Пехник Алевтіна Валентинівна
Кройтор Артем Вікторович
Бахметьєв Андрій Євгенійович

Deepfake та аудіо-підробки: нова ера політичної дезінформації

УДК 316.77
DOI https://doi.org/10.24195/2414-9616.2026-1.21

Пехник Алевтіна Валентинівна
кандидат політичних наук, доцент,
завідувач кафедри журналістики,
реклами та медіакомунікацій
Одеського національного університету
імені І. І. Мечникова
Французький бульвар, 24/26,
Одеса, Україна
ORCID: 0000-0003-2534-7652

Кройтор Артем Вікторович
кандидат політичних наук, доцент,
в.о. директора
Відокремленого структурного
підрозділу «Одеський фаховий коледж
комп'ютерних технологій Одеського
національного університету імені
І. І. Мечникова»
вул. Корольова академіка, 5/2,
Одеса, Україна
ORCID: 0000-0003-4652-7441

Бахметьєв Андрій Євгенійович
кандидат політичних наук,
старший викладач кафедри
журналістики, реклами
та медіакомунікацій
Одеського національного університету
імені І. І. Мечникова
Французький бульвар, 24/26,
Одеса, Україна
ORCID: 0000-0002-8509-0003

Активний розвиток генеративних технологій штучного інтелекту зумовив появу нових форм політичної дезінформації, зокрема deepfake-відео та аудіо-підробок, що суттєво трансформують механізми політичної комунікації в цифровому середовищі. Надреалістичні аудіовізуальні фальсифікації створюють принципово нові виклики для демократичних інститутів, оскільки підривають довіру до медіаконтенту як джерела верифікованої інформації та ставлять під сумнів сам інститут візуального доказу.

Метою статті є комплексний аналіз deepfake та аудіо-підробок як інструментів політичної дезінформації та оцінка їхнього впливу на функціонування політичної комунікації і демократичні процеси. Методологічна основа дослідження базується на поєднанні загальнонаукових і спеціальних методів. Застосовано бібліографічний аналіз для окреслення теоретичних підходів до вивчення цифрової дезінформації; методи аналізу, синтезу та узагальнення – для систематизації основних форм і механізмів використання deepfake у політичній практиці; структурно-функціональний підхід – для визначення їхньої ролі в системі політичної комунікації; порівняльний метод – для зіставлення випадків застосування таких технологій у різних політичних контекстах; контент-аналіз та елементи кейс-стаді – для дослідження конкретних прикладів цифрових підробок і реакцій на них.

У статті розкрито специфіку deepfake як високотехнологічного інструменту інформаційно-психологічного впливу, здатного інтегруватися у стратегії політичної боротьби та гібридного протистояння. Показано, що ключова загроза полягає не лише у створенні окремих фальсифікованих матеріалів, а у формуванні середовища тотального сумніву, в якому будь-який аудіовізуальний контент може бути поставлений під питання. Така трансформація інформаційного простору сприяє поширенню логіки постправди, посиленню поляризації та зниженню рівня раціонального політичного вибору. Проаналізовано вплив deepfake-технологій на електоральні процеси, репутаційні ризики політичних акторів та стійкість державних інститутів.

У підсумку обґрунтовано необхідність формування комплексної системи протидії, що передбачає розвиток механізмів превентивної комунікації, удосконалення процедур кризового реагування, підвищення стандартів журналістської верифікації та запровадження нормативного регулювання у сфері маркування контенту, створеного із застосуванням штучного інтелекту.

Ключові слова: *deepfake, аудіо-підробки, політична дезінформація, штучний інтелект, політична комунікація, гібридна війна, постправда, інформаційна безпека.*

Вступ. Сучасний інформаційний простір перебуває у стані кризи довіри, спричиненої стрімким розвитком генеративних технологій штучного інтелекту. Поява та доступність інструментів Deepfake та Voice Cloning створили безпрецедентну загрозу для політичної комунікації, електоральних процесів та національної безпеки. Якщо раніше фальсифікація вимагала значних ресурсів і залишала помітні сліди, то сьогодні створення надреалістичного відео чи аудіо, що імітує політичних лідерів, доступне широкому колу суб'єктів, включно з ворожими державними та недержавними акторами. Як зазначають О.Піцун Н. Мельник та Х. Ліп'яніна-Гончаренко «зростання популярності генеративного інтелекту також породжує серйозні виклики, зокрема у вигляді діпфейків – відео та зображень, які маніпулюють реальністю, зображуючи людей у неправдивому контексті» [4]. Проблема полягає не лише у виявленні фальшивок, а й у ерозії інституту довіри до автентичного контенту, що переводить політичний дискурс у сферу пост-правди та вічного сумніву. Це вимагає негайного переосмислення методів фактчекінгу, етичних норм журналістики та стратегій кризового PR у політиці.

Аналіз останніх досліджень і публікацій. Проблематика Deepfake та його політичних наслідків формується на перетині політології, теорії комунікацій, інформаційної безпеки та права. Класичні концепції масової психології Г. Лебона та теорії еліт В. Парето набувають нового змісту в умовах

цифрової маніпуляції, коли технології дозволяють відтворювати переконливі, але фальсифіковані образи політичних акторів. Сучасні дослідження розглядають Deepfake не лише як інструмент поширення дезінформації, а як фактор системної кризи довіри до аудіовізуальних доказів і масових комунікацій загалом.

Іванова, Єфремова та Чулков (2025) аналізують Deepfake як феномен соціальних і масових комунікацій, підкреслюючи його амбівалентність – поєднання технологічних можливостей і суттєвих ризиків для формування громадської думки. Охримович (2024) звертає увагу на специфіку функціонування Deepfake в інтернет-культурі, де алгоритмічні механізми поширення посилюють його вплив. Вальорска (2020) акцентує на прикладному вимірі проблеми, наголошуючи на необхідності вироблення практичних інструментів підвищення медіастійкості.

Правовий аспект також посідає важливе місце в наукових дискусіях. Юртаєва (2021) окреслює межу криміналізації використання Deepfake, визначаючи критерії, за яких цифрова підробка набуває ознак правопорушення. Hachkevych (2025) пропонує концептуальні підходи до розмежування шкідливого та нейтрального контенту, що є принциповим для збереження балансу між протидією дезінформації та свободою слова.

Окремий напрям досліджень пов'язаний із технічними механізмами верифікації. Піцун, Мельник та Ліп'яніна-Гончаренко (2024) розробляють гібридні методи виявлення Deepfake на основі аналізу біометричних характеристик обличчя, що свідчить про перехід до більш точних алгоритмічних моделей детекції.

У ширшому теоретичному контексті Deepfake розглядається як інструмент інформаційно-психологічних операцій та складова гібридної війни, що поглиблює феномен постправди. Його функціонування корелює з концепціями agenda-setting та фреймінгу, оскільки технологія дозволяє не лише поширювати вигадані повідомлення, а й конструювати порядок денний і емоційне тло сприйняття подій. Попри наявність міжнародних ініціатив щодо маркування ШІ-контенту та розвитку програм автентифікації, темпи практичного впровадження регуляторних і технічних рішень поки що не відповідають швидкості розвитку генеративних технологій, що зумовлює необхідність подальшого комплексного аналізу цієї проблематики.

Метою даної **статті** є аналіз впливу технологій Deepfake та аудіо-підробок на політичний процес та інститути демократії, визначення політологічних ризиків у контексті гібридної війни та обґрунтування необхідності впровадження стратегічних протоколів протидії на рівні медіа, PR-команд та держави.

Методи дослідження. У роботі використано метод бібліографічного аналізу наукових джерел, аналітичних звітів і нормативно-правових актів (для визначення теоретичних підходів до вивчення цифрової дезінформації, генеративного ШІ та інформаційної безпеки); аналіз, синтез і узагальнення (для систематизації форм, механізмів та інструментів застосування deepfake і аудіо-підробок у політичній комунікації); структурно-функціональний метод (для визначення ролі технологій синтетичного контенту у системі політичної взаємодії та їх впливу на інституційну стійкість демократії); порівняльний метод (для зіставлення практик використання deepfake у різних політичних та національних контекстах); контент-аналіз (для дослідження прикладів маніпулятивних аудіовізуальних матеріалів у медіапросторі та соціальних мережах); елементи кейс-стаді (для аналізу конкретних випадків застосування технологій цифрової підробки під час виборчих кампаній і кризових політичних ситуацій); прогностичний метод (для оцінки потенційних ризиків подальшого розвитку генеративних технологій для демократичних процесів).

Результати. Як слушно підкреслює М. Вальорска, політично вмотивована дезінформація не є новим феноменом, однак саме сучасний технологічний розвиток суттєво спростив і пришвидшив процес створення та поширення маніпулятивного контенту. Дослідниця наголошує, що завдяки алгоритмам штучного інтелекту відеоматеріали сьогодні можна фальсифікувати швидко й відносно недорого (йдеться про діпфейки), причому без необхідності володіння спеціалізованими знаннями [1].

К. Юртаєва зауважує, що витоки технології Deepfake сягають 1990-х років, проте тоді відповідні інструменти були доступні переважно професіоналам зі створення спецефектів у кіноіндустрії. Суттєвий прорив відбувся у 2014 році після розробки алгоритму машинного навчання – генеративної змагальної мережі (GAN), створеної студентом Стенфордського університету Яном Ґудфеллоу спільно з колегами. Принцип роботи GAN полягає у взаємодії двох нейронних мереж: одна генерує штучні зразки, тоді як інша (дискримінативна мережа D) у процесі своєрідної «антагоністичної гри» намагається відрізнити справжній контент від підробленого, що сприяє постійному вдосконаленню якості фальсифікації [5].

Сам термін Deepfake утворений поєднанням понять «deep learning» (глибинне навчання) та «fake» (підробка). Йдеться про технологію створення надзвичайно реалістичних відео-, аудіо- чи візуальних матеріалів, у яких обличчя, голос або міміка однієї особи замінюються характеристиками іншої, або ж формується повністю новий, але правдоподібний образ чи звуковий ряд. В основі процесу – нейронні мережі, зокрема GAN, які навчаються на великих обсягах даних: одна частина системи створює фейковий матеріал, інша – здійснює перевірку та «критику», підвищуючи рівень правдоподібності результату.

Т. Іванова, О. Єфремова та А. Чулков зазначають, що Deepfake здатний адаптувати контент до потреб конкретної аудиторії, зокрема формуючи динамічні відеоновини на основі текстових матеріалів. Як приклад наводиться ШІ-модель BloombergGPT, призначена для аналізу фінансової інформації, яка потенційно може автоматично генерувати відеоновини, що значно пришвидшує виробництво медіаматеріалів і дозволяє персоналізувати інформацію для окремого глядача [2]. А. Гадкевич виділяє дві ключові складові цього інструменту – технологічну та інтелектуальну – і підкреслює, що до діпфейків доцільно відносити не лише зображення та відео, а й аудіозаписи, які дедалі частіше використовуються у практиці маніпуляцій [6].

Основна загроза полягає не тільки у створенні неправдивих матеріалів, а й у підриві довіри до автентичного контенту як такого. Політичні актори отримують можливість заперечувати навіть справжні викривальні матеріали, посилаючись на те, що це може бути Deepfake. Технологія дає змогу фабрикувати відео із нібито провокаційними заявами, фактами корупції чи неетичною поведінкою політиків. Навіть за умови швидкого спростування такий контент здатен завдати суттєвої репутаційної шкоди, особливо в період виборчих кампаній. Аудіопідробки, попри меншу помітність, становлять ще більшу небезпеку, адже їх легше створити й складніше перевірити; вони ефективні у кризових комунікаціях (імітація термінових переговорів, наказів чи офіційних заяв).

Таким чином, Deepfake не лише продукує нові фейки, а й руйнує саму довіру до візуальних доказів. Головний політологічний ризик полягає у можливості повного заперечення об'єктивної реальності під приводом потенційної фальсифікації. Це формує ситуацію перманентного сумніву, в якій встановлення факту стає проблематичним. Технологія посилює феномен постправди, де емоційні реакції та ідеологічні настанови домінують над перевіреними даними, а політична конкуренція трансформується у боротьбу за сприйняття та конструювання симулякрів.

Deepfake активно інтегрується в інструментарій інформаційно-психологічних операцій та гібридної війни. Поширення фальшивих заяв від імені державних лідерів (наприклад, щодо капітуляції, зміни зовнішньополітичного курсу чи запровадження надзвичайного стану) може бути використане для провокування паніки, дестабілізації фінансових ринків або дискредитації військового керівництва. Зовнішні актори можуть застосовувати цю технологію для впливу на електоральну поведінку – зокрема, поширюючи фальшиві зізнання кандидата напередодні голосування, коли оперативне спростування вже неможливе. Окрім цього, Deepfake здатен загострювати соціальні, етнічні чи ідеологічні конфлікти шляхом створення вигаданих заяв від імені окремих груп.

Перед демократичними державами постає складне завдання: виробити ефективні механізми правового регулювання, не порушуючи свободи слова. Більшість національних законодавств, включно з українським, поки що не містять чіткого визначення Deepfake та спеціалізованої відповідальності за його створення і поширення. Норми щодо наклепу чи образи часто виявляються неефективними через стрімкість розповсюдження контенту. У межах політологічного дискурсу порушується питання про підвищення відповідальності великих технологічних компаній (Meta, Google, X) за моніторинг і своєчасне видалення таких матеріалів, що водночас породжує дебати щодо співвідношення цензури та безпеки. Також обговорюється доцільність запровадження обов'язкового маркування контенту, створеного або модифікованого за допомогою штучного інтелекту.

Безперечно, медіа та політичний PR відіграють ключову роль у системі протидії Deepfake. Відтак доцільно розглянути базові стратегії реагування на такі загрози. Передусім ідеться про стратегію превентивного PR (Prebunking), спрямовану на формування своєрідного когнітивного імунітету аудиторії. Попереджувальне спростування є інструментом управління очікуваннями та мінімізації ефекту раптовості й когнітивного шоку. Теоретичним підґрунтям цього підходу є теорія щеплення В. Мак-Гвайра, згідно з якою стійкість до переконання формується через попереднє ознайомлення індивіда зі «слабкою дозою» потенційно деструктивних аргументів у поєднанні з контраргументами. У практиці політичного PR це означає не просто інформування про існування Deepfake, а системне пояснення механізмів їх створення та логіки маніпуляції. Водночас важливо послідовно формувати наративи автентичності й доброчесності політика, щоб потенційна фальшивка суперечила вже усталеному образу. Якщо ж ідентифіковано можливого ворожого актора, попереднє позначення його як джерела дезінформації істотно знижує довіру до майбутніх інформаційних вкидів.

Іншою важливою моделлю є стратегія миттєвого реагування, що функціонує в межах кризового комунікаційного менеджменту, адаптованого до надшвидкого цифрового середовища. У цьому контексті застосовується логіка ситуаційної теорії кризових комунікацій, де основна мета полягає в оперативному зменшенні репутаційних втрат. У випадку Deepfake це потребує переходу від пасивної реакції до проактивних дій. Спростування доцільно здійснювати у форматі прямого ефіру або якісного відеозвернення з чіткими маркерами автентичності (актуальна дата, час, згадка про поточні події, унікальні деталі). Таким чином створюється новий підтверджений факт, що конкурує з фальшивкою. Критично важливим є дотримання принципу «золотої години» – реагування в межах 30–60 хвилин для запобігання віральному поши-

 ISSN 2414-9616 (Print), ISSN 2664-6013 (Online)

ренню. Паралельно має запускатися юридична процедура фіксації правопорушення (звернення до правоохоронних органів і цифрових платформ), що посилює легітимність публічного спростування.

Окремим стратегічним напрямом виступає медіаграмотність як складова національної безпеки. Її трансформація з освітньої категорії у площину державної стратегічної комунікації означає визнання захисту когнітивного простору елементом суверенітету. У контексті ІПСО Deepfake постає як інструмент зовнішнього впливу, а отже держава має інвестувати не лише в освітні програми, а й у спеціалізовані дослідницькі центри, що займаються моніторингом технологій, розробкою інструментів верифікації та впровадженням критичного мислення на всіх рівнях освіти. На рівні публічної політики доцільним є законодавче регулювання за моделлю ЄС (Digital Services Act), яке зобов'язує технологічні платформи (Meta, Google) вкладати ресурси у виявлення та маркування контенту, створеного ШІ. На необхідності прозорого використання технологій штучного інтелекту та відповідного нормативного врегулювання наголошує й В. Охримович [3].

Для ефективної протидії політичній дезінформації, створеній за допомогою Deepfake, фактчекери мають застосовувати комплекс технічних методів перевірки:

1. Аналіз метаданих (Metadata) – перевірка дати, часу, пристрою та місця зйомки. Навіть з огляду на можливість фальсифікації метаданих, їхня відсутність або невідповідність є вагомим сигналом ризику.

2. Аналіз освітлення – виявлення дисонансу між освітленням обличчя та фоном, невідповідності тіней і джерел світла, що може свідчити про монтаж.

3. Аналіз міміки та рухів – оцінка природності моргання, рухів очей і м'язів обличчя; неприродна циклічність або порушена частота Blink Rate часто вказують на втручання нейромережі.

4. Аналіз пікселів і артефактів – пошук дефектів стиснення, спотворень по контуру обличчя, збоїв у деталях (волосся, зуби), характерних для роботи GAN-мереж.

5. Спеціалізоване програмне забезпечення – використання професійних інструментів верифікації (зокрема ініціатив на кшталт Adobe Content Authenticity Initiative та університетських розробок), які аналізують математичні патерни втручання ШІ.

Поряд із технічними аспектами виникає широкий спектр етичних викликів. Проблема виходить за межі індивідуальної репутації та стосується легітимності політичних інститутів. Демократія передбачає раціональний вибір на основі достовірної інформації, тоді як Deepfake підриває сам інститут візуального доказу. Етична відповідальність медіа й PR-фахівців, особливо тих, хто працює в інтересах держави, полягає у захисті суспільного блага (Public Good), а не в максимізації сенсаційності.

Так званий «ефект Deepfake» проявляється в ситуації, коли політик може безпідставно оголосити справжній компромат фальшивкою. Це зміщує фокус відповідальності: медіа мають не лише викривати підробки, а й доводити автентичність реальних матеріалів, що ускладнює журналістську практику.

Для журналістів ключовим принципом стає утримання від публікації сумнівного контенту до завершення повної верифікації. Навіть матеріал із застереженням може стати каталізатором його поширення. Етично відповідальне спростування передбачає детальне пояснення механізмів виявлення підробки (аналіз освітлення, рухів, артефактів), перетворюючи його на інструмент підвищення медіаграмотності. Оптимальною стратегією є мінімізація експозиції фейку та оприлюднення спростування лише після технічного підтвердження.

Політичні PR-фахівці, працюючи з кандидатами чи урядовими структурами, зобов'язані уникати маніпулятивних практик і не вдаватися до «дзеркальних» технологій у відповідь на атаку. Використання власних Deepfake як контрінструменту суперечить етичним стандартам. Натомість пріоритетом має бути прозорість і негайна публічна поява політика у прямому ефірі для підтвердження автентичності. Така стратегія демонструє відкритість, впевненість і готовність протидіяти дезінформації в публічному полі.

В цілому, можна зробити висновок, що технологія Deepfake знаменує новий етап дезінформації та глибоку політологічну кризу легітимності та довіри. Як показано у статті, головна загроза полягає не стільки у створенні окремих фальшивок, скільки у ерозії інституту "візуального доказу", що є фундаментом раціонального вибору та демократичного дискурсу. Це переводить політичну боротьбу у сферу пост-правди та симулякрів.

Політологічні наслідки є стратегічними – це зовнішній вплив та внутрішня деструкція. Deepfake стає високотехнологічною зброєю гібридної війни та ІПСО, здатною до маніпуляції електоральною поведінкою та дестабілізації державних інститутів у критичні моменти. В свою чергу, технологія підживлює вічний сумнів («Ефект Deepfake»), дозволяючи політичним акторам заперечувати автентичний компромат та ухилятися від відповідальності.

Протидія має бути комплексною та багаторівневою, охоплюючи технологічні, правові та етичні площини:

• На державному рівні: необхідна законодавча ініціатива щодо обов'язкового маркування контенту, створеного ШІ, та посилення відповідальності технологічних платформ (Meta, Google, X). Медіаграмотність повинна бути визнана елементом національної безпеки, що вимагає системного державного фінансування для формування когнітивного імунітету населення.

• На рівні медіа та PR: Професійне співтовариство має керуватися етичним стандартом і не публікувати сумнівний контент до його повної технічної верифікації, оскільки публікація лише збільшує його віральність. Журналісти повинні застосовувати стратегію прозорості, перетворюючи спростування на освітній контент шляхом чіткого пояснення виявлених артефактів та невідповідностей.

• На рівні комунікаційних стратегій: Політичні PR-команди повинні впроваджувати стратегію Prebunking (теорія щеплення), а також протокол миттєвого реагування, діючи в межах «золотої години» (30–60 хвилин) для відновлення автентичності політика через публічні виступи у прямому ефірі.

Таким чином, в умовах стрімкого поширення технологій Deepfake етика комунікації трансформується з професійного стандарту у стратегічний ресурс забезпечення демократичної стабільності та національної безпеки. Дотримання принципів достовірності, прозорості та відповідальності стає не лише питанням репутації окремого медіа чи політичного актора, а чинником збереження інституційної довіри та легітимності політичної системи загалом.

Deepfake підриває базову передумову демократичного процесу – можливість громадян ухвалювати раціональні рішення на основі перевіреної інформації. У ситуації, коли межа між автентичним і сфабрикованим контентом розмивається, ключовим завданням медіа, журналістів і PR-фахівців стає не лише спростування фейків, а системне формування культури критичного сприйняття інформації. Саме тому етична комунікація має розглядатися як елемент стратегічних комунікацій держави, спрямований на зміцнення когнітивної стійкості суспільства.

Ігнорування професійних стандартів, поширення неперевірених матеріалів або свідоме використання маніпулятивних технологій є не просто помилкою – це фактор, що об'єктивно послаблює демократичні інститути та може бути використаний зовнішніми або внутрішніми деструктивними акторами. У цьому сенсі безвідповідальна інформаційна поведінка фактично сприяє руйнуванню суспільної довіри, яка є фундаментом політичної стабільності.

Отже, формування системи протидії Deepfake повинно поєднувати технічні інструменти верифікації, правове регулювання, стратегічний PR та високі етичні стандарти професійної діяльності. Лише комплексний підхід дозволить зберегти цілісність інформаційного простору та забезпечити стійкість демократичних процесів у цифрову епоху.

ЛІТЕРАТУРА:

1. Вальорска М. Діпфейк та дезінформація : практичний посібник / пер. з нім. В. Олійника. Київ : Академія української преси ; Центр Вільної Преси, 2020. 36 с. [in Ukrainian].

2. Іванова Т., Єфремова О., Чулков А. Deepfake як технологія сфери соціальних та масових комунікацій: можливості та загрози. *Наукові записки Інституту журналістики.* 2025. № 86 (1). С. 21–34. https://doi.org/10.17721/2522-1272.2025.86.2 (дата звернення: 24.02.2026). [in Ukrainian].

3. Охримович В. Діпфейк у комунікативному просторі інтернет-культури: особливості застосування. *Художня культура. Актуальні проблеми.* 2024. № 20 (2). С. 118–125. https://doi.org/10.31500/1992-5514.20(2).2024.318261 (дата звернення: 24.02.2026). [in Ukrainian].

4. Піцун О., Мельник Н., Ліп'яніна-Гончаренко Х. Гібридний підхід до визначення діпфейків на основі людського обличчя. *Herald of Khmelnytskyi National University. Technical Sciences*. 2024. № 341 (5). С. 371–375. https://doi.org/10.31891/2307-5732-2024-341-5-55 (дата звернення: 24.02.2026). [in Ukrainian].

5. Юртаєва К. Кримінологічний аналіз використання технології deepfake: коли фейк стає злочином. *Вісник кримінологічної асоціації України*. 2021. № 1 (24). С. 31–42. [in Ukrainian].

6. Hachkevych A. Deepfakes: Definition of the Concept and Criteria for Distinguishing Between Harmful and Harmless Deepfakes. *Veritas: Legal and Psychological-Pedagogical Research.* 2025. Vol. 1 (2). P. 12–20. https://doi.org/10.23939/veritas2025.02.012 (дата звернення: 24.02.2026).

REFERENCES:

1. Valjorska M. (2020) Deepfake ta dezinformatsiya: praktychnyi posibnyk [Deepfake and disinformation: a practical guide]. Kyiv: Academy of the Ukrainian Press; Center for the Free Press. (in Ukrainian)

2. Ivanova T., Yefremova O., Chulkov A. (2025) Deepfake yak tekhnolohiya sfery sotsialnykh i masovykh komunikatsii: mozhlyvosti ta zahrozy [Deepfake as a technology of the sphere of social and mass communications: opportunities and threats]. Naukovi zapysky Instytutu zhurnalistyky [Scientific notes of the Institute of Journalism], vol. 86, no. 1, pp. 21–34. https://doi.org/10.17721/2522-1272.2025.86.2

3. Okhrimovych V. (2024) Deepfake v komunikatyvnomu prostori internet-kultury: osoblyvosti zastosuvannya [Deepfake in the communicative space of Internet culture: features of application]. Mystetska kultura. Aktualni problemy [Art culture. Current problems], vol. 20, no. 2, pp. 118–125. https://doi.org/10.31500/1992-5514.20(2).2024.318261

4. Pitsun O., Melnyk N., Lipyanina-Honcharenko H. (2024) Hibrydnyi pidkhid do identyfikatsii deepfake na osnovi oblychchya lyudyny [Hybrid approach to identifying deepfake based on human face]. Visnyk Khmelnytskoho natsionalnoho universytetu. Tekhnichni nauky [Herald of Khmelnytskyi National University. Technical Sciences], no. 341, vol. 5, pp. 371–375. https://doi.org/10.31891/2307-5732-2024-341-5-55

5. Yurtaeva K. (2021) Kryminolohichnyi analiz vykorystannya tekhnolohii deepfake: koly feik staye zlochynom [Criminological analysis of the use of deepfake technology: when a fake becomes a crime]. Visnyk Kry-

minolohichnoi asotsiatsii Ukrainy [Bulletin of the Criminological Association of Ukraine], no. 1 (24), pp. 31–42.

6. Hachkevych A. (2025) Deepfakes: vyznachennya ponyattya ta kryterii rozmezhuvannya shkidlyvykh i neshkidlyvykh deepfake [Deepfakes: Definition of the concept and criteria for distinguishing between harmful and harmless deepfakes]. Veritas: pravovi ta psykholoho-pedahohichni doslidzhennya [Veritas: Legal and Psychological-Pedagogical Research], vol. 1, no. 2, pp. 12–20. https://doi.org/10.23939/veritas2025.02.012

Deepfake and audio forgeries: the new era of political disinformation abstract

Pekhnyk Alevtina Valentynivna

Candidate of Political Sciences, Associate Professor,
Head of the Department of Journalism, Advertising and Media Communications
Odesa I. I. Mechnikov National University
Frantsuzkyi Blvd Blvd, 24/26, Odesa, Ukraine
ORCID: 0000-0003-2534-7652

Kroitor Artem Viktorovych

Candidate of Political Sciences, Associate Professor, Acting Director
Separate Structural Unit "Odesa Professional College of Computer Technologies of Odesa I. I. Mechnikov National University"
Akademika Korolova str., 5/2, Odesa, Ukraine
ORCID: 0000-0003-4652-7441

Bakhmetiev Andriy Evgeniyovych

PhD in Political Science,
Senior Lecturer at the Department of Journalism, Advertising and Media Communications
Odesa I. I. Mechnikov National University
Frantsuzkyi Blvd, 24/26, Odesa, Ukraine
ORCID: 0000-0002-8509-0003

The active development of generative artificial intelligence technologies has led to the emergence of new forms of political disinformation, in particular deepfake video and audio fakes, which significantly transform the mechanisms of political communication in the digital environment. Hyperrealistic audiovisual falsifications create fundamentally new challenges for democratic institutions, as they undermine trust in media content as a source of verified information and call into question the very institution of visual evidence.

The aim of the article is a comprehensive analysis of deepfake and audio fakes as tools of political disinformation and an assessment of their impact on the functioning of political communication and democratic processes. The methodological basis of the study is based on a combination of general scientific and special methods. Bibliographic analysis was used to outline theoretical approaches to the study of digital disinformation; methods of analysis, synthesis and generalization - to systematize the main forms and mechanisms of using deepfake in political practice; a structural-functional approach - to determine their role in the system of political communication; comparative method – to compare cases of application of such technologies in different political contexts; content analysis and elements of case studies – to study specific examples of digital forgeries and reactions to them.

The article reveals the specifics of deepfake as a high-tech tool of informational and psychological influence, capable of being integrated into strategies of political struggle and hybrid confrontation. It is shown that the key threat lies not only in the creation of individual falsified materials, but in the formation of an environment of total doubt in which any audiovisual content can be called into question. Such a transformation of the information space contributes to the spread of post-truth logic, increased polarization and a decrease in the level of rational political choice. The impact of deepfake technologies on electoral processes, reputational risks of political actors and the stability of state institutions is analyzed.

As a result, the need to form a comprehensive countermeasure system is substantiated, which involves the development of preventive communication mechanisms, improving crisis response procedures, raising standards of journalistic verification and introducing regulatory regulation in the field of labeling content created using artificial intelligence.

Key words: *deepfake, audio fakes, political disinformation, artificial intelligence, political communication, hybrid war, post-truth, information security.*

Дата першого надходження статті до видання: 17.02.2026
Дата прийняття статті до друку після рецензування: 19.03.2026
Дата публікації (оприлюднення) статті: