть бути викликані фейковими новинами та дезінформацією в Інтернеті. Згадаємо також останній приклад маніпулювання громадською думкою та дискредитацію опонетки, коли Дональд Трамп звинуватив Камалу Харріс у використанні ШІ задля підробки чисельності присутніх людей під час транслявання її мітингу у Мічигані. Примітно також, що під час повномасштабного вторгнення росії на територію України одним з інструментів країни агресорки є використання інструментів ШІ як «зброї», задля ведення пропагандистської роботи, а саме для поширення фейкових новин. Одним з очікуваних дідфейків мав стати виступ Президента України про капітуляцію українських військ.

З огляду на це очевидно, що сьогодні виник ще один тип загрози, який пов'язаний з впровадженням дідфейків в життя людей – т. зв. «дивіденд брехуна». Це поняття описує ситуацію, коли у світі дідфейків людина, яка говорить неправду, може використовувати ці технології задля того, щоб дискредитувати правдиві докази, називаючи їх підробкою. І якщо в Україні тільки зараз у 2024 році з'явився прецедент судової справи щодо захисту авторських прав стосовно використання ШІ, то в Європі вже існують Стратегії та документи, які узагальнюють основні ризики та виклики, пов'язані із застосуванням інструментів ШІ.

До прикладу, для протидії дезінформації у 2020 р. було розроблено Європейську цифрову стратегію (Europe's Digital Decade). Вона передбачає, що до 2030 року 80% європейських громадян повинні володіти, принаймні, базовим рівнем цифрових навичок, що сприятиме розвитку цифрової грамотності та розпізнавання дезінформації. Безпосередньо, в ЄС 25 серпня 2023 року набув чинності важливий законодавчий акт – Digital Services Act (DSA). Однією з його цілей є зменшення системних ризиків та дезінформації, а однією з ключових особливостей DSA є модерація контенту. В акті вказується, що

«постачальники повинні приділяти особливу увагу тому, як їхні послуги використовуються для поширення або розширення оманливого вмісту, включаючи дезінформацію. Цей акт демонструє намір ЄС створити більш безпечний цифровий простір, в якому захищені основні права всіх користувачів цифрових послуг» [див.: 2].

У підсумку зазначимо, що застосування дідфейків у середовищі Інтернет-культури, у т.ч. соціально-комунікативних платформах, впливає на різноманітні аспекти існування сучасного суспільства. Використовуючи досягнення штучного інтелекту у цифровій обробці відео та зображень, дідфейки демонструють, як технології можуть впливати на усталені соціокультурні норми, політичні процеси та, взагалі, на сприйняття реальності. Технологія дідфейк змінює наше уявлення про креативність, автентичність культурного продукту, ідентичність та цінності, а також трансформує норми та моделі поведінки, актуалізуючи нові форми комунікативних взаємодій в медійному просторі Інтернет-культури.

Відтак, дідфейки стають не лише засобом творчого самовираження, але й інструментом маніпуляцій в медіа, порушуючи питання правди і брехні. Це викликає низку питань щодо відповідальності за достовірність інформації та регулювання способів використання інструментів ШІ в динамічному і відкритому середовищі Інтернет-культури. Адже неконтрольоване поширення дідфейків впливає на суспільну довіру до медіа, влади та міжособистісних стосунків, що може призвести до зростання скептицизму та зневіри в соціальних і культурних інститутах.

Ми вважаємо, що використання інструментів ШІ повинно бути прозорими та маркуватися. Закон повинен захищати права митців перед медіа-гігантами, програми з виявлення дідфейків повинні бути в безкоштовному доступі користувачів та розроблені заздалегідь. На нашу думку, оскільки наразі рівень медіаграмотності суспільства у розрізненні автентичних і фальшивих матеріалів в медіа «не встигає» за розвитком дідфейк-технологій, потрібно призупинити розробку інструментів ШІ, які створюють хаос у всіх сферах життя, та спочатку презентувати якісні програми з розпізнавання дідфейків, створивши таким чином ефективні допоміжні інструменти, які б стали для людства опорою в оцінці реальності.

Література:

1. Андрощук Г. О. Штучний інтелект: економіка, інтелектуальна власність, загрози. *Інтелектуальна власність в Україні*. 2021. № 2. С. 56–74.

2. Петрів О. Дезінформація та штучний інтелект: (не)видима загроза сучасності. Центр демократії та верховенства права (CEDEM). 2023. URL: <https://cedem.org.ua/analytics/dezinformatsiya-shtuchnyi-intelekt/> (дата звернення: 27.09.2024).

3. SingularityNET: A Decentralized, Open Market and Inter-Network for AIs. *Whitepaper 2.0*. URL: https://media.abnnewswire.net/media/en/docs/93410-AGI_whitepaper.pdf (дата звернення: 25.09.2024).