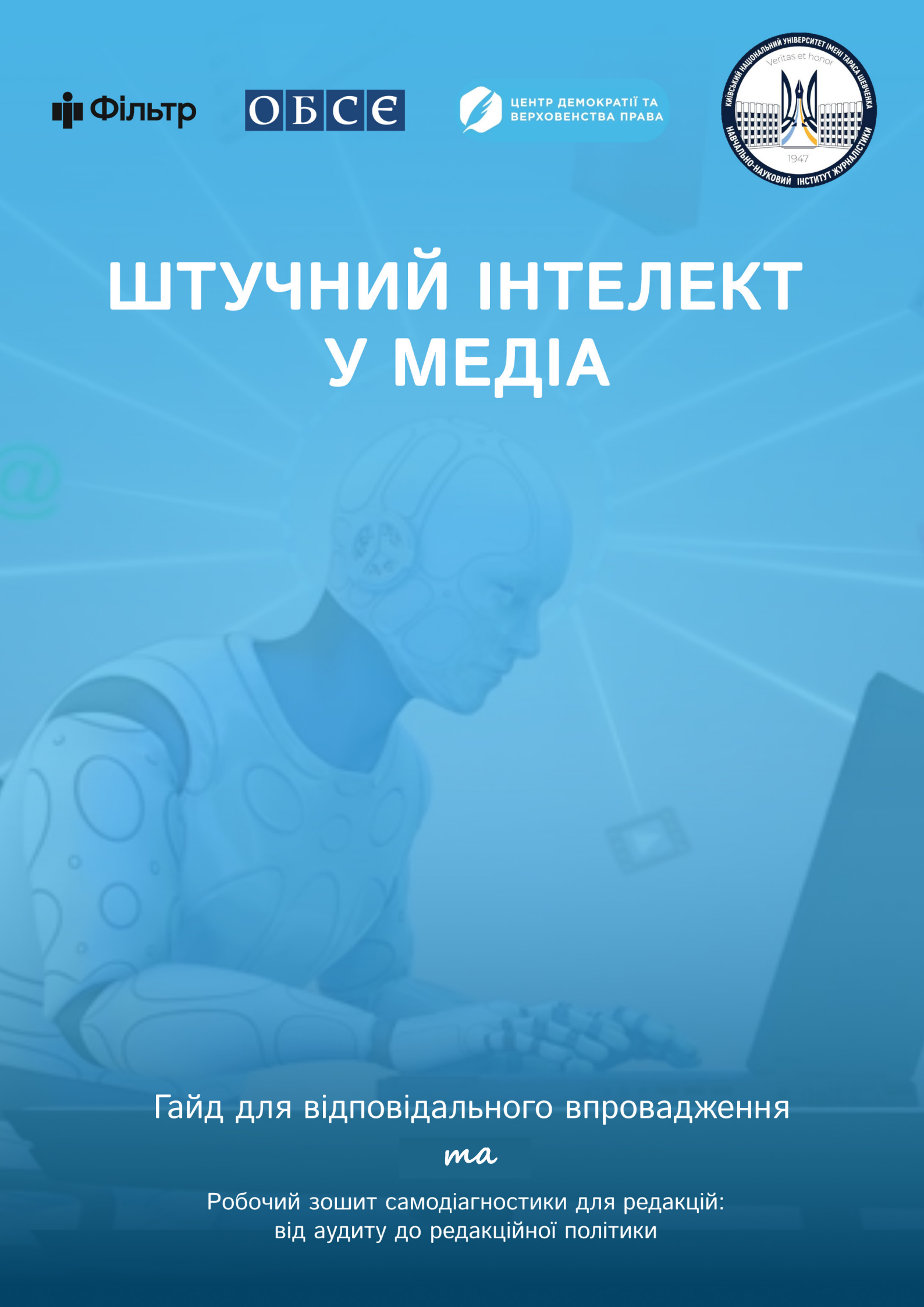
Фільтр

ОБСЄ

ЦЕНТР ДЕМОКРАТІЇ ТА ВЕРХОВЕНСТВА ПРАВА

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА
Veritas et honor
1947
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ЖУРНАЛІСТИКИ

ШТУЧНИЙ ІНТЕЛЕКТ У МЕДІА



Гайд для відповідального впровадження

та

Робочий зошит самодіагностики для редакцій:
від аудиту до редакційної політики

УДК: 070:004.8
Ш95

Ш95 Штучний інтелект у медіа: гайд для відповідального впровадження / Колектив авторів: **Горська К.**, **Дуднік Т.**, **Кравченко О.**, **Нестеренко А.**, **Романюк А.**, **Шевченко Т.** — Київ : Програма підтримки ОБСЄ для України, ННІЖ КНУ імені Тараса Шевченка, Проєкт «Фільтр» МКСК, Центр демократії та верховенства права, 2025. — 77 с.

Це практичне керівництво для медіа, що стоять перед викликом інтеграції ШІ-технологій у свою роботу. Видання призначене для журналістів, редакторів, керівників медіа — усіх, хто прагне зберегти якість та етичні стандарти журналістики в епоху штучного інтелекту. Робочий зошит самодіагностики, що міститься в гайді, допоможе редакції структуровано підійти до оцінювання поточних практик, виявлення потенціалу та ризиків, а також формування власної редакційної політики щодо ШІ.

© Програма підтримки ОБСЄ для України, 2025
© Навчально-науковий інститут журналістики КНУ імені Тараса Шевченка, 2025
© Національний проєкт з медіаграмотності «Фільтр»
Міністерства культури та стратегічних комунікацій України, 2025
© Центр демократії та верховенства права, 2025
© Колектив авторів, 2025

ПРО ГАЙД

Кожна редакція сьогодні стоїть перед вибором: як працювати зі штучним інтелектом? Ігнорувати нові технології, сліпо копіювати чужий досвід чи знайти власний шлях? Цей гайд створений спеціально для тих, хто обрав третій варіант.

Штучний інтелект (далі — ШІ) вже не майбутнє — він тут і зараз. Ми стикаємося з ним щодня: у підказках телефона, рекомендаціях соцмереж, навігаторах. Медіа теж активно експериментують із цими можливостями. Але між «ШІ — це круто» і «ШІ нас замінить» лежить величезна територія практичних питань.

Цей гайд допоможе вашій редакції розібратися в хаосі інформації про ШІ та зробити усвідомлений вибір. Тут ви знайдете відповіді на запитання: які інструменти підійдуть саме вашій команді, як оцінити ефективність використання ШІ, що робити з етичними дилемами та правовими нюансами.

Від базових понять до складних викликів, від вибору технологій до створення внутрішніх правил роботи з ШІ — ми зібрали практичний досвід і структурували його, використавши найсучасніші досягнення ШІ-технологій, щоб ви могли застосувати ці знання вже завтра.

Бонус: у гайді ви знайдете робочий зошит для самодіагностики редакції. Це практичний інструмент, який допоможе системно оцінити готовність вашої команди до роботи з ШІ, виявити сильні сторони і проблемні зони, а головне — створити власну редакційну політику, яка відповідатиме вашим цінностям і завданням.

Бо, зрештою, ШІ — це просто інструмент. А те, як ми його використаємо, залежить від нас.

ЗМІСТ

Розділ 1. Що варто знати про ШІ

- Базові принципи роботи ШІ та його види
- Впровадження ШІ в українських і зарубіжних медіа
- Як ШІ може допомогти на різних етапах роботи редакції
- Де шукати корисні інструменти для виконання редакційних завдань

Базові принципи роботи ШІ та його види

Популярний у соціальних мережах міф стверджує, що штучний інтелект автоматично знає особисті дані всіх користувачів, включно з їхнім місцезнаходженням. Насправді різні ШІ-системи мають різні рівні доступу до інформації — від повністю ізольованих моделей до тих, що інтегровані з різними сервісами та базами даних. Тож розберемося, як працюють ці системи.

Штучний інтелект — це не суперробот із кіно чи чарівна програма, яка знає про вас усе. ШІ — це насамперед інструмент, який бере багато інформації, аналізує її, шукає закономірності й видає результат, який вам підходить. І що більше він працює, то краще розуміє, що саме вам потрібно. Простими словами, він постійно вдосконалюється.

Згідно з **визначенням**, яке пропонує команда IBM (International Business Machines Corporation) — один із провідних світових розробників корпоративних рішень штучного інтелекту та програмного забезпечення для бізнесу,

«Штучний інтелект — це технологія, яка дає змогу комп'ютерам і машинам імітувати людське навчання, розуміння, розв'язання проблем, прийняття рішень, креативність та автономію».

Кожен з нас щодня стикається з роботою штучного інтелекту, навіть якщо здається, що такі технології взагалі не використовуються. Наприклад, коли ви пишете повідомлення, а телефон автоматично підказує слова, — це ШІ. Коли Instagram чи TikTok показує саме ті відео, які вам цікаві, — це теж робота ШІ. Google Maps прокладають найзручніший маршрут, Siri відповідає на ваші запитання чи навіть жарти, пошта автоматично фільтрує спам, а камера на телефоні впізнає обличчя — усе це приклади використання штучного інтелекту в повсякденному житті.

Сьогодні часто можна почути пересторороги, що штучний інтелект — це загроза, яка може кардинально змінити ринок праці та суспільні процеси. Втім, замість страхів, варто підходити до цієї технології усвідомлено. Насправді ШІ — це потужний інструмент з безліччю можливостей, який можна використовувати як у повсякденному житті, так і в професійній діяльності. А які види ШІ існують, розглянемо далі.

Види штучного інтелекту

Уже згадана компанія IBM, що розробляє ШІ по всьому світу, **пропонує** виділяти види штучного інтелекту за двома критеріями: можливостями та функціональністю.

Види ШІ за можливостями:

Вузький (Narrow) ШІ — це єдиний вид ШІ, який наявний зараз. Його так називають, бо він спеціалізується на виконанні одного конкретного завдання — і робить це дуже швидко та якісно, навіть часто краще за людину. **Генеративний ШІ** є різновидом **Вузького (Narrow) ШІ**. Хоча він може створювати різноманітний контент (тексти, зображення, відео, музику), він все одно спеціалізується на конкретному завданні — генерації контенту на основі навчальних даних. Він не може вийти за межі цього завдання й виконувати будь-які інші розумові настанови без додаткового програмування. Приклади такого ШІ — голосові помічники Siri від Apple чи Alexa від Amazon, система IBM Watson і ChatGPT, який спеціалізується на спілкуванні текстом у чаті. Програмісти використовують генеративний ШІ, щоб писати коди. Маркетологи можуть легко придумувати ідеї для рекламних кампаній.

Вузький ШІ допомагає журналістам автоматизувати рутинні завдання, наприклад, швидко знаходити релевантні факти або статистику для статей, розпізнавати мову в аудіо- чи відеозаписах і створювати чорнові варіанти текстів. Для медійників це можливість зекономити час на технічних деталях і зосередитися на створенні оригінального контенту. Загальний (General) ШІ — це поки

що лише ідея, теоретична модель. Такий ШІ міг би виконувати будь-які розумові завдання без додаткового втручання людини.

Супер (super) ШІ — це ще більш розвинений тип штучного інтелекту, який поки що існує тільки в уяві, фактично фантастика. Якщо колись його створять, він зможе мислити і приймати рішення набагато краще за людину. Уявіть собі, такий ШІ навчиметься сам, розумітиме складні ситуації, відчуватиме емоції і матиме власні цілі та переконання. Це означає, що він не просто виконуватиме завдання, а діятиме, ніби має власну свідомість і почуття. Наразі це лише теорія і ніхто не знає, чи стане це реальністю і коли. Звучить як фантастика, але це, звісно, питання часу.

Види ШІ за функціональністю:

Реактивний машинний ШІ — це найпростіший тип штучного інтелекту. Такі системи не мають пам'яті і не можуть згадувати, що було раніше. Вони працюють тільки з тими даними, які є в моменті, і виконують конкретне завдання. Цей ШІ використовує математичні правила, щоб швидко аналізувати інформацію і давати результат. Прикладами реактивного машинного ШІ є шаховий комп'ютер IBM Deep Blue, який у 1997 р. переміг чемпіона світу з шахів, швидко аналізуючи положення фігур на дошці і обираючи найкращий хід, а також система рекомендацій Netflix, яка на основі вашої історії переглядів підбирає фільми і серіали, що можуть сподобатися.

ШІ з обмеженою пам'яттю. На відміну від реактивного ШІ, який працює тільки з даними тут і зараз, ШІ з обмеженою пам'яттю може запам'ятовувати нещодавні події та інформацію. Він використовує ці дані, щоб приймати більш точні рішення, хоча й не зберігає їх назавжди.

Прикладами штучного інтелекту з обмеженою пам'яттю є генеративні інструменти, як ChatGPT, Bard та DeepAI, які використовують цю пам'ять для передбачення наступних слів або зображень у створеному контенті. Також сюди належать віртуальні помічники і чатботи, такі як Siri, Alexa, Google Assistant, Cortana та IBM Watson Assistant, які розуміють запити користувачів і дають відповіді, використовуючи обробку природної мови і пам'ять. Автономні автомобілі, тобто ті, що керуються без водія, теж працюють на основі такого ШІ, аналізуючи інформацію про оточення в реальному часі, щоб приймати рішення про рух — коли прискорюватися, гальмувати або повертати.

Також до цієї категорії належить генеративний штучний інтелект — технологія, яка створює новий контент: тексти, зображення, відео, музику. Такі системи використовують контекст попередніх запитів та взаємодій, щоб генерувати більш релевантний і послідовний контент. Генеративний ШІ допомагає журналістам, SMM-фахівцям, програмістам та маркетологам автоматизувати творчі завдання, хоча результат потребує людського контролю та доопрацювання.

ШІ з теорією розуму. Це тип штучного інтелекту, який матиме здатність розуміти психічні стани людей — їхні думки, емоції, наміри та переконання. Такі системи зможуть розпізнавати емоційний стан користувача й адаптувати свою поведінку відповідно до ситуації. Наприклад, ШІ зможе визначити, що людина засмучена, і змінити тон спілкування або запропонувати підтримку. На відміну від сучасних систем, які лише імітують розуміння через аналіз тексту або голосу, ШІ з теорією розуму справді розумітиме мотиви та емоційні потреби людей. Наразі це концептуальна модель, над створенням якої працюють дослідники у сфері когнітивних наук та машинного навчання.

Самосвідомий ШІ. Найвищий теоретичний рівень штучного інтелекту, який матиме власну свідомість, самоусвідомлення і здатність до рефлексії. Такі системи зможуть не лише розуміти емоції та наміри інших, а й усвідомлювати власні ментальні процеси, формувати особисті цілі та переконання.

Самосвідомий ШІ був би здатний до самоаналізу, розуміння власних сильних і слабких сторін, а також до формування унікальної «особистості». Це найбільш гіпотетичний тип ШІ, створення якого потребуватиме революційних проривів у розумінні природи свідомості та самосприйняття. Наразі це лише ідея і таких систем ще не створили.

Впровадження ШІ в українських і зарубіжних медіа

Штучний інтелект дедалі активніше інтегрується в роботу медіа по всьому світу — від систематизації даних до генерації контенту. За даними дослідження **PwC Global AI Study**, понад 64% медіакомпаній уже використовують ШІ у тій чи тій формі. Серед основних сфер застосування — розшифрування інтерв'ю, створення візуального контенту, генерація текстів, аналітика авдиторії, персоналізація новин та SEO-оптимізація.

Штучний інтелект справді спрощує роботу медійників. Однак разом із тим, як зазначає експерт Джек Шейфер з видання **Politico**, ШІ не здатен замінити живе спілкування з людьми і роботу журналіста безпосередньо на місці події. Адже багато важливої інформації просто не знайдеш у відкритих джерелах або через пошук у ChatGPT. Водночас, на його думку, ШІ може допомогти зробити більше новин. Проте це не завжди означає покращення якості контенту[1].

Як в Україні, так і в інших країнах медіа поступово починають впроваджувати ШІ у щоденну роботу. У 2024 р. ринок ШІ в медіа та розвагах аналітики **оцінювали** майже у 20 мільярдів доларів, і очікується, що вже до 2033 р. він зросте у шість разів! Україна, згідно з «**Аналізом секторального напряму та первинного бачення розвитку сфери ШІ**» 2025 р., вже «посідає 2-ге місце в Східній Європі за кількістю компаній, що працюють у галузі ШІ». До 2030 р. у держави амбітна мета — увійти в трійку лідерів за рівнем інтеграції та впровадження ШІ-рішень.

Використання ШІ в редакціях: іноземний досвід

Ми вже бачимо, як розвиток штучного інтелекту змінює медіаландшафт, створюючи як безпрецедентні можливості, так і серйозні виклики для журналістів і контент-креаторів. З одного боку, медійники використовують ШІ-сервіси для полегшення роботи: транскрибування аудіозаписів, редагування текстів, пошуку ідей, створення інфографіки, подкастів та відео. З іншого, все більше користувачів замість пошуку інформації через Google користуються ChatGPT, що, зі свого боку, призводить до зниження трафіку на сайтах. Це впливає на доходи від реклами, впізнаваність і довіру до медіа загалом. Автоматизація редакційних процесів запускає процеси звільнення власниками медіа своїх працівників або скорочення штату по всьому світу. Одним із перших гучних кейсів 2023 р. стало оголошення компанії BuzzFeed (США) про інтеграцію генеративного ШІ у створення контенту. Генеральний директор Йона Перретті в січні 2023-го **заявив**, що ШІ стане частиною основного бізнесу BuzzFeed та використовуватиметься для персоналізації інтерактивного контенту (наприклад, вікторин), підказок для мозкового штурму і створення розважальних матеріалів на основі напрацювань редакції. Зокрема, вже протягом 2023 р. BuzzFeed запустив серію ШІ-згенерованих вікторин та тревел-гайдів для своєї авдиторії. Інвестори сприйняли цю новину оптимістично — **акції компанії стрімко зросли** після оголошення про співпрацю з OpenAI та Meta, яка передбачала залучення ШІ до контент-процесів BuzzFeed.

India Today Media Group, один з найбільших медіахолдингів Індії, **у 2023 році представив ШІ-ведучу новин на ім'я Sana**. В ефірі вона розповідала короткі новинні дайджести кількома мовами. Проєкт позиціювався як перший подібний в Індії, і компанія наголошувала на «співпраці людини і ШІ» в редакції: Sana працювала під наглядом продюсерів і в парі з відомим ведучим новин Судхіром Чаудхарі у його вечірньому шоу[2].

---

1 Jack Shafer. How AI Is Already Transforming the News Business. Politico. 27.02.2024.

2 BW Marketing World. AI Newsreader «Sana» To Join Sudhir Chaudhary's Show On Aaj Tak. BW Marketing World. 2023 (точна дата на сайті не вказана).

Джерело: **India Today**

Sana може миттєво переключатися між мовами (англійська, гінді тощо), не має проблеми втоми, що дозволяє їй часто оновлювати випуски новин протягом дня. Її впровадження було частиною стратегії India Today щодо цифрової трансформації ньюзруму. Кейс отримав багато позитивної уваги. У 2024 р. Sana принесла India Today **міжнародну нагороду INMA** (Global Media Award) за лідерство в ШІ-трансформації ньюзруму. Журі відзначило, що впровадження ШІ-ведучої суттєво підвищило залученість авдиторії та диференціацію контенту медіагрупи. Представники компанії відзначають, що Sana стала «революційною силою» в медіаландшафті Індії, привернувши увагу молодої авдиторії до теленовин[3].

У жовтні 2024 року радіостанція OFF Radio Kraków провела тижневий експеримент — **запустила першу в Польщі радіостанцію, створену майже повністю штучним інтелектом**[4] [5]. Програми були представлені трьома ШІ-персонажами, що уособлювали представників молодого покоління та мали привернути їхню увагу. Музичні плейлісти створювалися за допомогою інструментів штучного інтелекту (під наглядом режисера-людини), а частину журналістів, які працювали на OFF Radio Kraków звільнили. Невдовзі після запуску проєкт OFF Radio Kraków показав уявне інтерв'ю, згенероване штучним інтелектом, з Віславою Шимборською, польською поетесою, лауреаткою Нобелівської премії, яка померла 2012 року. Планувалося інтерв'ю з Юзефом Пілсудським. Втім, експеримент викликав широке обговорення, хвилю критики і негативу. Матеуш

3 Ajmal Abbas. India Today Group's AI news anchor, Sana, wins global media award for AI-led newsroom transformation. India Today. 27.04.2024.

4 MediaMaker. Польська радіостанція замінила журналістів на ШІ-ведучих. 22.10.2024.

5 Notes from Poland. Controversy after Polish radio station replaces human presenters with AI.22.10.2024.

Демський, журналіст і кінокритик, який вів програму на OFF Radio Kraków до початку експерименту, опублікував петицію із протестом проти «заміни працівників штучним інтелектом», яку за кілька днів підписали понад 15 000 людей[6]. Згодом редакція була змушена закрити цей проєкт.

Ведучий радіопроєкту, створений за допомогою ШІ. Джерело: **Focus.ua**

Американський технологічний сайт CNET став одним із перших великих медіа, що спробували використовувати **ШІ для написання статей** — і зіткнулися з негативними наслідками. У кінці 2022 — на початку 2023 р. CNET тихцем публікував статті, згенеровані нейромережею[7] (переважно на фінансові теми, такі як поради з економії, опис фінансових термінів тощо). Виявилося, що матеріали під авторством «CNET Money Staff» писав внутрішній ШІ-інструмент. Коли в січні 2023-го це з'ясувалося завдяки журналістам видання Futurism[8], CNET опинився в центрі скандалу. Аналіз показав численні помилки у таких текстах: із 77 опублікованих ШІ-статей понад половину — 41 — довелося виправляти. Серед виявлених помилок були як фактологічні неточності (некоректні відсоткові показники, формули тощо), так і стилістичні недоліки, а в окремих випадках зафіксовано текстові збіги, що можуть свідчити про плагіат. Після хвилі критики керівництво спершу оголосило про паузу в публікації ШІ-контенту. Головна редакторка Конні Гульєльмо у відкритій заяві захищала експеримент, зазначивши, що метою було розширити обсяг інформаційного сервісу для читачів, але визнала «недопрацювання» та пообіцяла вдосконалити процес перевірки фактів[9]. Зрештою CNET не відмовився від ШІ повністю — компанія Red Ventures (власник CNET) повідомила, що продовжить тестувати ШІ-інструменти для допомоги редакції, але з більшою прозорістю і контролем.

---

6 Notes from Poland. Polish broadcaster shuts down AI-run radio station after a week following backlash. 28.10.2024.

7 Roth, Emma. CNET found errors in more than half of its AI-written stories. The Verge. 25.01.2023.

8 Futurism. CNET quietly published dozens of AI-generated articles, then retracted many of them. Futurism. 24.01.2023.

9 Roth, Emma. CNET found errors in more than half of its AI-written stories. The Verge. 25.01.2023.

Кейс CNET став показовим негативним прикладом: він продемонстрував ризики сліпої довіри генеративному ШІ в журналістиці — без ретельної модерації це призводить до фактичних помилок і підриву довіри до медіа.

Ще один показовий випадок стався в Німеччині у квітні 2023 р.: журнал Die Aktuelle (щотижневий таблоїд) опублікував нібито «ексклюзивне інтерв'ю» з легендарним гонщиком Міхаелем Шумахером, який з 2013 р. не з'являється на публіці через тяжку травму. Насправді, як з'ясувалося, текст інтерв'ю був повністю згенерований штучним інтелектом — редакція використала ШІ для імітації відповідей Шумахера на поставлені запитання[10].

Публікація викликала значний резонанс: родина Шумахера подала позов до суду, а громадськість різко засудила такий неетичний підхід. Видавництво було змушене принести публічні вибачення родині чемпіона, а головний редактор журналу втратив посаду. Інцидент чітко окреслив межі допустимого: використання ШІ для створення фальшивих висловлювань реальних людей є формою дезінформації та грубим порушенням журналістських стандартів.

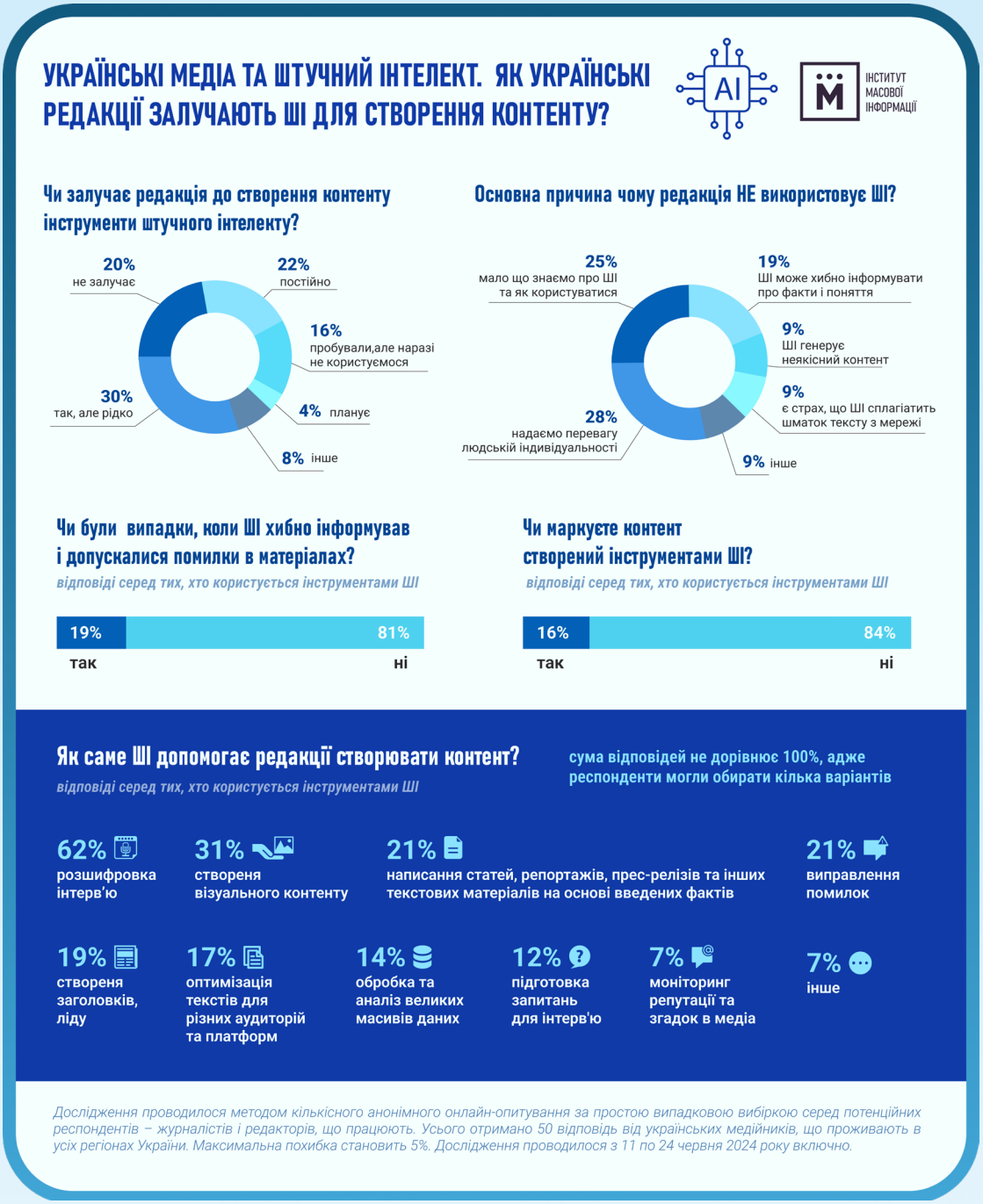
Скандал спонукав інші німецькі медіа посилити внутрішні процедури контролю за використанням ШІ-технологій. Ситуація з Die Aktuelle стала важливим уроком для всієї галузі: безвідповідальне застосування штучного інтелекту заради створення сенсаційного контенту може завдати непоправної шкоди довірі до медіа.

Використання ШІ в редакціях: українські практики

Попри зростання інтересу до ШІ в міжнародному медіасередовищі, в українських редакціях його впровадження поки що обмежене. Згідно з опитуванням Інституту масової інформації (IMI), «лише 22 % українських медійників повідомили, що їхня редакція на постійній основі використовує хоча б один інструмент штучного інтелекту»[11].

10 Kingsley, Patrick. German tabloid Bild to replace range of editorial jobs with AI // The Guardian. 20.06.2023.

11 Яна Машкова. Українські медіа та штучний інтелект. Як редакції залучають ШІ для створення контенту?/ IMI. 01.07.2024.



Джерело: **Інститут масової інформації (IMI)**

Втім, незважаючи на загалом невеликий відсоток впровадження інструментів ШІ в українських редакціях, українські медіа теж демонструють успішні практики його використання. Так, наприклад, полтавське видання Kolo.news **використовує ШІ-згенеровану ведучу** на ім'я Наталка-Полтавка для зачитування новин[12].

12 Коло. Всі новини Полтави. Дорогі читачі, хочемо представити вам нову журналістку редакції Kolo.news Наталку Полтавку, яку ми згенерували за допомогою… // Facebook. 11.03.2024.

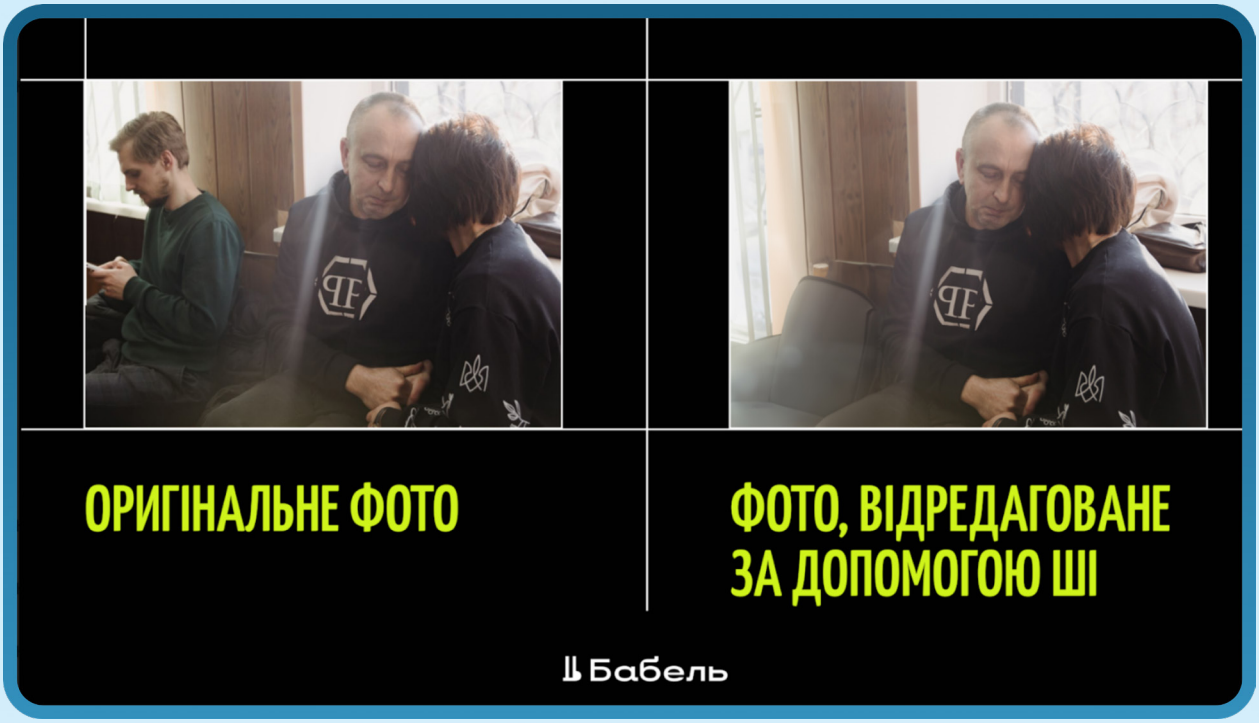
ШІ-ведуча Наталка-Полтавка. Джерело: **видання «Коло»**

Українське онлайн-видання про технології **Gagadget провело серію експериментів** із впровадження штучного інтелекту. Зокрема, у 2023 р. сайт запустив автоматичний переклад новин різними мовами (на піку було 12 мовних версій ресурсу), що мало на меті розширити авдиторію за рахунок багатомовності контенту. Також редакція створила власного чатбота, який допомагав читачам взаємодіяти з контентом видання. Ці інновації продемонстрували потенціал ШІ для технічних завдань — автоматизації перекладу та агрегування новин, однак потребували ретельного людського контролю якості. Досвід Gagadget показав, що ШІ може ефективно виконувати рутинні операції, але не замінює редакторів і журналістів у питаннях точності та творчого підходу.

Інше українське онлайн-видання «Бабель» використовує штучний інтелект як помічника для різнопланових завдань. Редакція свідомо підходить до інтеграції ШІ, використовуючи його лише там, де це «справді спрощує життя і приносить користь»[13]. Серед базових завдань, які «Бабель» довіряє ШІ, — **транскрибація аудіо/відео в текст** і переклад. Команда створила власного телеграм-бота «Dedel» на основі API OpenAI (ChatGPT 3.5) — цей бот вміє розшифровувати аудіозаписи й одразу перекладати їх потрібною мовою приблизно за дві хвилини.

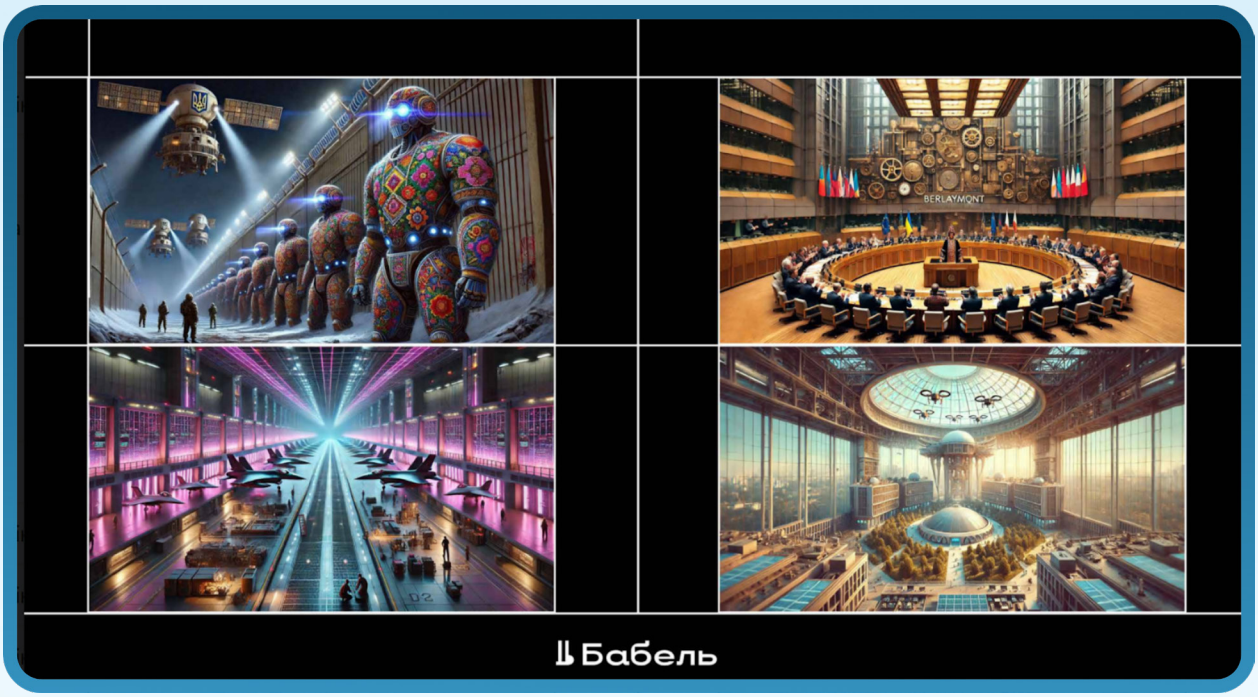
Обробка зображень у «Бабелі» теж частково покладена на ШІ: у Photoshop журналісти користуються вбудованими нейромережевими функціями для редагування фото, як-от розширення фону, видалення зайвих об'єктів або людей з кадру тощо.

13 Нановська, Вероніка. Редакційний різнороб. Які завдання «Бабель» доручає штучному інтелекту. MediaMaker. 24.09.2024.



Джерело фото: **Медіамейкер**

Іноді «Бабель» експериментує з **генерацією нових зображень**, але робить це рідко, здебільшого для спецпроєктів. Наприклад, до Дня Незалежності редакція запропонувала читачам уявити майбутнє України і згенерувала серію картинок на основі їхніх описів (через ШІ-генератори, під'єднані до ChatGPT).



Джерело фото: **Медіамейкер**

Для **пошуку інформації** журналісти видання використовують ШІ-пошуковик **Perplexity** — це допомагає швидко збирати довідкові факти, шукати потенційних спікерів чи підказки щодо тем, не покладаючись цілковито на традиційний пошук.

Підготовка грантових заявок — ще одна царина, де «Бабель» залучає ШІ. Оскільки мова грантів часто перевантажена канцеляризмами, журналісти іноді використовують ChatGPT (версії 4 і вище), щоб швидше підготувати чернетки заяв в офіційному стилі.

Новинний ресурс «24 канал» інтегрував ШІ для **оптимізації як контенту**, так і внутрішніх процесів. Використання інструментів ШІ з 2023 р. суттєво покращило показники сайту — глибину перегляду та час, проведений на сторінці, що було підтверджено A/B-тестуванням. Одне з головних нововведень — впровадження короткого аудіоогляду новин: редактори готують відібрані текстові стислі новини, а ШІ озвучує їх синтезованим голосом (без імітації реальних диктторів). Станом на середину 2024 р. такі аудіоновини додаються вже до ~70 % матеріалів сайту[14].

Крім того, ШІ **автоматизує розшифрування інтерв'ю та відео**: алгоритм перетворює відео на текст, а редактори лише редагують і звіряють його з оригіналом, що суттєво економить час. Рутинні тексти — такі як офіційні пресрелізи — журналісти інколи доручають нейромережі перефразувати для підвищення читабельності (усуваючи канцеляризми та додаючи оригінальності), а також генерувати варіанти заголовків і підбирати SEO-ключові слова. У розважальних розділах сайту («Life», «Men», «Pets» тощо) редакція навіть створює нескладні тести/вікторини за допомогою ШІ — нейромережа генерує текст питань і відповідей, які потім редагуються журналістом перед публікацією[15]. Усі згенеровані матеріали мають дисклеймери про використання ШІ.

Команда тернопільського телеканалу TV-4 застосовує ШІ для підбору тем, що цікавлять авдиторію, переформатування новин для сайту, а також у написанні структури сценаріїв для телепрограм. У тернопільській редакції видання «20 хвилин» ШІ активно використовують[16] для полегшення щоденної роботи: автоматизують рутинні завдання, допомагають з відеомонтажем, озвучуванням, редагуванням текстів і навіть плануванням робочого дня. Це дає змогу команді заощаджувати час і зосереджуватися на творчих процесах.

Поповнюють ряди тих, хто впроваджує ШІ, і державні установи. Так, Міністерство закордонних справ України створило **цифрову особу Вікторію Ші**, яка офіційно має коментувати консульську інформацію для медіа. За словами міністра,

«створення цифрової консульської представниці — це частина ширшої стратегії з системного впровадження передових технологій з використанням штучного інтелекту в МЗС України»[17].

Створення таких цифрових аватарів порушує питання безпеки даних та верифікації інформації. МЗС акцентує, що передбачило кілька рівнів захисту Вікторії Ші від цифрових підробок:

14 Аліна Єлевтєрова. Креативник і оптимізатор. До якої роботи і на яких умовах редакція 24 каналу залучає ШІ. MediaMaker. 15.08.2024.

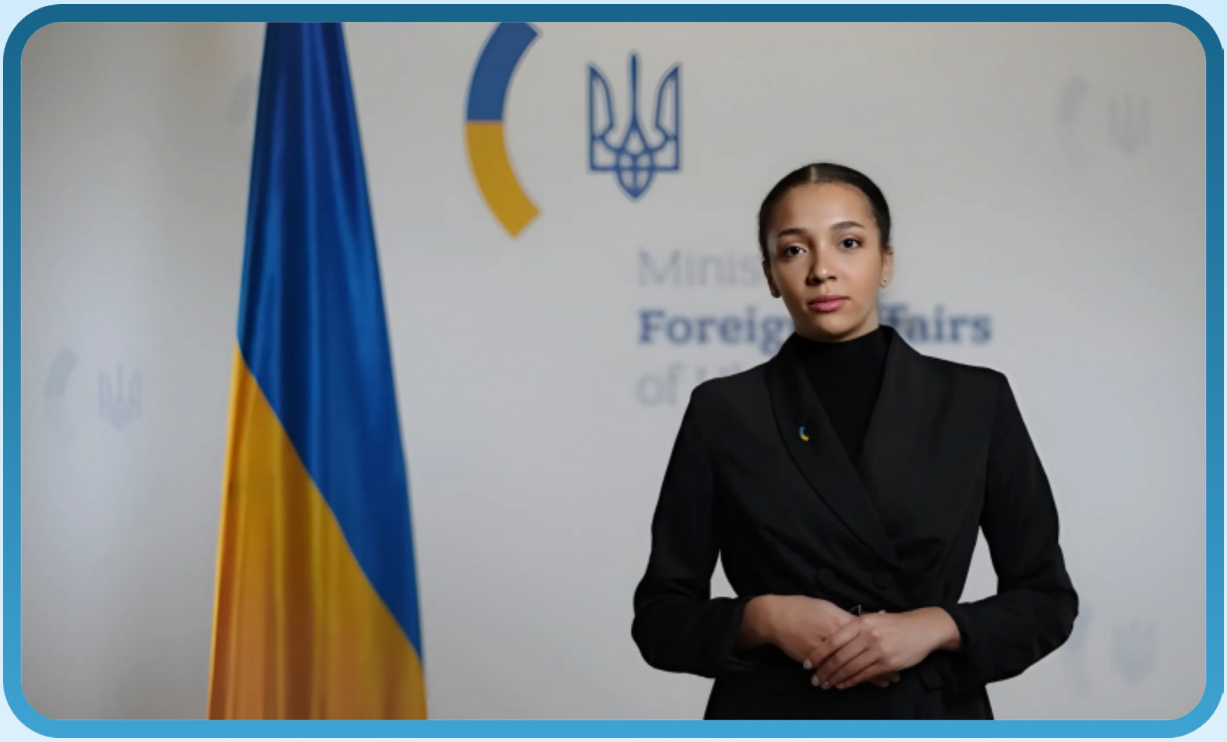
15 Аліна Єлевтєрова. Креативник і оптимізатор. До якої роботи і на яких умовах редакція 24 каналу залучає ШІ. MediaMaker. 15.08.2024.

16 Богдан Звоник. Пересторороги, експерименти, сфери застосування: як українські регіональні медіа «запускають» до себе ШІ. MediaMaker. 20.10.2024.

17 Міністерство закордонних справ України. МЗС України призначило «цифрову особу» для інформування щодо консульських питань. 01.05.2024.

«на кожному оригінальному відео консульської представниці розміщений QR-код, який веде на текстову версію того ж самого коментаря на офіційній сторінці МЗС. У разі відсутності такого QR-коду чи переходу на будь-який інший сайт відео не вважається достовірним»[18].

Втім, будьмо відверті, на бажання, підробити QR-код та створити односторінковий сайт, який матиме такий самий вигляд, як сайт МЗС, теж достатньо легко.



Цифрова особа Вікторії Ші від Міністерства закордонних справ України. **Джерело: МЗС України**

18 Міністерство закордонних справ України. МЗС України призначило «цифрову особу» для інформування щодо консульських питань. 01.05.2024.

Як ШІ може допомогти на різних етапах роботи редакції

Штучний інтелект — це, звісно, не чарівна паличка, але він дійсно може полегшити життя команді редакції. Він не замінює журналістів, але допомагає з рутинними завданнями, економить час і відкриває нові можливості. ШІ слід розглядати як ефективний інструмент, який автоматизує технічні процеси та звільняє ресурси для створення якісного контенту і поглибленої аналітичної роботи.

Ось як він може бути корисним на різних етапах роботи:

Знаходити цікаві теми

ШІ швидко «проглядає» багато новин і соцмереж, щоб підказати, що зараз найцікавіше для людей. Завдяки цьому журналісти можуть легко знаходити гарні ідеї для своїх матеріалів і писати про них до того, як про них напишуть усі.

Перекладати з різних мов

ШІ допомагає перекладає тексти різними мовами. Хоч часто і не ідеально, і потрібна перевірка професіоналом, однак все одно це дуже зручно, якщо треба швидко опрацювати матеріали з іноземних джерел.

Впорядковувати інформацію

ШІ може швидко звірити тексти чи цитати і загалом впорядкувати інформацію.

Писати і редагувати тексти

ШІ може зробити чернетку статті, запропонувати структуру матеріалу, цікаві заголовки або скоротити текст статті для соцмереж. Він також допомагає виправити помилки і зробити текст легшим для читання.

Працювати з відео й аудіо

ШІ може не лише створювати презентації, а й монтувати відео, перетворювати текст в аудіо або навпаки. Це спрощує створення мультимедійного контенту.

Перевіряти факти і джерела

ШІ швидко перехресно перевіряє інформацію з різних джерел, допомагає знаходити первинні документи та виявляє суперечності в даних. Це особливо корисно у процесі роботи з великими обсягами статистичної інформації.

Аналізувати аудиторію

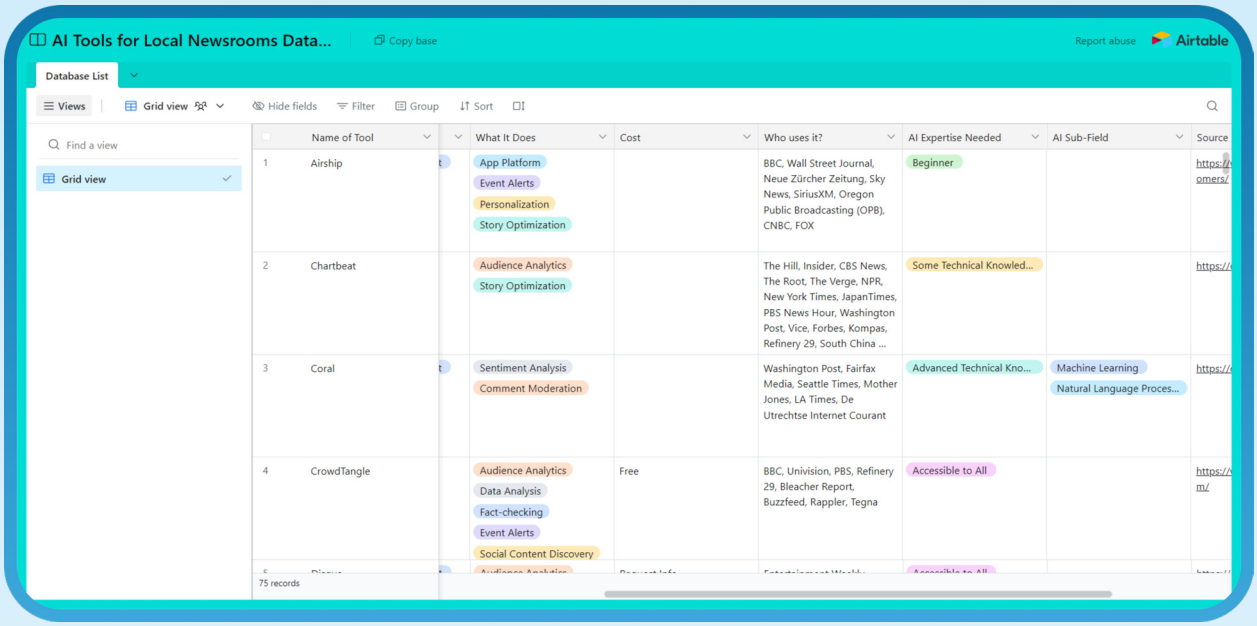
ШІ вивчає поведінку читачів, визначає найпопулярніші теми й формати контенту, що допомагає редакції краще розуміти потреби своєї авдиторії та планувати майбутні публікації.

Допомагати з публікацією і поширенням

ШІ підказує, коли і де краще публікувати матеріали, щоб їх побачило більше людей. Детально впровадження ШІ на різних етапах редакційних процесів розглянемо в **Розділі 2**.

Де шукати корисні інструменти для виконання редакційних завдань

Сьогодні існує велика кількість сервісів, де представлені інструменти для редакцій. Наприклад, на сайті **Airtable**, який допомагає створювати бази даних і керувати ними, розміщена зручна база ШІ-інструментів для медіа, що буде корисною й українським медійникам. Тут є чіткий опис кожного інструменту, інструкції щодо його застосування, а також приклади медіа, які вже використовують ці рішення.



Скрин інтерфейсу бази інструментів ШІ на **Airtable.com**

Далі даємо короткий список інструментів за функціоналом, що можуть бути корисними вашій редакції:

Аудіо та відео в текст

Transkriptor — платформа для автоматичного перетворення аудіо та відео в текстовий формат, що суттєво полегшує роботу журналістів під час підготовки матеріалів на основі записів.

Генерація, адаптація та пошук відео

Runway — креативна платформа, яка дає змогу створювати відеоматеріали через штучний інтелект без традиційного процесу знімання.

HeyGen — технологічне рішення для локалізації відеоконтенту, яке дає змогу адаптувати матеріали для різних мов, зберігаючи природність звучання оригінального мовця.

Filmot — спеціалізований пошуковий інструмент, який допомагає знаходити відеоматеріали за конкретними словами або фразами, що в них згадуються.

Lumen5 — конвертер контенту, який трансформує текстові матеріали блогів у динамічні відеоформати.

Faceless.video — автоматизована система для ведення акаунтів у соціальних мережах без постійного втручання користувача.

Predis.ai — сервіс для автоматизації процесу створення рекламних відеоматеріалів та планування публікацій через ШІ-технології.

Синтез мовлення та аудіо/музика

Suno — музичний генератор на базі ШІ, який дає змогу миттєво створювати оригінальні композиції з унікальним вокалом та мелодією.

ElevenLabs — передова технологія синтезу мовлення, яка забезпечує швидке та якісне створення аудіонаративу для різноманітного медіаконтенту.

Універсальні асистенти та контент-редактори

ChatGPT — універсальний ШІ-асистент для журналістської діяльності, який охоплює широкий спектр завдань – від створення текстового контенту до аналізу матеріалів та підготовки інтерв'ю.

Vista Social — інтегрує можливості ChatGPT для автоматичного створення контенту та підписів до публікацій у соціальних мережах.

HootSuite — традиційна платформа управління соціальними мережами, доповнена ШІ-асистентом Owly для генерації текстового контенту.

Маркетинг, аналітика і стратегії

Scalenut — комплексне рішення для розроблення контент-стратегій, що включає інструменти для планування кампаній та аналізу їхньої результативності.

Sprout Social — маркетингова платформа, яка об'єднує штучний інтелект з аналітичними інструментами для створення природного контенту для промокампаній.

Anyword — інтелектуальна платформа для створення маркетингових текстів із вбудованими аналітичними можливостями для оцінювання їхньої ефективності.

Приклади з українських та світових медіа демонструють, що ШІ вже став повноцінною частиною редакційної екосистеми — у ролі як помічника, так і потенційного джерела ризиків. Найбільшу цінність він має у сфері автоматизації рутинних процесів, локалізації контенту, генерації промоматеріалів і аналітики, тоді як експерименти із цілковитою заміною журналістів часто завершуються репутаційними втратами. Ці кейси підкреслюють необхідність виваженого й етичного впровадження інструментів ШІ в редакціях — з чітким фокусом на прозорість, якість і професійні стандарти.

Розділ 2. Оптимізація редакційних процесів за допомогою ШІ

- Генерування та перевірка ідей для матеріалів
- Автоматизація рутинних завдань
- Аналіз даних і фактчекінг
- Персоналізація контенту
- Інструменти моніторингу та аналітики

Генерування та перевірка ідей для матеріалів

У кожній редакції генерування ідей — це частина щоденної роботи, яка відбувається за інерцією. Зустрічі, де редактори перебирають стрічки новин, Telegram-канали, дивляться аналітику за переглядами. Але що робити, коли новий контекст вимагає тем і форматів, які раніше не використовувалися? Як реагувати на інформаційні хвилі, які піднімаються і падають буквально за години?

У 2023 р. команда **Reuters** створила внутрішню платформу Open Arena — захищене середовище, у якому журналісти можуть безпечно експериментувати з великими мовними моделями або LLM (від Large Language Model) (далі — LLM). Платформа дає змогу тестувати варіанти заголовків, структурувати майбутні історії, пропонувати альтернативні кути подання тем. Усі результати роботи LLM у цій системі проходять обов'язкову редакційну перевірку перед використанням у журналістських матеріалах.

BBC News Labs, інноваційний інкубатор, який відповідає за інновації для BBC News, пішли своїм шляхом. Вони створили прототип для генерації експлейнерів через ChatGPT-3. Як підкреслювала **Міранда Маркус**, виконавиця обов'язків керівника BBC News Labs, тексти, створені у такий спосіб, не публікуються автоматично: *«Жоден текст не потрапить до авдиторії, якщо журналіст вручну не вибере його і не доопрацює»*. Це ключовий принцип, якого дотримуються практично всі відповідальні медіа у світі.

Але важливо, що велика мовна модель не тільки допомагає з текстом, а й здатна

- підказати нові підходи до подання теми;
- проаналізувати інформаційне поле, запропонувавши, які формати «працюють» зараз для певної авдиторії;
- допомогти створити семантичні мапи — карту понять і пов'язаних тем, які журналіст може використати для розширення кута зору. Така мапа допомагає ширше подивитися на потенційний матеріал, уникнути банального кута подання та побачити, як тема «розгалужується» у різні сфери.

Наприклад, якщо журналіст працює з темою повернення військових до цивільного життя, LLM може допомогти окреслити семантичне поле цієї теми. Серед можливих напрямів — соціальна адаптація, психологічна реабілітація, проблеми з працевлаштуванням, сприйняття ветеранів у суспільстві, взаємини з родиною, доступ до медичної допомоги, участь у ветеранських об'єднаннях тощо. У кожному з цих напрямів — свої герої, свої виклики, свої джерела інформації.

Семантична мапа не є готовим редакційним планом, але вона дає змогу

- сформувати загальне бачення теми;
- побачити, де є «незаповнені ніші», які ще не були широко висвітлені;
- зрозуміти, які групи авдиторії можуть бути дотичні до проблематики;
- визначити, які експерти, герої чи формати можуть бути доречними.

Це дає змогу отримати структурований перелік, який стане відправною точкою для глибшого планування та подальшого «розробляння» теми.

Ризики використання великої мовної моделі (LLM):

- **LLM може стандартизувати подання**. Якщо сліпо слідувати пропозиціям і не вдосконалювати промпти (запити, з якими ви працюєте з ШІ), є ризик отримувати одноманітні пропозиції щодо тем, сформовані на англомовному масовому контенті.
- **Упередження**. Моделі можуть просувати твердження, які не відповідають українському контексту, або підсилювати вже наявні стереотипи.

- **Фальшива оригінальність**. LLM може генерувати «нові» підходи, які насправді є кальками з давно відомих форматів.

Саме тому у Reuters, BBC, Gannett усі експерименти йдуть за принципом «human-in-the-loop», де людина — центр процесу, а LLM лише помічник.

Отже, ШІ може допомогти зібрати перші чорнові підходи, запропонувати варіанти заголовків чи структури — але не замінити журналістську інтуїцію, досвід і контекстне мислення. Ідея, підказана LLM, — це лише початкова заготовка. Її унікальність та глибина з'являються тільки тоді, коли журналіст пропускає тему через себе: аналізує, ставить запитання, перевіряє джерела і формулює новий ракурс.

Що це означає для українських редакцій?

Багато медіа в Україні вже експериментують з LLM на етапі ідей. Але важливо відразу встановити правила гри:

- **Документувати використання ШІ** — журналіст має чітко розуміти, коли і як він використовує LLM для генерації ідей.
- **Фільтрувати результат** — жодна ідея не переходить у продакшн без редакційної оцінки.
- **Пояснювати авдиторії**, що контент створюється за участі журналістів, а не «генерується машиною».

На практиці це може мати такий вигляд: зібрана за допомогою LLM семантична карта (semantic map) допомагає редакції сформувати ширший спектр потенційних підходів до теми, зібрати контексти й порушити пов'язані питання. Далі долучається класична журналістська робота: редактор визначає релевантні напрями, журналіст розробляє матеріал, а фактчекери верифікують дані.

Чекліст: як оцінити ідею, згенеровану за допомогою ШІ

Перед тим як використовувати тему або ракурс, запропонований великою мовною моделлю (LLM), журналіст має оцінити їх за допомогою низки запитань. Такий самоперевірочний чекліст дає змогу відсіювати поверхові або повторювані ідеї, уникати неусвідомленого відтворення стереотипів і фокусуватися на темах, які справді допоможуть створити цінний для авдиторії контент.

Критерії, за якими варто оцінювати ідею:

- **Релевантність для авдиторії**

  Чи відповідає ця тема інтересам, потребам і життєвому контексту моїх читачів або глядачів?

- **Новизна або недоговореність**

  Чи є в темі новий ракурс, недосліджений аспект або спосіб подання, який не був широко представлений у ЗМІ?

- **Суспільна значущість**

  Чи допоможе ця тема краще зрозуміти складну ситуацію, зменшити рівень дезінформації або дати практичну користь авдиторії?

- **Етична безпека**

  Чи не містить ідея прихованих упереджень або ризику відтворення стереотипів (наприклад, гендерних, етнічних, вікових), які можуть мимоволі зашкодити?

- **Реалістичність реалізації**

  Чи має моя редакція ресурси, досвід, доступ до джерел, героїв або експертів, щоб якісно реалізувати цю тему?

Рекомендація щодо дій:

Якщо на щонайменше **три з п'яти запитань відповідь «так», ідею варто розвивати далі**. Якщо ж **більшість відповідей «ні» або «не впевнені», доцільно переформулювати запит до ШІ або обрати іншу тему**.

Автоматизація рутинних завдань

Окрім творчої роботи з темами, значна частина редакційної діяльності — це повсякденна рутина: транскрибування інтерв'ю, зведення статистики, перевірка біржових котирувань, моніторинг соцмереж, створення коротких дайджестів. За даними **опитування** Reuters Institute, «74 % медіалідерів вважають, що в найближчі 10 років генеративний ШІ допоможе робити ефективніше лише деякі речі: автоматизує бази, виконуватиме рутинні завдання, визначатиме потреби та бажання авдиторії»... «39 % респондентів розповіли, що в їхніх медіа вже працюють над тим, щоб залучити можливості ШІ у свій робочий процес; 21 % все ще розмірковує про таку можливість, а у 29 % вже узгоджені принципи використання нейромережі»[19].

Це частково пояснюється браком локалізованих інструментів, технічної підтримки та обмеженими ресурсами редакцій. До того ж, важливо чітко окреслити межі застосування ШІ: у журналістській практиці автоматизація насамперед стосується повторюваних, технічних завдань — таких як транскрибування, агрегування даних або переклад. Аналітичні і творчі аспекти роботи залишаються у виключній відповідальності журналіста. Такий розподіл не лише зберігає контроль над якістю, а й дає змогу редакціям ефективніше використовувати обмежені ресурси.

Типові рутинні завдання, які можна автоматизувати

Розшифрування аудіо- та відеозаписів. Автоматичні сервіси транскрипції (наприклад, Whisper, Otter.ai, AssemblyAI, PinPoint) дають змогу швидко отримати текстову версію інтерв'ю, брифінгу або виступу. Це особливо актуально при великому обсязі матеріалу. Водночас результати потребують редакторської перевірки — через можливі помилки в передаванні термінології, імен або назв.

Агрегація новин і моніторинг інформаційного поля. Інструменти на кшталт Feedly або NewsNow дають змогу автоматично збирати новини за ключовими словами, джерелами чи тематикою. Це корисно для підготовки оглядів, моніторингу суспільних настроїв або виявлення тенденцій. При використанні агрегаторів важливо зберігати ручний контроль за перевіркою джерел.

Генерування чернеток текстів і адаптація повідомлень. Моделі на зразок GPT-4 або Claude можна використовувати для створення первинних версій заголовків, анотацій, соцмережевих постів чи спрощених пояснень. У міжнародній практиці ці моделі також застосовують для адаптації текстів під конкретні авдиторії. Усі результати генерації потребують професійного редагування та перевірки.

Створення базових візуалізацій із даних. Сервіси на кшталт Flourish або Datawrapper дають змогу швидко побудувати графіки, діаграми чи часові шкали на основі відкритих даних. Це особливо

19 Changing Newsrooms 2023: Media leaders struggle to embrace diversity in full and remain cautious on AI disruption. Reuters Institute Research.

корисно під час роботи зі статистикою, опитуваннями або економічними звітами. Частина інструментів підтримує автоматичне оновлення даних із Google Sheets або API.

Робота з візуальним контентом. ШІ також активно використовується для підготовки візуального контенту. Наприклад, автоматична генерація субтитрів, підготовка коротких промороликів, адаптація довгих відеоматеріалів під короткий формат соцмереж.

Переклад іноземного контенту. Інструменти на базі штучного інтелекту (наприклад, DeepL або ModernMT) дають змогу отримати якісний попередній переклад текстів. Це зручно для первинного аналізу іноземних джерел або підготовки швидкого огляду. Автоматичний переклад не рекомендують використовувати без подальшого втручання редактора, особливо в разі публікації матеріалу.

Глибоке дослідження теми

Сучасні LLM, як-от GPT-4, Claude або Perplexity AI, здатні швидко підготувати огляд теми з ключовими тезами, історією питання, основними точками конфлікту, позиціями сторін, типологією дезінформації (за потреби). Наприклад:

- Скласти хронологію подій із теми.
- Виділити ключові джерела: міжнародні звіти, аналітичні центри, провідні медіа.
- Сформулювати список потенційних підтем, які можуть бути недооцінені у висвітленні.

Водночас будь-яка інформація, згенерована ШІ, не є остаточно достовірною і має бути обов'язково перевірена за першоджерелами.

Підготовка щоденних довідок або інфоаналітики

LLM можуть бути налаштовані на щоденне виконання завдань, наприклад, підготовку щоденної довідки про новини медіаграмотності, ціни на нафту, заяви щодо війни Росії проти України тощо. Усі такі матеріали потребують **вичитування та редагування**, а також обов'язкової перевірки першоджерел, адже ШІ властиві "галюцинації" та вигадування фактів.

Як почати: три кроки для впровадження автоматизації в редакції

Автоматизація редакційних процесів не вимагає повної перебудови роботи. Навпаки, починати краще з простих рішень, які можна протестувати без ризику. Ось базова стратегія, яку можуть застосувати як великі медіаорганізації, так і невеликі локальні редакції.

Крок 1. Визначте завдання, які «з'їдають» найбільше часу

Проведіть короткий внутрішній аудит: які дії у вашій редакції є рутинними, повторюваними та не вимагають складної аналітики? Це може бути транскрибування інтерв'ю, переклад пресрелізів, підготовка дайджестів, моніторинг соцмереж або зведення статистичних даних. Саме ці завдання найчастіше піддаються ефективній автоматизації.

Крок 2. Оберіть і протестуйте 1–2 інструменти без інтеграції в систему

Для старту не потрібно змінювати CMS чи наймати програмістів. Достатньо обрати кілька інструментів, якими можна почати користуватися швидко і без додаткових налаштувань. Протестуйте їх на реальних завданнях протягом кількох тижнів і оцініть зміни у витратах часу.

Крок 3. Проведіть обговорення з командою

Навіть найефективніші інструменти не мають сенсу без зворотного зв'язку. Зустріньтеся з командою та обговоріть, що спростилося, а що викликає труднощі. Чи справді використання ШІ скоротило час виконання завдань? Чи зберігається якість? Чи є потреба в доопрацюванні процесу? На основі цих відповідей можна ухвалювати рішення: масштабувати, змінювати або відмовитися.

Аналіз даних та фактчекінг

Штучний інтелект дедалі активніше інтегрується в роботу журналістів і редакцій, зокрема у сферах аналізу великих обсягів даних (big data), розпізнавання дезінформації та автоматизації фактчекінгу. Такі системи ШІ дають змогу обробляти велику кількість сторінок документів за лічені хвилини, виявляючи ключові тези, тональність, маніпулятивні наративи та кореляції між подіями.

Поява великих обсягів відкритої інформації — від реєстрів компаній до цифрових слідів у соціальних мережах — вимагає нових підходів до обробки та інтерпретації. Саме тому алгоритми штучного інтелекту дедалі частіше інтегруються у щоденну практику редакцій як інструменти для аналізу даних, пошуку зв'язків, виявлення тенденцій та формування новинного порядку денного на основі фактів.

Однією з найбільш затребуваних технологій є обробка природної мови (Natural Language Processing, NLP) — підгалузь ШІ, яка дає змогу «читати» великі обсяги тексту і виявляти в них ключові теми, зміни тональності, повторювані патерни. Аналітика даних на основі ШІ також активно використовується в журналістиці розслідувань. Один із прикладів — проєкт **Reuters**, у якому журналісти проаналізували понад 100 тисяч судових рішень США за допомогою NLP та машинного навчання. Алгоритми виявили закономірності упередженого ставлення судів до певних соціальних груп, що стало підґрунтям для серії публікацій про системну нерівність у правосудді.

У галузі цифрового споживання новин ШІ дає змогу журналістам відстежувати, які теми є «вірусними», як змінюється авдиторія і які дані викликають найбільшу зацікавленість. Видання **The Markup використовує** інструменти збирання даних про алгоритмічні стрічки соціальних платформ, щоб аналізувати, як контент поширюється серед різних груп користувачів. У межах проєкту Citizen Browser було проаналізовано мільйони інформаційних одиниць для виявлення інформаційної сегрегації у Facebook.

ШІ також використовують для «витягування» даних зі складно розпізнаваних форматів — наприклад, із фінансових звітів, PDF-файлів або таблиць. **Bloomberg** щодня створює тисячі коротких новин на основі даних фондових ринків і корпоративної звітності за допомогою системи Cyborg, яка автоматично «читає» фінансові звіти та формує текстові новини.

Інструменти для аналізу даних із відкритих джерел (OSINT) — ще один приклад використання ШІ в медіа. Організація **Bellingcat** поєднує супутникові знімки, геодані, соцмережі та автоматичне розпізнавання зображень для підтвердження подій у зонах бойових дій.

Інший напрям — генеративна журналістика на основі даних. Деякі редакції використовують ШІ для автоматичного створення звітів із наборів статистичних даних. Наприклад, **шведське інформагентство TT** розробило систему, яка автоматично генерує спортивні новини, базуючись лише на цифрах матчів. Подібні технології використовуються для виборчої аналітики та фінансових новин, де основою тексту є числа, а не цитати.

ШІ-системи демонструють здатність до надшвидкої обробки великих обсягів інформації, що є недосяжним для людських когнітивних можливостей. Така функціональність дає змогу ефективно структурувати дані й ідентифікувати пріоритетні елементи.

До прикладу, у звіті **компанії OpenAI** від травня 2024 р. «ідеться про те, що під час розслідувань щодо діяльності дезінформаційних кампаній впливу в мережі OpenAI розробляє та використовує власні моделі на основі ШІ. Технології ШІ допомогли їм ефективніше працювати багатьма мовами та суттєво покращили аналітичні можливості, скоротивши деякі робочі процеси з годин або днів до кількох хвилин»[20].

Водночас штучний інтелект додає новий вимір до проблеми дезінформації. ШІ **знижує вартість** проведення інформаційних операцій впливу, зокрема через різноманітні безкоштовно доступні інструменти з написання фейкових матеріалів, імітації голосів реальних людей і створення фотографій і відео, які майже неможливо відрізнити від справжніх, тощо.

У звіті **Всесвітнього економічного форуму у 2024 році «Глобальні ризики»** дезінформацію, створену за допомогою ШІ, назвали однією з найсерйозніших загроз світовій спільноті у коротко- та довгостроковій перспективі. Ця проблема стрімко набуває більших масштабів через використання передових технологій, які значно полегшують створення, адаптацію та поширення неправдивої інформації.

Водночас штучний інтелект відкриває нові можливості для фактчекінгу та підвищення медіаграмотності користувачів. Він може ефективно аналізувати інформацію, швидко перевіряючи достовірність контенту й допомагаючи користувачам відрізняти правдиву інформацію від оманливої.

Системи фактчекінгу на основі штучного інтелекту можуть використовувати обробку природної мови та великі мовні моделі для розуміння контексту, лінгвістичних нюансів та навіть настрою, що стоїть за висловлюваннями. Це дає змогу проводити перевірку фактів з детальнішим підходом, виходячи за межі простої класифікації «правда» чи «неправда».

Основою фактчекінгу на базі ШІ є розвинені алгоритми, здатні аналізувати великі масиви даних. Ця технологія включає обробку природної мови (NLP) для розуміння контексту, моделі машинного навчання для виявлення патернів дезінформації та складні бази даних, які зберігають достовірну інформацію для перевірки фактів. Ці системи навчаються на великих текстових масивах, намагаючись відрізняти фактичні твердження від тих, що є сумнівними або неправдивими.

Одним із перших інструментів автоматизованого фактчекінгу стала **система ClaimBuster**, розроблена Університетом Техасу в Арлінгтоні. Вона ідентифікує потенційно перевірювані твердження в політичних дебатах або виступах і відправляє їх до фактчекінгових організацій. Система використовувалася CNN, FactCheck.org та Politifact під час президентських кампаній у США.

Інший приклад — **Full Fact**, незалежна британська організація, яка розробила API для фактчекінгу в реальному часі. Їхній алгоритм відстежує політичні виступи, трансляції та статті, і якщо виявляє вже перевірене твердження, автоматично додає примітку з відповідною інформацією. Платформа також аналізує застосовувані аргументи і визначає, чи були вони раніше спростовані.

Інтеграція **технології ClaimReview**, ініційованої Google та Schema.org, дає змогу редакціям позначати фактчекінгові статті структурованими тегами. Це дає змогу пошуковим системам автоматично показувати попередження або виноски, коли користувач стикається з фейковою інформацією. Google Fact Check Explorer є прикладом того, як такі дані з багатьох редакцій збираються в єдиний глобальний інструмент.

У Франції медіаорганізація AFP Fact Check використовує алгоритми для виявлення повторюваних фейків. Наприклад, система автоматично аналізує зображення та відео, які активно поширюються в соцмережах, і порівнює їх із попередніми спростованими кейсами в базі.

---

20 Вплив штучного інтелекту на розвиток медіаграмотності. Звіт. Фільтр. IMS. ЦЕДЕМ.

Інструменти для багатомовного фактчекінгу, як-от TruLens або WeVerify, особливо важливі у глобальних кампаніях протидії дезінформації. Вони поєднують машинне перекладання, автоматичне порівняння джерел та метадані зображень.

Проєкт WeVerify, фінансований ЄС, розробив модулі для журналістів, які дають змогу перевіряти автентичність фото з воєнних регіонів, аналізувати походження матеріалів та ділитися результатами між редакціями.

Ще одним цифровим інструментом, який автоматизує процес фактчекінгу за допомогою штучного інтелекту, є **Facticity AI**. Сервіс здатен аналізувати текстовий, відео- та аудіоконтент, виявляти перевірювані твердження та зіставляти їх із достовірними джерелами інформації. У роботі Facticity AI використовуються методи обробки природної мови (NLP), машинного навчання та великі мовні моделі (LLM). Після «прочитання» або «прослуховування» контенту система визначає ключові факти, які можна перевірити, і проводить пошук підтверджень або спростувань у базах даних, новинних архівах та відкритих джерелах. Результатом є оцінка достовірності висловлювань із поясненням, на чому вона базується.

Навіть великі мовні моделі, як GPT-4, сьогодні активно використовуються для попередньої перевірки інформації. Деякі редакції, зокрема The Washington Post і Bloomberg, тестують інтеграцію ШІ в редакційні CMS-системи для первинного аналізу заяв політиків або сумнівних джерел.

З початком повномасштабної війни в Україні автоматизований фактчекінг набув критичного значення. Українська платформа **StopFake** застосовує ШІ-інструменти для аналізу текстових матеріалів російських пропагандистських медіа. Також використовують сервіси розшифрування текстів, машинного перекладу, автоматичного пошуку та аналізу фото/відео, що значно підвищує ефективність перевірки фактів. Це дає змогу щомісяця оперативно аналізувати десятки тисяч матеріалів офіційних і пропагандистських джерел, швидко реагувати на нові наративи й оперативно спростовувати їх.

Персоналізація контенту

У швидкозмінному цифровому середовищі персоналізація контенту є ключовою стратегією залучення користувачів та покращення їхнього загального досвіду. Персоналізація — це процес адаптації вмісту, формату або рекомендацій відповідно до індивідуальних вподобань, поведінки та історії взаємодії конкретного користувача. Ідеться, зокрема, про підбір новин за темами, які цікавлять читача, пріоритетне розміщення матеріалів певного формату (відео, текстів, подкастів), а також показ заголовків і рубрик, з якими людина, ймовірно, найбільше взаємодіятиме.

Оскільки люди мають орієнтуватися у величезному інформаційному просторі, адаптація контенту до конкретних уподобань і поведінки стала ключовим фактором. В основі цього трансформаційного процесу лежить інтеграція технологій штучного інтелекту. Персоналізовані системи на основі ШІ покращують користувацький досвід, сприяючи глибшій взаємодії з платформою, підвищенню задоволеності контентом і формуванню довготривалої лояльності.

Алгоритми штучного інтелекту аналізують поведінку користувачів на основі їхньої активності — кліків, пошукових запитів, часу взаємодії та маршрутів навігації. Ці дані сприяють виявленню стійких патернів дій і формуванню уявлення про індивідуальні уподобання та інтереси. На цій основі системи створюють персоналізовані рекомендації та контент, підвищуючи релевантність інформації, яку бачить користувач. Такий підхід посилює відчуття персонального досвіду, коли користувач відчуває, що платформа розуміє його потреби й реагує на них. Персоналізація на основі штучного інтелекту зазвичай використовує певну комбінацію машинного навчання (ML), обробки природної мови (NLP) та генеративного штучного інтелекту. Дані аналізуються алгоритмами штучного інтелекту, які виявляють закономірності й тенденції в поведінці користувачів.

Як правило, штучний інтелект також групує користувачів у сегменти на основі схожих характеристик та поведінки в процесі, відомому як сегментація авдиторії. Аналізуючи ці сегменти та поведінку користувачів, штучний інтелект рекомендує продукти, послуги або контент, що відповідає вподобанням і демографічним показникам користувачів.

Соціальні мережі та цифрові платформи широко застосовують алгоритми штучного інтелекту, щоб підвищити релевантність контенту. Вони аналізують історію переглядів, кліки та інші поведінкові сигнали, щоб пропонувати користувачам матеріали, які найімовірніше їх зацікавлять.

The New York Times **використовує** машинне навчання для персоналізації своєї головної сторінки для кожного окремого користувача, виділяючи статті, які можуть бути цікавими на основі їхньої історії читання та поведінки. Це допомогло компанії підвищити залученість користувачів і збільшити кількість підписок.

Медіакомпанія BBC використовує ШІ для сегментації авдиторії та персоналізації мультимедійного контенту. Проєкт **MyBBC** дає змогу створити «профіль споживання», який формує стрічку новин, подкастів і відео з урахуванням локації, часу доби, уподобань та навіть темпу перегляду контенту. Цей досвід спрямований на поглиблення довіри авдиторії шляхом надання релевантної та корисної інформації.

Ще один приклад — видання The Washington Post, яке **використовує** внутрішню ШІ-систему під назвою ModBot, що автоматично коригує порядок новин на головній сторінці залежно від інтересів читачів у реальному часі. ModBot не створює новини, але змінює їхній порядок і подання для кожного сегмента авдиторії, що істотно впливає на залучення та глибину споживання контенту.

ШІ також використовують для персоналізації форматів подання інформації. Наприклад, деякі мобільні додатки використовують голосовий ШІ (TTS — text-to-speech), щоб читати новини в стилі подкастів для тих, хто більше слухає, ніж читає. Цей підхід застосовує **медіа Quartz** — їхній додаток самостійно адаптує подання новин у вигляді коротких аудіоблоків на основі звичок користувача.

Виклики персоналізації

Однак персоналізація контенту на основі ШІ викликає і низку етичних питань. Зокрема, існує ризик створення «інформаційних бульбашок» — коли користувач отримує лише ті новини, що підтверджують його погляди, уникаючи альтернативних поглядів. Це може обмежувати об'єктивність і плюралізм медіапростору. Алгоритми, навчені «підлаштовуватися» під поведінку читача, починають фільтрувати альтернативні погляди, непопулярні теми або навіть критичну інформацію, що не узгоджується з профілем користувача. Це призводить до когнітивної ізоляції, коли людина сприймає спотворену версію реальності, підкріплену лише «зручними» джерелами.

До того ж, ШІ-системи персоналізації здебільшого є «чорними скриньками»: вони не пояснюють, чому саме той чи той контент був рекомендований, що ускладнює редакціям контроль над процесом і позбавляє авдиторію можливості критично оцінювати свою інформаційну стрічку. Як наслідок, редакційна відповідальність зміщується до алгоритмів, які не мають етичних норм, а працюють на оптимізацію метрик: кліків, часу взаємодії, конверсій. Це створює ситуацію, коли користувачеві показують не найважливіше чи найправдивіше, а найпривабливіше для утримання його уваги.

Попри наявні ризики та етичні дилеми, персоналізація контенту на основі ШІ залишається одним із найвпливовіших напрямів трансформації сучасної журналістики. Вона змінює не лише інформаційне наповнення, яке бачить користувач, а й саму логіку взаємодії з новинами — від формату, стилю і часу подання до індивідуального глибокого залучення. У цій новій екосистемі саме алгоритми визначають інформаційний ландшафт для кожного окремого читача, формуючи персоналізований досвід споживання новин, який дедалі більше впливає на наші погляди, поведінку та розуміння світу.

Інструменти моніторингу та аналітики

Зміни у швидкості поширення інформації, обсягах цифрових даних та кількості платформ для комунікації радикально ускладнили завдання медіа — не лише висвітлювати події, а й вчасно їх виявляти, відстежувати динаміку й аналізувати вплив. У таких умовах ручний моніторинг інформаційного поля втрачає ефективність, і журналісти дедалі частіше звертаються до інструментів на основі штучного інтелекту, які здатні працювати в режимі реального часу з великими обсягами розрізнених даних.

ШІ-технології дають змогу виявляти новинні сигнали ще до того, як вони потраплять у традиційні стрічки медіа, відстежувати поширення наративів у соцмережах, аналізувати зображення та документи, а також автоматично класифікувати контент за джерелами, тоном і потенційним впливом. Це розширює можливості редакційної роботи не лише у щоденному новинному потоці, а й у довгострокових розслідуваннях, кризових ситуаціях чи під час аналізу інформаційних кампаній.

Одним із найвідоміших інструментів є NewsWhip — платформа, яка використовує ШІ для моніторингу новинних трендів і активності в соціальних мережах. Система оцінює, які теми набирають популярності, аналізує динаміку поширення контенту та прогнозує, що стане вірусним. Журналісти використовують **NewsWhip**, щоб оперативно реагувати на події або готувати матеріали «на злеті» тренду.

У сфері аналітики інформаційних кампаній варто виділити **інструмент Pulsar**, який об'єднує соціальну аналітику, NLP-модулі та візуалізацію взаємозв'язків. Платформа дає змогу не лише моніторити події, а й зрозуміти, які наративи та ключові повідомлення поширюються в авдиторіях. Pulsar використовують як журналісти, так і аналітичні центри для дослідження дезінформаційних атак або моніторингу реакції на суспільно важливі теми.

У регіональному та локальному моніторингу ШІ також має значення. Платформи, як-от **Meltwater**, дають змогу журналістам відстежувати медіа в понад 100 країнах, аналізувати тональність новин, виявляти повторювані патерни у висвітленні тем та знаходити джерела дезінформації. Алгоритми класифікують повідомлення за географією, впливом, охопленням і дають редакціям можливість побачити «повітряну карту» інформаційного поля.

ШІ-технології також дають змогу аналізувати метадані і візуальний контент. Наприклад, **Google Pinpoint**, створений спеціально для журналістів, допомагає знаходити ключові фрази, імена, організації та теми в масивах документів, PDF-файлів, сканів і аудіозаписів. Редакції використовують його у великих розслідуваннях, де обсяг джерел перевищує можливості ручного опрацювання. Сервіс також автоматично розпізнає мову в аудіо- та відеоконтенті, перетворюючи її на текст для подальшого аналізу й пошуку, що значно спрощує роботу з неструктурованими даними.

Інструмент Data Miner допомагає журналістам швидко отримувати структуровану інформацію з вебсторінок, що зручно для розслідувань або збирання даних.

Інструменти моніторингу та аналітики на основі ШІ стали критично важливими для сучасної журналістики, яка працює в умовах надшвидкого інформаційного обігу та зростання дезінформаційних загроз. Вони дають змогу не лише оперативно відстежувати теми, що набирають обертів, а й глибше розуміти механізми поширення контенту, його джерела та емоційний вплив на авдиторію. Алгоритми ШІ знімають із журналістів навантаження щодо рутинної обробки даних, даючи більше часу для аналізу, перевірки та креативності.

У поєднанні з критичним мисленням і редакційними стандартами, ці технології трансформують не лише інструментарій, а й саму логіку новинного виробництва. ШІ дає змогу журналістам працювати швидше й точніше.

Розділ 3. Етичні аспекти використання ШІ в медіа

- Коли говорити «так», а коли «ні» штучному інтелекту
- Забезпечення прозорості у використанні ШІ
- Мінімізація упереджень і дискримінації
- Чому людина завжди має залишатися головною
- Баланс між автоматизацією та креативністю

Коли говорити «так», а коли «ні» штучному інтелекту

Системи штучного інтелекту дедалі частіше використовуються для автоматизації ухвалення рішень. Водночас із поширенням ШІ-рішень зростає потреба в розробленні прозорих систем ухвалення рішень, щоб уникнути упередженості, дезінформації або неетичного впливу на авдиторію. Перший критичний крок — оцінювання ризиків перед упровадженням таких систем.

У цій сфері все частіше застосовуються **AI Impact Assessments** — оцінки впливу алгоритмів, зокрема в контексті захисту прав користувачів та редакційної доброчесності. Наприклад, якщо онлайн-медіа запускає ШІ-модуль, що автоматично визначає, які новини показати на головній сторінці, необхідно проаналізувати, чи не призводить це до надмірного посилення «інформаційних бульбашок» або системного замовчування соціально важливих тем.

Для виявлення та оцінення таких ризиків у практиці застосовуються **матриці ризиків** (**risk matrices**), що дають змогу систематизувати рівень імовірності та серйозності негативного впливу алгоритмів. Наприклад, матриця може показати, що невелика зміна алгоритму ранжування новин може призвести до значного зниження наявності регіональних тем або матеріалів незалежних редакцій. На основі таких висновків можна адаптувати параметри ШІ-моделі або ввести коригувальні правила.

Крім того, важливою практикою стає **запровадження audit trails** — журналів або логів, які фіксують усі кроки алгоритмічного прийняття рішень. Це дає змогу ретроспективно простежити, як і чому система обрала той чи інший контент, що особливо важливо у випадках скарг, підозри в упередженості або втручання в редакційну незалежність. Audit trails також дають можливість здійснювати внутрішній або зовнішній моніторинг відповідності системи етичним стандартам, вимогам конфіденційності чи стандартам інформаційної безпеки.

Наступне важливе питання — чи доцільно взагалі застосовувати ШІ у конкретному контексті. Наприклад, варто уникати автоматичного модераційного блокування новин про чутливі теми без людської перевірки. Системи можуть неправильно ідентифікувати контент як «шкідливий» або «небезпечний», особливо в умовах мовних, культурних чи політичних нюансів. Таким чином, рішення про впровадження ШІ в редакційні чи аналітичні процеси мають базуватися не лише на ефективності, а й на потенційному впливі на інформаційне середовище.

Центральним елементом відповідального використання є наявність людини в процесі ухвалення рішень (human-in-the-loop). На практиці це означає, що редактор або модератор залишається фінальним арбітром: наприклад, ШІ може автоматично згенерувати зведення новин, але публікація відбувається лише після вичитування журналістом. Це забезпечує контроль за точністю, стилістикою і фактичністю контенту, а також зменшує ризик дезінформації чи неетичних формулювань.

Забезпечення прозорості у використанні ШІ

Одним із ключових принципів відповідального використання ШІ є прозорість — тобто відкрите і зрозуміле пояснення авдиторії, де і як у вашій редакційній практиці застосовуються алгоритми та моделі, а також передбачає розуміння його етичних, правових та соціальних наслідків. Це питання набуває особливої ваги в медіа, освіті, політиці й державному управлінні, де від прозорості ШІ залежить довіра громадськості та збереження демократичних норм.

Непрозоре використання ШІ створює ризики для маніпуляцій, поширення дезінформації та підриву довіри до традиційних медіа. Наприклад, користувач може не знати, що читає текст, згенерований ШІ, або що новини у стрічці розміщені за рекомендацією алгоритму, який базується на комерційних або політичних інтересах.

Етичні стандарти ШІ від Європейської комісії

Європейська комісія розробила керівні принципи для створення етичного штучного інтелекту, сформулювавши сім фундаментальних критеріїв, що ґрунтуються на європейських моральних засадах і мають застосовуватися при створенні ШІ-технологій.

7 основних принципів етичного штучного інтелекту:

Контроль людини: ШІ-технології мають підтримувати людські здібності та сприяти ухваленню виважених рішень, захищаючи основні права громадян. Обов'язковою є участь людини в керуванні такими системами.

Технічна надійність і безпечність: ШІ-рішення повинні не тільки ефективно розв'язувати поставлені завдання, а й гарантувати безпеку користування з обов'язковим планом дій у критичних ситуаціях.

Конфіденційність і керування даними: Громадяни мають зберігати абсолютний контроль над власною інформацією, доступ до якої має бути правомірним та безпечним для користувачів.

Прозорість: Користувачі повинні отримувати повну інформацію про функціональні можливості та обмеження ШІ-систем.

Різноманітність, недискримінація і справедливість: ШІ-технології мають забезпечувати однаковий доступ для всіх категорій населення без будь-яких форм дискримінації.

Добробут суспільства і навколишнього середовища: Використання ШІ має сприяти виключно конструктивним суспільним перетворенням та збереженню довкілля.

Підзвітність: Необхідно створити ефективні інструменти для забезпечення відповідальності за функціонування ШІ-систем.

Джерело: **Ethics guidelines for trustworthy AI**

Законодавче регулювання

З 1 серпня 2024 р. почав діяти перший в історії всеосяжний законодавчий акт ЄС щодо штучного інтелекту **Artificial Intelligence Act** (далі — AI Act). Цей нормативний документ спрямований на підтримку розвитку відповідального ШІ як у Європі, так і в глобальному масштабі, гарантуючи дотримання базових прав людини, принципів безпеки та етики, а також на мінімізацію ризиків від певних ШІ-моделей. Законодавство встановлює обов'язок для створювачів і постачальників ШІ-систем інформувати користувачів про взаємодію з автоматично створеним або обробленим контентом.

За класифікацією професорів Лундського університету Швеції, існує **три рівні прозорості ШІ**:

Алгоритмічна прозорість зосереджена на поясненні логіки, процесів та алгоритмів, що використовуються системами штучного інтелекту. Вона дає уявлення про те, як працює система штучного інтелекту: які дані використовуються, як відбувається обробка, які моделі застосовуються, які є потенційні ризики, як ухвалюються рішення та враховуються будь-які фактори, що впливають на ці рішення. Цей рівень — базовий, але найскладніший для досягнення, адже сучасні моделі, особливо ті, що базуються на глибокому навчанні, є достатньо складними для розуміння їхнього функціонування. Прагнення до алгоритмічної прозорості означає створення документації, що пояснює, як були зібрані дані, чи є в них упередження і які обмеження має модель.

Прозорість взаємодії стосується комунікації між користувачами та системами штучного інтелекту. Це про те, чи знає користувач, що він має справу не з людиною, а з алгоритмом. Чи маркований контент, створений ШІ? Чи чатбот чесно вказує, що він бот? Цей рівень має вирішальне значення для довіри та етичного використання технологій. Наразі великі платформи, зокрема Google, Meta та TikTok, поступово впроваджують політики маркування відео- та аудіоконтенту, створеного ШІ. Втім, поки що весь контент не маркується автоматично, а самостійно розпізнати його користувачам важко. Тут важлива й візуальна зрозумілість — маркування має бути чітким і помітним.

Соціальна прозорість виходить за межі технічних аспектів і зосереджується на ширшому впливі систем штучного інтелекту на суспільство в цілому. Сюди входять питання економічного впливу (наприклад, автоматизація праці), соціального (вплив на доступ до послуг, розширення нерівності), політичного (використання ШІ для маніпуляцій, вплив на вибори), а також екологічного (енергоспоживання великих моделей). Цей рівень прозорості вимагає міждисциплінарного підходу, співпраці технологічних компаній, урядів, громадських організацій та науковців[21].

Серед базових елементів прозорості у використанні ШІ виділяють:

- **Позначення контенту, згенерованого ШІ**: наприклад, у журналістиці або освітньому середовищі має бути чітке маркування: «Цей текст створений за допомогою штучного інтелекту». Компанії, як-от Google і OpenAI, уже впроваджують такі позначення в контенті, створеному мовними моделями або генераторами зображень.
- **Опис алгоритмів і логіки прийняття рішень**: платформи мають пояснювати, як працюють їхні алгоритми персоналізації або модерації. Наприклад, Meta дає користувачам можливість переглянути, чому певна публікація відображається у стрічці (опція «Why am I seeing this post?»).
- **Доступність політики ШІ**: користувачі повинні мати доступ до простої й чіткої інформації про політику компанії щодо ШІ. Наприклад, **BBC оприлюднила** редакційні правила щодо використання ШІ в журналістиці. Суспільне Мовлення **відкрито опублікувало** свої принципи застосування штучного інтелекту.
- **Освітні кампанії**: оскільки не всі користувачі можуть одразу розпізнати ШІ-контент, важливо проводити кампанії з цифрової та медіаграмотності. Наприклад, ініціатива **AI Literacy Framework** від Єврокомісії допомагає впроваджувати навички критичного розуміння ШІ у школах.

Мінімізація упереджень і дискримінації

Технології штучного інтелекту схильні до створення алгоритмічних деформацій, які можуть генерувати неточні, заангажовані чи дискримінаційні рішення. Підґрунтям таких проблем часто стає відсутність збалансованих, всеосяжних та інклюзивних датасетів для тренування систем. У журналістиці, де питання об'єктивності й інклюзії особливо чутливі, використання ШІ вимагає постійного контролю за тим, аби автоматизовані рішення не відтворювали або не посилювали дискримінаційні практики.

У Звіті «Вплив штучного інтелекту на розвиток медіаграмотності», підготовленому Проєктом «Фільтр», Центром верховенства права та IMS, наведено кілька прикладів.

21 Haresamudram K.; Larsson S.; Heintz F. Three Levels of AI Transparency. 2023.

Серед **найрезонансніших випадків, пов'язаних з алгоритмічною упередженістю**, варто згадати інцидент 2016 р. з чатботом Tay від Microsoft. Цей ШІ-асистент мав удосконалювати свої комунікативні навички через спілкування у Twitter, однак протягом декількох годин інтеракції в соцмережі він почав генерувати расистські та антисемітські висловлювання, включно із запереченням Голокосту. Компанія була змушена припинити роботу бота впродовж доби та офіційно вибачитися за його поведінку. Дослідження ЮНЕСКО «**ШІ та Голокост: перепис історії? Вплив штучного інтелекту на розуміння Голокосту**» акцентує увагу на тому, що технології штучного інтелекту можуть бути використані для поширення фейків, стимулювання ворожості та спрощення багатовимірних історичних подій, що становить серйозну загрозу для збереження історичної правди[22].

ШІ також дедалі частіше застосовують для поширення пропаганди та історичних міфів. **Російська мережа «Правда»** — приклад такої практики: фейкові новинні портали, що імітують достовірні джерела, потрапляють до тренувальних вибірок мовних моделей ШІ. У результаті чатботи на основі великих мовних моделей можуть відтворювати проросійські наративи, зокрема щодо війни в Україні, посилаючись на нібито легітимні джерела. Це підриває довіру до ШІ, спотворює суспільне сприйняття подій і створює нові виклики для прозорості та перевірки контенту.

Аналітичний центр DFRLab, який у партнерстві з фінською компанією CheckFirst викрив механізми роботи цієї мережі, **зазначає**, що у світовому масштабі мережа «Правда» опублікувала понад 3,7 мільйона статей, зосереджуючись переважно на Франції, Німеччині, Україні, Молдові та Сербії. До того ж, DFRLab **створила** інтерактивну карту для кращого оцінювання впливу російської мережі на конкретні країни.

Серед інших прикладів: Amazon створив модель для відбору кандидатів на вакансії, яка виявила гендерне упередження, віддаючи перевагу чоловікам, оскільки більшість історичних резюме належала чоловікам. Відомий кейс з використанням сервісів на кшталт **Midjourney**, які можуть генерувати зображення із стереотипами (наприклад, секретар завжди жінка).

Водночас якість такого оманливого контенту буде досить високою. Адже інструменти ШІ дають змогу створювати більш переконливі наративи, ніж здатна людина, бо вони базуються на великій кількості інформації, яку людині осягнути неможливо.

Також технології ШІ, які працюють над стрічками соцмереж і агрегаторами новин, можуть пропонувати користувачам інтернету лише той контент, що узгоджується з їхніми наявними переконаннями, які можуть бути хибними. Тобто ШІ без злого наміру може маніпулювати контентом, створювати інформаційні бульбашки та обмежувати доступ до різних поглядів, що призводить лише до посилення упереджень.

Упередження у штучному інтелекті виникають через особливості даних, на яких навчається модель. Вони можуть нести ризики дискримінації або несправедливості у процесі використання ШІ в прийнятті важливих рішень.

Як виникають упередження в ШІ

Причини виникнення упереджень можуть бути різноманітні:

- Незбалансовані або необ'єктивні набори даних.
- Включення в дані стереотипів і упереджень суспільства.
- Ненавмисні помилки під час підбору навчальних даних.

22 Вплив штучного інтелекту на розвиток медіаграмотності. Звіт. Фільтр. IMS. ЦЕДЕМ.

Типові приклади упереджень

- **Гендерні упередження**: наприклад, якщо система штучного інтелекту тренується на даних про вакансії або кар'єрний розвиток за останні десятиліття, де більшість керівних посад традиційно обіймали чоловіки, то коли такий ШІ починає автоматично відбирати кандидатів — наприклад, у процесі рекрутингу або в системах рекомендацій — він може частіше пропонувати чоловіків, навіть якщо жінки мають такі самі або й кращі професійні якості.
- **Расові упередження**: системи розпізнавання облич часто помиляються при ідентифікації осіб певних етнічних груп, якщо модель навчалася на обмеженому наборі даних.
- **Вікові упередження**: алгоритми можуть дискримінувати людей певних вікових груп у процесах рекрутингу або кредитування. До прикладу, в автоматизованій системі оцінення кредитоспроможності люди похилого віку або, навпаки, дуже молоді особи можуть отримувати гірші умови, навіть якщо не мають об'єктивних проблем з фінансовою відповідальністю.

Щоб зменшити вплив упереджень у штучному інтелекті, рекомендують:

- Використовувати різноманітні і збалансовані набори даних під час навчання моделей.
- Регулярно перевіряти ШІ-системи на наявність упереджень та проводити аудит алгоритмів.
- Залучати різноманітні команди фахівців для оцінювання та перевірки результатів роботи ШІ.
- Впроваджувати прозорість алгоритмів, щоб користувачі розуміли, як саме формуються рішення моделі.

Упередження є серйозною проблемою сучасних ШІ-систем, яку потрібно активно розв'язувати, використовуючи комплексний підхід, що включає прозорість, контроль і регулярні перевірки даних та алгоритмів. Це допоможе запобігти дискримінації і забезпечити справедливість у застосуванні штучного інтелекту.

Чому людина завжди має залишатися головною

Упровадження штучного інтелекту у медіасферу має супроводжуватися збереженням критично важливого елемента — людського нагляду. Хоча алгоритми здатні автоматизувати рутинні процеси, від генерації заголовків до автоматичної модерації коментарів, остаточне рішення, особливо в редакційно чутливих контекстах, повинно залишатися за людиною. Ідеться не лише про технічну перевірку — це про відповідальність, етику та довіру до медіа як інституції.

Концепція **human-in-the-loop (HITL)** у журналістиці означає, що журналіст або редактор бере участь у кожному критичному кроці, де алгоритм може вплинути на інформаційний вибір. Наприклад, якщо ШІ пропонує ранжування матеріалів на головній сторінці, редактор має змогу переглянути, відредагувати або змінити автоматичне рішення, враховуючи суспільну цінність, актуальність та етичні наслідки публікації. Такий підхід дає змогу уникати перекосів, які можуть виникати через алгоритмічні упередження або нестачу контексту.

Важливо також закріпити відповідальність за рішення, ухвалені з участю ШІ. Якщо медіа використовує автоматизовану систему для перевірки фактів, але публікує некоректну інформацію, відповідальність за це несе не ШІ, а редакція. Наявність чіткої політики щодо того, коли і як застосовується ШІ, хто перевіряє його результати та які дії вживаються у разі помилок, — критично важлива для підтримки довіри авдиторії і прозорості роботи медіа. HITL дозволяє мінімізувати ризики у кілька способів:

- **Контроль достовірності**: редактор перевіряє факти та джерела, запропоновані ШІ, і коригує помилки. Наприклад, якщо GPT-модель пропонує короткий дайджест новин, журналіст повинен перевірити його на фактологічну точність і збалансованість.
- **Редакційний контекст**: людина в ланцюгу ухвалення рішень забезпечує відповідність контенту тональності видання, стандартам мови та стилю, політиці щодо чутливих тем (наприклад, гендер, насильство, війна).
- **Юридична й етична відповідальність**: у разі спірного або неправдивого матеріалу відповідальність несе саме редакція, а не розробник ШІ-моделі. HITL дає змогу уникнути репутаційних ризиків у ситуаціях неточної інтерпретації даних штучним інтелектом.

Окремої уваги потребує також навчання редакційного персоналу: медіафахівці мають розуміти, як працює ШІ, які його обмеження, де можливі упередження і як правильно інтерпретувати його «рекомендації». Це особливо актуально в умовах швидкого розвитку генеративних моделей, які можуть створювати візуальний чи текстовий контент, що має правдоподібний вигляд, але не відповідає дійсності.

Баланс між автоматизацією та креативністю

Через те що інструменти штучного інтелекту все глибше інтегруються в роботу редакцій, медіафахівці стикаються з новим викликом: збереження креативності в умовах дедалі ширшої автоматизації. З одного боку, ШІ дає змогу значно оптимізувати рутинні процеси, з іншого — ставить під загрозу унікальний авторський стиль, журналістське бачення та креативний підхід до висвітлення тем.

ШІ вже здатен створювати короткі новинні зведення, переписувати релізи, формувати заголовки, а також генерувати інфографіку, субтитри й навіть відео. Проте ці інструменти зазвичай працюють лише з шаблонним контентом, залишаючи поза увагою глибші культурні, соціальні та етичні нюанси.

Креативність у медіа — це не лише красива мова чи емоційне подання. Це глибоке розуміння теми, здатність сформулювати несподіваний ракурс, поставити правильні запитання. Ці якості складно автоматизувати. Наприклад, у процесі підготовки аналітичного інтерв'ю чи розслідування редактор має врахувати контекст, досвід спікерів тощо — усе це вимагає гнучкості мислення й журналістського чуття.

Баланс досягається тоді, коли автоматизація допомагає, але не підміняє роботу журналіста. Наприклад, генеративний ШІ можна використовувати як спаринг-партнера для брейнштормів — пропонувати теми, підбірки джерел чи структуру тексту. Але авторський стиль журналіста і його відповідальність за зміст повинні залишатися незмінними.

Ще одна загроза надмірної автоматизації — уніфікація контенту. Алгоритми, які генерують новини чи репортажі, часто покладаються на «усереднений стиль» — лаконічний, нейтральний, передбачуваний. З часом це може призвести до зниження медіавиразності, коли всі тексти починають звучати однаково. Саме тому творчий контроль редактора, зокрема на етапі редагування ШІ-згенерованого матеріалу, є критично необхідним.

Творча інерція — ще один виклик. Якщо журналісти регулярно використовують ШІ для генерування ідей, заголовків чи структур, вони ризикують погіршити свої навички сторителінгу, критичного мислення та авторського бачення. Згідно з **дослідженням науковців** Массачусетського технологічного інституту (MIT), використання штучного інтелекту впливає на когнітивні здібності користувачів. Хоча інструмент спрощує роботу, він водночас зменшує глибину мислення, пам'ять і креативність. Автоматизація повинна не підміняти креативність, а лише підтримувати її.

Важливо також враховувати культурні аспекти, адже ШІ-моделі тренуються на глобалізованих текстах і часто не враховують місцеві аспекти, гумор чи культурні особливості. Журналіст у такому разі виступає як «культурний модератор», який адаптує чи переписує контент відповідно до цільової авдиторії. Без цього навіть технічно коректний текст може видаватися недоречним або «неживим».

Деякі редакції розробляють спеціальні гайдлайни, що регламентують межі застосування ШІ — наприклад, дозволено використання лише для чернеток, а остаточна версія готується людиною. У The Guardian та BBC діє правило, що жоден матеріал, створений ШІ, не може бути опублікований без людської валідації й доопрацювання.

Професорка журналістики даних Нью-Йоркського університету **Мередіт Бруссар стверджує**, що «креативність та емпатія залишаються важливими для ефективної журналістики та не можуть бути повністю замінені технологіями».

Розділ 4. Що кажуть закони

- Українське законодавство про штучний інтелект
- AI Act: єдиний підхід ЄС до регулювання систем ШІ
- Досвід США, судова практика та курс на співрегулювання
- Хто відповідає за помилки штучного інтелекту
- Загальні рекомендації з правомірного використання ШІ

Стрімкий розвиток ШІ вимагає від користувачів постійного підвищення власних навичок в опануванні його інструментів, а від держав — отримання відповідей на важливі правові питання: хто є автором контенту, згенерованого за допомогою ШІ, чи підлягає він авторсько-правовій охороні, хто несе відповідальність за помилки та дезінформацію, згенеровану ШІ?

Відповіді на ці питання сьогодні перебувають на рівні дискусій між розробниками систем ШІ, державами, авторами та іншими гравцями сфери та мають бути зафіксовані в нормативно-правовому регулюванні — як на державному, так і на міжнародному рівні.

Регулювання сфери ШІ — не гальмо прогресу, натомість це важлива умова безпечного, справедливого і відповідального використання інструментів ШІ, що дасть змогу захистити та збалансувати інтереси користувачів, розробників, держав. Натомість відсутність регулювання створює простір для зловживань: використання ШІ для масового створення фейкових новин, deepfake-контенту, порушень прав авторів.

Наявність належного регулювання сфери ШІ є особливо важливою для сфери медіа, де від якості інформації залежить не лише репутація медіа, сприйняття його суспільством та збереження довіри авдиторії, а й **формування поваги до авторських прав та недопущення поширення дезінформації**:

- Під час використання інструментів ШІ журналісти мають дотримуватися авторських прав і принципу прозорості. Генеруючи контент — тексти, зображення, музику, відео, — ШІ сьогодні ставить під сумнів традиційне людиноцентричне розуміння авторського права. Хто є автором твору: ШІ, який згенерував текст за промптом журналіста, сам журналіст, або ж цей текст не має автора? І якщо на це запитання відповідь наразі відносно очевидна — автором може бути лише людина, — то залишається питання розпорядження майновими правами на результат роботи ШІ. Хто має право дозволяти або забороняти використання згенерованого контенту — розробник інструменту ШІ, компанія-власник чи користувач інструменту? Без належного правового регулювання існує правова невизначеність щодо використання такого контенту та ризик порушення прав творців.
- Іншою проблемою є те, що ШІ може генерувати помилки, дезінформацію або контент із викривленими фактами. А безконтрольне поширення медіа такого контенту та відсутність відповідальності за наслідки неприпустимі для демократичного суспільства. Адже якість і достовірність інформації є критично важливими, зокрема, в умовах війни, коли інформація стає стратегічною зброєю. Тому важливо нести відповідальність за наслідки поширення згенерованого контенту та діяти відповідно до принципу прозорості — давати авдиторії інформацію про те, чи був контент створений або змінений із застосуванням систем ШІ.

Держави в умовах швидкого розвитку технологій ШІ повинні знайти баланс — не гальмувати інновації надмірним регулюванням, але й гарантувати захист основоположних прав людини. Співпрацюючи, держави мають знайти правові відповіді на проблемні питання, які постають з активним використанням ШІ, зокрема що стосується захисту авторських прав та генерування дезінформації. Сьогодні кожна держава має право на власний розсуд вирішувати, який правовий режим застосовувати. Втім, зважаючи на питання, які постають уже сьогодні перед судами (судові процеси в США та Китаї), та регіональні ініціативи (ухвалення в ЄС AI Act, регулювання в Україні використання згенерованого ШІ контенту), можна очікувати на розвиток законодавства щодо використання ШІ. Поки ж законодавчих рамок немає, користувачі мають бути обачними та дотримуватися певних правил щодо використання інструментів ШІ.

Навіщо журналістам дотримуватися правил правового й етичного використання інструментів ШІ? Медіа несуть повну відповідальність за контент, який вони публікують, навіть якщо його згенерував ШІ. На це звертає увагу журналістів і Комісія з журналістської етики. Порушення

авторських прав, дифамація, розпалювання ворожнечі або дезінформація — усе це може призвести до судових позовів, штрафів, блокування публікацій у соцмережах. Поширення ж дезінформації, неперевіреного контенту, який згенеровано ШІ, підриває довіру авдиторії, шкодить репутації видання, сприяє поширенню фейків[23].

Українське законодавство про ШІ

Першою державою у світі, яка закріпила норми щодо охорони прав на згенеровані системами ШІ об'єкти, є Україна. Тож журналістам слід ознайомитися з цим регулюванням як мінімум для того, щоб розуміти: **якщо матеріал чи інший контент було згенеровано журналістом за допомогою ШІ, до такого матеріалу можуть застосовуватися нові положення, що визначені у статті 33 Закону України «Про авторське право і суміжні права»**. Далі детальніше розберемо зміст цих положень.

Тож стаття 33 запроваджує спеціальне право особливого роду «sui generis», яке регулює питання використання об'єктів, які згенеровані в результаті функціонування системи ШІ, та їх правову охорону.

«**Sui generis**» — це щось своєрідне, єдине у своєму роді. Тому цей набір спеціальних положень в Законі України «Про авторське право і суміжні права» захищає нові сучасні об'єкти (об'єкти, згенеровані ШІ), що відрізняються від традиційних (літературних, музичних, аудіовізуальних та інших творів).

Чому ж використання об'єктів, які згенеровані ШІ, регулюється саме Законом України «Про авторське право і суміжні права»? ШІ не має правосуб'єктності, він є лише інструментом, який виконує дії на основі алгоритмів, розроблених людьми. Але як і традиційні твори, контент, згенерований ШІ, може бути скопійований, змінений, використаний без дозволу. Закон України «Про авторське право і суміжні права» пропонує інструменти для захисту інтересів осіб, які мають правовий зв'язок із контентом, який згенерований ШІ, — розробників програм, користувачів, правовласників. Цей закон є найбільш логічним і ефективним механізмом для врегулювання нових реалій «цифрової творчості».

Оскільки контент, згенерований ШІ, широко використовується, зокрема медіа, виникає потреба визначити, хто має право розпоряджатися згенерованими об'єктами, як саме ними розпоряджатися, отримувати прибуток, дозволяти та забороняти використання.

І якщо авторське право регулює відносини навколо автора та його твору, то право особливого роду «**sui generis**» регулює відносини навколо об'єктів, що згенеровані ШІ, та суб'єктів, які зробили свої інвестиції (фінансові інвестиції, інтелектуальні зусилля, час тощо) у створення цього об'єкта.

Об'єкт, що згенерований ШІ, в Законі України «Про авторське право і суміжні права» називається «неоригінальний об'єкт, згенерований комп'ютерною програмою»[24]. Він має певні ознаки: відрізняється від наявних подібних об'єктів, утворений у результаті функціонування комп'ютерної програми без безпосередньої участі людини в його створенні. Він не вважається оригінальним твором, а отже, на нього не поширюється класичне авторське право. Натомість такий об'єкт охороняється через право особливого роду «sui generis», що передбачає інший обсяг прав та строки їхньої дії.

На неоригінальні об'єкти, створені ШІ, також поширюються винятки щодо вільного використання. Відомо, що медіа може вільно використовувати певний обсяг творів для інформування населення.

23 Комісія з журналістської етики.

24 Закон України «Про авторське право і суміжні права».

Що ж стосується неоригінальних об'єктів, то їх не можна вільно використовувати лише тому, що вони не є «творами» в традиційному сенсі. **Порушення правил використання тягне за собою відповідальність**.

Оскільки, **відповідно до норми закону**, неоригінальні об'єкти, згенеровані комп'ютерною програмою, виникають без творчого внеску людини, **особисті немайнові права** на них **не поширюються**. Тому, наприклад, використовуючи згенерований ШІ текст, **журналіст не може вимагати від редакції зазначати його автором цього тексту, навіть якщо він ініціював його створення**.

Сьогодні дискусії точаться щодо питання про подальші дії людини навколо згенерованого об'єкта: **чи може людина вважатися автором згенерованого об'єкта, якщо вона зробила значний творчий внесок в кінцевий результат?** Остаточної відповіді наразі не існує. Втім, значно поширена наступна думка: якщо людина згенерувала об'єкт за допомогою ШІ, значно доопрацювала його, креативно видозмінила, то вона може заявити про себе як про автора кінцевого результату.

Хто ж може розпоряджатися правами на згенерований об'єкт — дозволяти або забороняти його використання, отримувати дохід від розпорядження правами? Це можуть бути автори комп'ютерної програми, їхні спадкоємці, особи, яким вони передали майнові права на програму (наприклад, за договором) або правомірні користувачі програми. Право розпоряджатися дозволяє цим суб'єктам запобігати незаконному використанню інструменту ШІ або згенерованого об'єкта та отримати певну вигоду. Це справедливо, адже автор програми зробив значні інвестиції в розробку інструменту ШІ, а користувач — у генерування результату. І коли йдеться про інвестиції, то варто розуміти, що під захист підпадають не лише фінансові інвестиції, але й час, професійні, людські, технічні та інші ресурси. Простіше кажучи, якщо журналіст оформив підписку на певну програму ШІ, щоб використовувати її систематично у своїй професійній діяльності, витратив час, щоб розібратися в її функціонуванні, сформував досвід для написання промптів та згенерував потрібний йому результат, він повинен мати право його використати.

Де дізнатися про обсяг прав на згенерований результат? Частіше за все відповідь на це питання можна знайти у **публічному договорі**, який міститься на сайті інструменту ШІ та визначає умови його використання. Договір може містити дозвіл на комерційне використання згенерованого результату, але за певних умов. Наприклад, від користувача можуть вимагати маркування контенту: «Цей текст створено частково за допомогою інструменту ____ GPT-3. Після створення журналіст переглянув, відредагував і доопрацював результат на свій розсуд і несе повну відповідальність за зміст цієї публікації». Також в умовах публічного договору може бути зазначено про те, що згенерований об'єкт буде зберігатися у «пам'яті» системи ШІ та використовуватися для генерування інших об'єктів іншими користувачами.

Рекомендації щодо відповідального використання ШІ: питання права інтелектуальної власності засвідчують те, що принцип свободи договору гарантує реалізацію волі сторін договору і може нести певні ризики для користувача системи ШІ у разі включення умов, які не відповідають його інтересам. Ідеться про те, що всі **користувачі повинні детально ознайомлюватися з умовами договору (публічний договір, ліцензійний договір) та аналізувати, чи погоджуються вони на визначені в ньому зобов'язання**.

Користувачі повинні формувати промпти відповідально для того, щоб мінімізувати ризик порушення авторських прав:

- не використовувати відсилки до вже наявних об'єктів права інтелектуальної власності: творів, чужих текстів, комерційних найменувань, ТМ тощо;
- не завантажувати до системи ШІ чужі авторські зображення, відео, музичні композиції, якщо не маєте відповідних прав;

- у разі підозри про те, що результат роботи ШІ містить охоронюваний об'єкт, не використовувати його.

Тож в Україні сьогодні вже створено теоретичне підґрунтя для регулювання використання об'єктів, що згенеровані ШІ. Наступний крок — формування практики щодо того, як буде реалізовуватися право особливого роду «sui generis», як його застосовувати в різних сферах. До **позитивних аспектів такого підходу** належать «гнучкість, незарегульованість, можливість визначати зміст регулювання відповідно до існуючих потреб, особливостей системи ШІ, його стрімкого розвитку на певному етапі [25].

AI Act: єдиний підхід ЄС до регулювання систем ШІ

На міжнародному рівні немає остаточного рішення щодо того, як регулювати використання об'єктів, які згенеровані ШІ. Найімовірніше, держави досягнуть консенсусу і залишать питання використання об'єктів, які згенеровані ШІ, **на розсуд сторін та договірне врегулювання**. Водночас сьогодні виникають і інші питання на перетині тем «штучний інтелект» та «авторське право», а саме — чи можуть системи штучного інтелекту використовувати твори, які захищені авторським правом, для генерування нових об'єктів і власного навчання?

ЄС наголошує на прозорості та контролі за використанням авторського контенту в системах ШІ, що передбачені в **Artificial Intelligence Act** . Акт регулює використання систем штучного інтелекту, котрі функціонують у Європейському Союзі. Регламент пропонує запровадження єдиної правової бази для розроблення, розміщення на ринку, введення в експлуатацію та використання систем ШІ. Крім того, AI Act передбачає механізми, за допомогою яких правовласники можуть заперечувати проти використання їхніх творів для навчання систем ШІ. ШІ здатен генерувати текст, зображення та інший контент, що безумовно відкриває унікальні можливості для інновацій. Водночас такі можливості систем ШІ є викликом для художників, авторів та інших творців, а також можуть нашкодити створенню, розповсюдженню, споживанню їхнього творчого контенту.

Проблема полягає в тому, що **розробляння та навчання генеративних моделей ШІ вимагають доступу до величезних обсягів текстів**, **зображень**, **відео та інших даних**, які можуть бути як у вільному доступі, так і захищені авторським правом. Для пошуку й аналізу такого контенту можуть використовуватися методи глибинного аналізу даних. Натомість будь-яке використання даних, які захищені авторським правом, вимагає **отримання дозволу правовласника**, якщо не застосовуються відповідні винятки. Наприклад, **Директива (ЄС) 2019/790 про авторське право і суміжні права на Єдиному цифровому ринку** запроваджує винятки, що дають змогу вільно використовувати твори для цілей глибинного аналізу даних, наприклад, з метою наукових досліджень.

З такими загрозами можуть стикнутися і медіа, зокрема журналісти, створивши матеріал та опублікувавши його в мережі, тож вони мають усвідомлювати, що з розвитком ШІ є ризик того, що система ШІ може «використати» цей матеріал для навчання без дозволу автора. AI Act, зі свого боку, прагне збалансувати інтереси правовласників (зокрема авторів) та постачальників моделей ШІ і вимагає від постачальників запровадити політику дотримання законодавства ЄС про авторське право та суміжні права — **виявляти застереження правовласників, які дозволяють або забороняють використовувати їхні твори для глибинного аналізу**. А з метою підвищення прозорості даних, що використовуються для навчання моделей ШІ, постачальники таких моделей повинні складати й оприлюднювати детальний підсумок змісту (стаття 53 AI Act), який використовується для навчання, що полегшить правовласникам захист своїх прав.

Зміст AI Act важливо розуміти не лише постачальникам моделей ШІ, але й користувачам, які

25 Олена Спесивцева. Право sui generis: як штучний інтелект адаптує під себе авторське право? 23.02. 2024.

активно застосовують у своїй діяльності згенеровані ШІ об'єкти. Варто усвідомлювати, що ШІ, використовуючи авторські твори для навчання, може генерувати результат, який містить елемент захищеного твору. І якщо постачальники мають певний обсяг зобов'язань перед авторами (наприклад, оприлюднювати звіти щодо використаних для навчання творів), то користувачі, застосовуючи згенерований ШІ результат з елементами авторського твору, можуть отримати скаргу від правовласника та нести юридичну відповідальність. Тому, ознайомившись з умовами публічної угоди, яка зазвичай міститься на сайті інструменту ШІ, варто особливу увагу звернути на пункти щодо відповідальності: частіше відповідальність за наслідки використання згенерованого контенту несе лише користувач, зокрема за порушення прав третіх осіб та норм чинного законодавства (включно зі сферою авторських прав).

Саме тому користувачі мають усвідомлювати ризики та, використовуючи інструменти ШІ, дотримуватися певних правил, які знизять вірогідність порушення прав третіх осіб:

- Необхідно завжди ознайомлюватися з умовами використання інструменту ШІ (публічна угода) перед використанням результату. Зокрема, звертати увагу на те, чи можливо використовувати результат з комерційною метою.
- Варто перевіряти згенерований результат щодо того, чи містить він частину авторського твору. Зокрема, здійснювати зворотний пошук подібних результатів та перевірку на плагіат за допомогою відповідних сервісів (Google Reverse Image Search, TinEye, Copyscape, Grammarly, CAI by Adobe).
- Важливо діяти прозоро та маркувати контент, створений за допомогою ШІ. Зазначати, що ви створили текст (інший об'єкт) за допомогою інструменту ШІ (вказувати, якого саме), перевірили, відредагували і доопрацювали результат на свій розсуд і несете повну відповідальність за зміст та наслідки оприлюднення.

Досвід США, судова практика та курс на співрегулювання

США поки не мають єдиного закону про ШІ, проте працюють над регулюванням цієї сфери, зосереджуючись на інноваціях, національній безпеці та захисті прав людини. Водночас великі ІТ-компанії, зокрема Google, Meta, OpenAI і Microsoft, добровільно погоджуються дотримуватись етичних принципів, що свідчить про курс на співрегулювання разом із бізнесом.

Активно розвивається судова практика судів США, які шукають відповіді на різноманітні питання, пов'язані зі сферою ШІ. Остаточних відповідей все ще немає, судові процеси тривають, втім, цікавими є питання, які перебувають на розгляді.

У справі **SARAH ANDERSEN v. STABILITY AI** компанія використала авторські твори для навчання ШІ без дозволу авторів і без виплати їм винагороди. Суд вирішує питання, чи може ШІ навчатися на авторських творах, чи вважається використання творів для навчання ШІ копіюванням та порушенням авторських прав і як бути з виключним правом авторів дозволяти/забороняти використання своїх творів?

У справі **GETTY IMAGES ПРОТИ STABILITY AI** Getty Images подала позов проти Stability AI, оскільки вважала, що компанія Stability AI використовувала базу зображень Getty Images, що захищені авторським правом, для навчання Stable Diffusion. Суд має знайти відповідь щодо ліцензування контенту для навчання ШІ: чи треба отримувати дозвіл правовласників або ж це випадок вільного використання (для навчання системи ШІ)?

Кількість судових справ зростає та збільшує вірогідність того, що найближчим часом буде запроваджене якщо не правове регулювання, то саморегулювання використання систем ШІ, що зменшить ризик порушення авторських прав. Крім того, кожен штат може взяти ініціативу у свої руки

та ухвалити відповідні нормативно-правові межі. Наприклад, 21 березня 2024 р. у штаті Теннессі було ухвалено закон під назвою **ELVIS Act**, який має на меті захист особистих немайнових прав публічних осіб (музикантів, митців, артистів) від несанкціонованого використання їхнього голосу та зображення, зокрема за допомогою інструментів ШІ. ШІ здатен точно імітувати голоси й образи людей — зокрема, зірок, та створювати «нові пісні» померлих артистів без дозволу спадкоємців. ELVIS Act встановлює заборону на використання ШІ для створення чи поширення імітацій голосу або зовнішності без згоди власника прав і відповідальність за порушення положень.

Хто відповідає за помилки штучного інтелекту

Штучний інтелект здатен генерувати надзвичайно реалістичні тексти, зображення, відео та аудіо, які іноді важко відрізнити від справжніх навіть фахівцям. Тому ШІ активно використовують для генерування й поширення дезінформації онлайн.

Регулювання цієї сфери є складним питанням, адже заходи проти дезінформації часто межують із цензурою. Тож для визначення правил щодо відповідальності за поширення дезінформації потрібно враховувати інтереси всіх учасників — від авторів контенту, медіа до платформ та споживачів.

Враховуючи кількість контенту, який сьогодні потрапляє до інфопростору, важливо не лише забезпечувати свободу слова, але й свободу бути вільним від дезінформації. Тому медіа, здійснюючи свою діяльність, мають бути обізнані про можливості ШІ та бути особливо обережними щодо ризику поширення дезінформації.

Для мінімізації ризику поширення дезінформації, яка згенерована ШІ, та уникнення відповідальності медіа необхідно врахувати: **за будь-який опублікований медіа матеріал юридичну відповідальність несе медіа** — чи то створений контент журналістом, чи згенерований ШІ. Сьогодні ШІ є лише інструментом, вибір щодо його використання, перевірка змісту й публікація результату — це свідомі дії редакції. Контент вважається «поширеним медіа», отже — підлягає загальним нормам про дифамацію, порушення авторських прав, наклеп тощо.

Навіть за відсутності юридичних наслідків медіа несе репутаційні ризики, якщо поширює дезінформацію, згенеровану ШІ — зокрема, якщо це призводить до маніпуляцій, порушення прав людини або дискредитації осіб.

Ба більше, **маркування про те, що контент було згенеровано ШІ, не знімає відповідальності у разі настання негативних наслідків та порушення прав третіх осіб**. Умови використання інструментів ШІ обачливо вказують на те, що компанія-власниця інструменту ШІ не несе відповідальності за наслідки використання згенерованого контенту, а юридична відповідальність лягає на користувачів.

За потреби та наявності сумнівів у достовірності інформації, яка згенерована ШІ, медіа може проводити консультації з експертами у відповідній галузі.

Загальні рекомендації з правомірного використання ШІ

Наразі не існує єдиних правил щодо використання інструментів ШІ та згенерованих результатів, а участь ШІ у професійній діяльності, зокрема журналістів, стає все частішою, тож варто дотримуватися певних правил. Вони дадуть змогу мінімізувати ризик порушення прав третіх осіб (зокрема авторських прав), поширення дезінформації та притягнення до відповідальності.

1. **Дотримуватися авторських прав**. Використовуючи ШІ, медіа повинні уникати генерації контенту на основі творів, захищених авторським правом, без дозволу, публікації згенерованого контенту без перевірки на потенційний плагіат. Важливо перевіряти згенеровані зображення, тексти або аудіо за допомогою спеціальних інструментів (Google Reverse Image Search, TinEye, Copyscape тощо).

2. Маркувати контент, який згенеровано ШІ. Якщо матеріал створено або змінено за допомогою ШІ, авдиторія має бути про це поінформована. Рекомендовано зазначати: «Цей текст (зображення тощо) створено із використанням ШІ-моделі ___, відредаговано журналістом, який несе відповідальність за зміст».

3. Вивчати умови використання ШІ. Перш ніж застосовувати ШІ у створенні матеріалу, ознайомтеся з умовами публічної угоди або ліцензії інструменту. Звертайте увагу на пункти про право комерційного використання результатів, про збереження та подальше використання згенерованого об'єкта іншими, про обмеження відповідальності платформи та покладення її на користувача.

4. Пам'ятати про відповідальність за публікацію. Редакція несе повну юридичну та етичну відповідальність за оприлюднений матеріал, навіть якщо він був згенерований ШІ. ШІ — лише інструмент. Рішення про використання та публікацію ухвалює людина, яка нестиме відповідальність за наслідки.

5. Уникати поширення дезінформації. Перевіряйте факти, подані в результатах ШІ, перш ніж оприлюднювати. За сумнівів у достовірності звертайтеся до експертів або фахових організацій. Пам'ятайте: навіть маркування «створено ШІ» не звільняє від відповідальності у разі завдання шкоди або порушення прав.

6. Дотримуватися положень щодо права особливого роду «sui generis». Якщо контент не є оригінальним твором, але створений ШІ без участі людини, на нього поширюються спеціальні правила «sui generis» (відповідно до статті 33 Закону України «Про авторське право і суміжні права»). Такі об'єкти не можна вільно використовувати лише тому, що вони не мають традиційного автора.

7. Розвивати внутрішні політики. Медіа варто запровадити внутрішні політики щодо використання ШІ в роботі редакції, розподілу відповідальності між авторами та редакторами, позначення, верифікації і зберігання ШІ-результатів.

Розділ 5. Як перевіряти інформацію від ШІ

- Критичний аналіз ШІ-генерованої інформації
- Чекліст перевірки походження контенту: людина чи ШІ
- Виявлення маніпуляцій і дезінформації
- Фактчекінг та верифікація ШІ-контенту
- Розпізнавання та протидія дипфейкам

Згідно з поточним **дослідженням**, лише 36 % користувачів застосовують критичне мислення під час взаємодії з сервісами штучного інтелекту. Інші ж, замість аналізу й оцінювання отриманих відповідей, схильні сприймати їх як беззаперечну правду. А між тим, це далеко від істини. Дослідження Центру цифрової журналістики Тоу, опубліковане в **Columbia Journalism Review (CJR)** у березні 2025-го, показало, що вісім генеративних інструментів пошуку на основі штучного інтелекту не змогли правильно визначити походження уривків зі статей у 60 % випадків[26]. А коли дослідники провели експеримент та попросили ChatGPT, Copilot, Gemini та Perplexity відповісти на запитання про поточні новини, використовуючи статті BBC як джерела, з'ясувалося, що 51 % відповідей чатботів містили «суттєві проблеми». У 19 % відповідей було виявлено додавання власних фактичних помилок, 13 % цитат були або змінені, або взагалі відсутні у цитованих статтях[27].

Критичний аналіз ШІ-генерованої інформації

Відповідно до **Digital News Report**, у багатьох країнах за останні роки зменшився рівень довіри до новин, створених із використанням ШІ. Це зниження пов'язують із занепокоєнням щодо того, що ШІ може створювати контент, який є неточним, упередженим або навіть неправдивим.

Дані опитування **Reuters Institute** про ШІ в новинах у США, Великій Британії та Мексиці демонструють, що в усіх перелічених країнах лише меншість (36 %) сприймають новини, створені за допомогою ШІ, і ще менша частина (19 %) комфортно сприймають новини, створені ШІ під наглядом людини. Більшість авдиторії, як показує опитування, не замислювалася над тим, як ШІ можна використовувати для новин і зазвичай ставиться до цього з підозрою та навіть страхом.

У цьому контексті відповідальність медіа підвищується, а використання ними технологій ШІ потребує додаткової уваги з урахування ставлення авдиторії та потенційних ризиків. Навіть маркування контенту, створеного за допомогою ШІ, не змінює ставлення споживачів. Так, великого розголосу свого часу набув випадок, коли **Верховна Рада у своєму акаунті в соцмережі проілюструвала трагедію** в м. Дніпро зображенням, згенерованим ШІ. Попри те, що в опис було додано текст: «Art: Visualizations», повідомлення викликало масове обурення. Допис згодом видалили[28].

---

26 6 Jaźwińska K., Chandrasekar A. We Compared Eight AI Search Engines. They're All Bad at Citing News. Columbia Journalism Review. 06.03.2025.

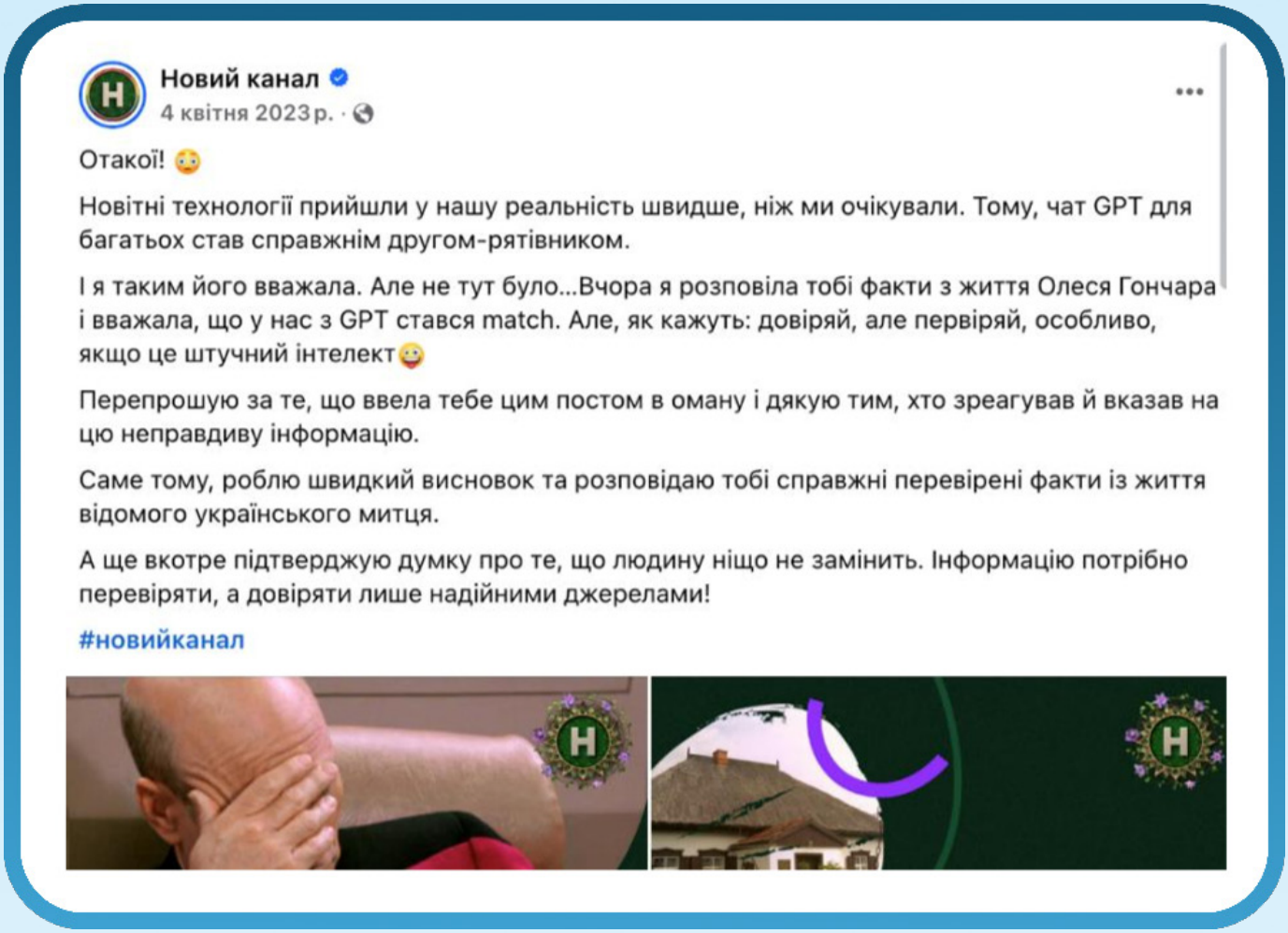
27 BBC. Representation of BBC News content in AI Assistants. BBC. February, 2025. 24 p.

28 Детектор медіа. Верховна Рада проілюструвала трагедію в Дніпрі зображенням від штучного інтелекту — в мережі обурені, допис видалили. Ms.detector.media. 16.01.2023.

Пост з акаунту Верховної Ради України від 16 січня 2023 р.

Інший приклад недалекоглядного використання ШІ в редакціях — кейс «Нового каналу». У квітні 2023 р. в інстаграм-акаунті телеканалу **з'явився допис**, присвячений 105-й річниці з дня народження українського письменника та громадського діяча Олеся Гончара. «Маловідомі факти» його життя виявилися галюцинацією штучного інтелекту, за який довелося вибачатися [29].

29 Семенюта І. «Новий канал» через ChatGPT поширив фейкову біографію Олеся Гончара». MediaSapiens. 04.04.2023.



Пост з акаунту телекомпанії «Новий канал» від 4 квітня 2023 р.

Основні виклики, які стоять перед журналістами

Широке впровадження штучного інтелекту в медіапростір створило багато нових викликів для журналістів. Сучасні редакції стикаються з проблемами верифікації контенту, боротьби з дезінформацією та необхідністю адаптації до технологій, які швидко розвиваються. Найгостріші з цих викликів можна згрупувати в кілька ключових категорій.

Масштаб. ШІ може генерувати величезні обсяги контенту за короткий час, що ускладнює ручну перевірку.

Реалістичність. Сучасні ШІ-моделі створюють контент, який часто неможливо відрізнити від реального без спеціалізованих інструментів. Та й спеціалізовані інструменти можуть помилково верифікувати штучно згенерований контент як справжній і навпаки.

Зловмисне використання. ШІ все частіше використовують для створення цілеспрямованої дезінформації, пропаганди та фейкових новин, а також для впливу на авдиторію за допомогою коментарів.

«Галюцинації» ШІ. Навіть без злого умислу ШІ-моделі можуть генерувати неправдиві або вигадані факти, посилаючись на неіснуючі джерела.

Отже, професійна компетентність сучасного журналіста обов'язково має включати здатність розпізнавати та верифікувати ШІ-контент. Медійник, який не володіє цими навичками, ризикує не лише власною репутацією, а й довірою до всієї галузі, оскільки поширення неперевіреної інформації підриває основи якісної журналістики.

Чекліст перевірки походження контенту: людина чи ШІ

Перевірка ШІ-генерованого контенту вимагає поєднання традиційних журналістських методів та новітніх технологічних підходів. Незалежно від типу контенту — текст, фото чи відео, — можемо виділити кілька основних стратегій верифікації, серед яких основні — це **логіка й точність**. ШІ-системи можуть «галюцинувати» або створювати контент, який видається правдоподібним на перший погляд, але містить логічні суперечності чи неточності при детальнішому аналізі. Наприклад, у тексті можуть бути посилання на вигадані дослідження чи організації, спотворені назви, а на зображенні — аномальні об'єкти (наприклад, листя на хвойних деревах) або нелогічне розташування об'єктів. Для текстів характерна надмірна формальність, відсутність емоцій, повторювані фрази, використання кліше або занадто досконала граматика без «людських» помилок. В аудіоматеріалах подібні ознаки проявляються через монотонність, відсутність інтонаційних нюансів або неприродні паузи.

ШІ-моделі, хоч і досконалі, часто залишають «цифрові відбитки» у згенерованому контенті. Детектори контенту на основі штучного інтелекту використовують алгоритми для аналізу лінгвальних та стилістичних особливостей, щоб виявити шаблони, типові для тексту, згенерованого штучним інтелектом. Незважаючи на наявність різних онлайн-інструментів, розроблених для цієї мети, їхня надійність часто сумнівна. Ці детектори мають труднощі з точним розрізненням контенту, згенерованого штучним інтелектом, та контенту, створеного людиною. Швидкий розвиток технологій штучного інтелекту призводить до хибнопозитивних результатів і неможливості швидко адаптуватися до нових моделей штучного інтелекту. Навіть якщо більшість цих детекторів здатні визначити, чи контент створений людиною чи штучним інтелектом, багато з них не вказують, який відсоток контенту є людським чи штучним. Якісний і добре структурований текст, який створений людиною, сервіси можуть розпізнавати як згенерований, і навпаки.

Інструменти для виявлення ШІ в тексті (пам'ятаймо, що вони часто дають хибнопозитивні результати):

Originality.AI — детектор штучного інтелекту з підтримкою широкої моделі NLP, зменшує фактичні помилки, є розширення для Chrome.

Turn itin — детектор ШІ з акцентом на виявлення недоброчесності у студентських та наукових роботах.

Winston AI — детектор ШІ, який не тільки перевіряє згенерований контент, але й виявляє перефразований та гуманізований контент, а також стратегії обходу, включно з перефразуванням контенту за допомогою таких інструментів, як Quillbot, або навіть гуманізаторів контенту на основі штучного інтелекту.

Copyleaks — детектор штучного інтелекту з перевіркою вихідного коду, позначає несанкціоноване використання та допомагає вам дотримуватися авторських прав.

Content at Scale — зосереджується виключно на сторінках вебсайтів, публікаціях у блогах, контенті соціальних мереж та маркетингових повідомленнях електронною поштою. Необхідно лише вставити свій контент і почати безкоштовне сканування. Також є перевірка на антиплагіат.

GPTZero — детектор зі штучним інтелектом та комплексною мовною підтримкою. Містить 7 компонентів, які обробляють текст, щоб визначити, чи був він написаний штучним інтелектом. Використовує багатоетапний підхід, метою якого є створення прогнозів з максимальною точністю та найменшою кількістю хибнопозитивних результатів.

Content Detector — детектор зі штучним інтелектом і перевіркою на плагіат, безлімітна кількість слів для перевірки.

Sapling — детектор на базі штучного інтелекту з можливістю встановлення окремого розширення для ChatGPT, а також аналізу PDF-файлів.

Scribbr — детектор пропонує розширені функції, такі як розрізнення контенту, написаного людиною, згенерованого штучним інтелектом та вдосконаленого штучним інтелектом, а також зворотний зв'язок на рівні абзаців для детальнішого аналізу тексту.

AI or Not — це онлайн-сервіс, який дозволяє визначити, чи було зображення створене штучним інтелектом, чи зроблене в реальності. Система аналізує зображення, використовуючи алгоритми машинного навчання, і визначає його ймовірне походження. Сервіс здатний виявляти зображення від найбільших ШІ-фотогенераторів: Midjourney, DALL-E та Stable Diffusion.

Важливо: жоден детектор ШІ не може забезпечити повну точність, оскільки мовні моделі продовжують розвиватися. Тож інструменти виявлення, а також інструменти ШІ-гуманізаторів, завжди «наздоганяють» ШІ-сервіси.

Перевірка фото

Штучний інтелект розвивається і в напрямі ідентифікації фотоконтенту, згенерованого нейромережами. Однак подібні сервіси часто «відстають» від технологій ШІ та видають релевантні результати лише за умови завантаження оригінальних фото (не скринів чи фото зменшеної якості). Цим активно користуються поширювачі дезінформації, що свідомо використовують скриншоти або спеціально погіршують якість зображень.

Далі ми пропонуємо низку інструментів для виявлення згенерованих ілюстрацій і показуємо, чи можуть ці сервіси розпізнати ШІ, якщо це скриншот згенерованого зображення. Одна зі світлин, яка стала вірусною в українському сегменті соціальних мереж, пов'язана з військовими, які сплять на снігу.

Фото, створене за допомогою ШІ

- **Aiornot.com** — сервіс виявляє зображення, текст, музику та відео, згенеровані штучним інтелектом. Наш скриншот промаркував як такий, що видається створеним за допомогою ШІ.
- **Wasitai.com** — використовує вдосконалені алгоритми для ідентифікації згенерованого контенту. Наш скриншот ідентифікував як зображення (або його частину), точно згенероване ШІ.
- **Hivemoderation.com** — ресурс, який дозволяє як генерувати контент за допомогою ШІ, так і верифікувати його. Умовно безкоштовний. Визначив зображення як таке, що згенероване ШІ.
- **Brandwell.ai** — на цьому ресурсі можна проаналізувати як зображення, так і текст, згенерований ШІ. Визначив фото як таке, що згенероване ШІ.
- **Aiimagedetector.org** — визначає, наскільки ймовірно контент згенерований ШІ чи створений людиною. У нашому випадку скриншот зображення ідентифікував з імовірністю 47 % згенерований ШІ і ймовірністю 33 % створений людиною.
- **Isitai.com** — ресурс, який має в безкоштовній версії обмеження в аналізі 15 зображень на місяць. Частину зображень, згенерованих ШІ, визначає чітко. Проте в нашому випадку ідентифікував скриншот як фото, створене людиною (90 % імовірності). Також ресурс дає змогу аналізувати текст, згенерований ШІ.
- **Illuminarty.ai** — ресурс, який найгірше впорався із завданням. Ідентифікація ШІ — всього 3,8 %. Проте з оригіналами, а не скриншотами впорується набагато краще.

Загалом сервіси з ідентифікації — це лише помічники. Вони не завжди дають правильні відповіді і схильні до хибних трактувань. Тому такі детектори не можуть бути основним методом, а в роботі із зображеннями слід використовувати різні інструменти та підходи, робити кросфактчекінг і пам'ятати, що технології з ідентифікації ШІ завжди з'являються після технологій створення ШІ.

Виявлення маніпуляцій та дезінформації

ШІ-згенерована інформація часто є інструментом інформаційних атак, джерелом дезінформації та маніпуляцій. Завдання журналіста — не лише виявити факт того, що контент згенерований ШІ, але й зрозуміти, з якою метою його створили. Для цього необхідно дотримуватися кількох етапів перевірки:

- перевірка джерел;
- фактчекінг текстів;
- аналіз зображень, аудіо і відео щодо маніпуляцій.

Аналіз достовірності джерел

Дезінформація та шкідливий контент, згенерований ШІ, часто поширюються через мережу скомпрометованих або фейкових джерел.

Як аналізувати джерело? Під час перевірки репутації джерела необхідно відповісти на такі запитання:

- Чи є джерело відомим і надійним медіа?
- Деякі джерела можуть бути створені спеціально для поширення дезінформації, тож необхідно з'ясувати, чи не було медіа раніше викрито у поширенні фейків.
- Чи є розуміння, кому належить джерело — медіа, телеграм-канал, акаунт у будь-якій соцмережі?

- Кожне джерело має певну упередженість. Важливо розуміти, як ця упередженість може впливати на представлення інформації. ШІ може посилювати наявні упередження, навчаючись на упереджених даних.
- Чи є імена і контакти редакторів, чи справжні вони?

Під час інформаційної атаки на журналістів «Бігус.інфо» маніпулятивне відео було розміщене на невідомому ресурсі «Народна правда». Хоча на сайті ресурсу було вказано контакти редакції, а також редакторів та журналістів, **ретельний аналіз показав**, що їхні фото та біографії були згенеровані ШІ [30].

Аналіз елементів фейкового медійного ресурсу. Джерело: проєкт НотаЄнота

30 Романюк А. Метод гнилого оселедця: хто і навіщо атакує журналістів. Detector.media. 16.01.2024.

Фактчекінг зазначеної на сайті інформації та перевірка фото були одним із важливих кроків, проте аналіз починається із того, що це за сайт, яку інформацію публікував раніше, чи є команда, яка наповнює ресурс.

Ресурси, які можуть допомогти у перевірці джерел:

- **Білий список медіа від Інституту масової інформації**
- **Мапа регіональних медіа від Інституту масової інформації та Детектора Медіа**
- **Список безпечних каналів отримання інформації від Центру протидії дезінформації та Кібердепартаменту СБУ**
- **Список інструментів поширення ворожої дезінформації від Центру протидії дезінформації**
- **100 російських ТГ-каналів, які мімікрують під українські від Центру стратегічних комунікацій**
- **Список каналів поширення ворожої пропаганди в соцмережі X**

Звісно, це не повний список якісних і неякісних джерел, проте вказані ресурси можна використовувати як відправну точку, від якої відштовхуватися у перевірці. Також можна звернутися до ресурсів таких фактчекінгових проєктів, як StopFake, VoxCheck чи НотаЄнота, які спеціалізуються на виявленні російської пропаганди та дезінформації в Україні.

Пошук першоджерела

Для пошуку першоджерела можна скористатися як ручним пошуком, так і автоматизованими сервісами. Наприклад, розробка українських фахівців під назвою **Osavul** дозволяє відстежувати хвилі поширення повідомлень і швидко віднаходити точку вкиду та першоджерело. Платформа дає унікальну можливість робити широкий і глибокий медіамоніторинг, з аналізом сентименту, акторів, які поширюють інформацію, та надає дані про всіх акторів і їхню афіліацію з державними та недержавними установами [31].

Також відстежувати поширення контенту допоможуть системи:

- **YouScan** — платформа для моніторингу й аналітики даних із соціальних медіа на основі ШІ. Також аналізує як тексти, так і візуалізації.
- **SemanticForce** — нейросимволічна платформа на основі штучного інтелекту (AI) для багатоканального моніторингу, розширеної аналітики та взаємодії з користувачами, містить у собі моніторинг новин, соціальних мереж, відгуків, месенджерів, цін, реклами, а також аналітику загроз у межах єдиної потужної екосистеми. Як зазначено на сайті сервісу, модуль SemanticForce DRP здатний виявити «сплановані кампанії, шкідливі домени, неавтентичну поведінку та дезінформацію».

Корисними ресурсами також стануть сервіси:

- **Domain Digger** — розширений вебінструмент, розроблений для надання детальної аналітики даних, пов'язаних із доменом.
- **Whois** — відомий багатьом медійникам ресурс, який дає змогу перевірити, на кого зареєстровано той чи інший ресурс.

31 Osavul. Профіль українського стартапу.

Фактчекінг та верифікація ШІ-контенту

Фактчекінг текстів

Робота з текстами, згенерованими ШІ, мало відрізняється від перевірки текстів загалом. Традиційно слід дотримуватися кількох правил:

- **Перехресна перевірка інформації**. Один з основних стандартів професії вимагає перевіряти інформацію з трьох незалежних і надійних джерел. Це можуть бути офіційні заяви, перевірені інформаційні агентства, академічні дослідження або свідчення безпосередніх очевидців, чиї слова підтверджені іншими доказами. Навіть коли ШІ-сервіс дає посилання на першоджерело, слід зайти на ресурс і перевірити, чи дійсно там є вказаний факт. Якщо подія, описана в тексті, не згадується великими новинними агентствами або офіційними організаціями, це привід для сумнівів.
- **Верифікація даних та статистики**. ШІ-моделі можуть легко генерувати вигадані статистичні дані, цитати з нереальних досліджень або посилання на фейкові джерела. Завжди перевіряйте джерела будь-яких цифр, графіків чи заяв. Звертайтеся до офіційних статистичних відомств (наприклад, Держстат, КМІС), наукових публікацій та досліджень у рецензованих журналах або звітів міжнародних організацій (ООН, Світовий банк тощо).
- **Аналіз цитат і заяв**. Якщо ШІ генерує текст, що містить цитати, перевірте, чи справді ці слова були сказані вказаною особою. Використовуйте пошукові системи для пошуку оригінальних заяв, стенограм виступів або офіційних пресрелізів. Навіть найменші неточності в цитуванні можуть вказувати на маніпуляцію.

У квітні 2025 р. низка українських медіа поширила цитати з інтерв'ю «Слідства.Інфо» з російським військовим, яких той не говорив. Тобто вони були вигадані невідомо ким — людиною чи штучним інтелектом. Видання «Букви» вказало, що «неточність сталася внаслідок того, що при написанні новини нами було взято за основу текстову публікацію іншого медіа без звіряння з першоджерелом, а не саме першоджерело — відео Слідство.Інфо»[32], проте багато користувачів і фахівців вважали, що контент перед публікацією опрацьовували сервіси ШІ, а журналісти не перевірили інформацію.

Автоматизована перевірка фактів (для тексту та даних)

Хоча автоматизовані інструменти не замінюють людський аналіз, вони можуть стати потужним помічником.

- **Google Fact Check Explorer** — ресурс, який дає змогу шукати інформацію в уже наявних спростуваннях фактчекерів з різних країн.
- **Originality.AI** — інструмент, який ми вже згадували вище, також допомагає перевіряти факти в режимі реального часу та уникати галюцинацій штучного інтелекту.
- **ClaimBuster** — це один із програмних засобів, який «розмічує» речення, щоб виділити твердження, які можна перевірити, та пропустити недоречні фрази.

Деякі платформи дають можливість аналізувати **тональність і стилістику тексту** — наприклад, SemanticForce та Osavul, про які йшлося вище. Хоча ШІ може імітувати різні стилі, часто можна

32 Детектор медіа. Низка медіа поширила недостовірні цитати з інтерв'ю російського солдата Слідству.Info.

виявити надмірну нейтральність, відсутність емоційних відтінків або аномальні мовні звороти, що відрізняються від типової людської мови.

Перевірка згенерованих зображень та відео

Сучасні ШІ-інструменти, що створюють графічні елементи, вдосконалюються завдяки навчанню на величезних наборах даних. З часом їхні результати виглядають дедалі правдоподібніше, тому відрізнити оригінальний контент від згенерованого стає складніше. Виявлення таких відмінностей часто перетворюється на непросте завдання.

Для роботи зі світлинами слід дотримуватися такого алгоритму:

- Уважно **дослідити зображення,** звернути увагу на найменші деталі: світло і тіні, вуха, окуляри, листя на деревах, елементи будівель тощо. ШІ достатньо розвинувся, тому жарти про шість пальців залишилися в минулому. Проте в нашвидкуруч згенерованих зображеннях часто трапляються помилки з написами.
- Збільшити фото і роздивитися все ще раз. Поставити собі питання: **що не так зі світлиною?**
- **Дізнатися, хто автор зображення**, де і коли воно було зроблено, хто опублікував зображення. Для цього можна скористатися сервісами зворотного пошуку зображень. Це може допомогти визначити, чи це оригінальне зображення/відео, чи воно було змінене, або чи було воно використано в іншому, можливо, фейковому контексті раніше.

    - **Google Images** — потужний інструмент зворотного пошуку зображень, який індексує зображення з усієї мережі.
    - **Плагін Rev Eye Reverse images search** для Google Chrome — дає змогу здійснювати пошук зображень одразу в чотирьох сервісах: Google Images, Tin Eye, Yandex та Bing.
    - **Tin Eye** — спеціалізується на зворотному пошуку зображень і надає детальну інформацію про те, де зображення було використано в інтернеті. Також є опція пошуку найбільшого за розміром зображення.
    - **PimEyes** використовує технології пошуку з розпізнаванням обличчя для виконання зворотного пошуку зображень. Розширення допомагає знайти ваші зображення в інтернеті та захистити себе від шахраїв, крадіжок особистих даних або людей, які незаконно використовують ваше фото.

- **Перевірити вихідні дані фото**. Перевірка EXIF-даних для зображень та відео може виявити аномалії, такі як невідповідність часу та дати створення, використання специфічного програмного забезпечення, що не є типовим для камер, або повну відсутність цих даних. Наприклад, відсутність GPS-координат на фото, зробленому нібито на вулиці, може бути підозрілою. Ті, хто поширює фото, можуть видаляти або змінювати метадані, їх наявність або відсутність є першим кроком. Для цього можна скористатися сервісами:

    - **ExifTool** — ресурс, який дає змогу вилучити всі метадані про завантажений або розташований в інтернеті об'єкт.
    - **Onlineexifviewer** — онлайн-переглядач EXIF створений для перегляду деталей EXIF-даних фотографій з метаданих більшості форматів фотографій, включно з файлами зображень JPEG, JPG, TIFF, PNG, WebP та HEIC.
    - **Pic2Map** — це онлайн-переглядач даних EXIF із підтримкою GPS, який допомагає знаходити та переглядати фотографії на карті. Після завантаження фото на сайт, крім надання інформації про місце і час створення фото, він покаже точку на мапі, де його зробили.

- **Перевірити світлини на помилки та шуми**. Для цього можна скористатися інструментами аналізу зображень, які спочатку використовували фотокриміналісти, а згодом почали використовувати фактчекери.

    - **FotoForensics** — надає дослідникам-початківцям та професійним слідчим доступ до передових інструментів для цифрової фотокриміналістики, дає змогу визначити рівень помилок та аналізує шуми у зображеннях, допомагає визначити маніпуляції.
    - **Forensically** — потужний сервіс для аналізу зображень, який допомагає виявляти клони, досліджувати помилки, переглядати метадані та інше.

- Перевірити геолокацію та деталі місцевості. Порівняйте видимі орієнтири, ландшафт, написи на вивісках або номери автомобілів з **Google Street View** чи пошукайте фото із вказаного місця. Важливо переконатися, що навколишнє середовище таке ж, як і на фото. Також зверніть увагу на погоду і пору року на світлині: чи збігається вона з датою знімка? Наприклад, листя на деревах, довжина тіней (можна приблизно зіставити з розташуванням сонця через **Suncalc**), одяг людей (зимовий або літній) — усе це допомагає підтвердити або спростувати заявлений контекст світлини.

- Проконсультуватися з фактчекерами. Якщо після власної перевірки залишилися сумніви, пошукайте фактчекерські матеріали чи спростування. Для цього можна використати спеціалізовані пошукові сервіси, наприклад, **Fact Check Explorer**, де можна знайти спростування за ключовими словами або зображенням. Ще один варіант — ввести опис фото у пошукові системи — можливо, інші медіа уже публікували розбір цього або подібного випадку.

- Завантажити світлини через **Google Lens** — сервіс, який допомагає ідентифікувати об'єкти за допомогою візуального аналізу на основі нейронної мережі. Також є різні мобільні додатки, які допоможуть визначити ймовірність фейку. Наприклад, Fake Image Detector або Photo Sherlock. Але їх краще використовувати як допоміжні інструменти.

Розпізнавання та протидія дипфейкам

Розпізнати дипфейк стає дедалі складніше, тому для виявлення застосовують комбінацію перевірок з різним інструментарієм, що забезпечує ідентифікацію синтетичного контенту, тобто створеного за допомогою ШІ.

Перевірка відео та зображень

- **InVID** — розширення для браузера Chrome, яке дає змогу аналізувати відео, розміщені на таких платформах як YouTube або Facebook, отримувати їхні метадані, а також автоматично розбивати відео на кадри і застосовувати до них зворотний пошук за зображеннями. Можна обрати, в яких пошукових системах шукати.

- **Sensity AI** — це одна з провідних компаній, що спеціалізується на виявленні дипфейків. Їхні інструменти використовують складні алгоритми для аналізу візуальних та аудіоартефактів.

- **Deepware Scanner** — онлайн-детектор deepfake, який сканує відео та фото на ознаки штучного втручання. Підтримує завантаження файлів або посилань із соцмереж, а результати видає у вигляді оцінки ймовірності модифікації ШІ.

- **DeepFake-o-Meter** — це онлайн-інструмент для виявлення та аналізу deepfake-контенту у відео й зображеннях. Сервіс використовує штучний інтелект для розпізнавання маніпуляцій та підробок у цифрових медіа.

- **Hive Moderation** — комерційна платформа, що спеціалізується на автоматичній перевірці контенту — зображень, відео, текстів.
- **GODDS (Global Online Deepfake Detection System)** — нова ініціатива, запущена у 2024 р. Північно-Західним університетом (США) спеціально для журналістів. Це безкоштовний сервіс, куди репортери після реєстрації можуть завантажити підозрілий файл (відео, аудіо або зображення) і отримати експертний висновок протягом 8–24 годин.

Перевірка аудіо

- **AI Voice Detector** — це сервіс на основі штучного інтелекту, який дозволяє перевіряти аудіозаписи на наявність штучних змін. Він розрізняє природну людську мову та згенеровані або відредаговані голоси, аналізуючи такі параметри, як тембр, частоти, інтонаційні особливості та сторонні артефакти.
- **DuckDuckGoose** — мовний детектор синтетичного звуку для виявлення штучно згенерованих голосів та аудіоманіпуляцій з точністю 96 %. Аналізує аудіо понад 16 мовами, щоб визначити, чи був запис створений або змінений штучним інтелектом.
- **Resemble Detect** — платформа використовує технології машинного навчання для ідентифікації синтетичного аудіоконтенту у різноманітних файлових форматах. Система здійснює миттєвий аналіз звукових записів, виявляючи ознаки штучно створеного або модифікованого голосу, що забезпечує можливість верифікації автентичності аудіоматеріалів і запобігає поширенню звукових підробок.

Варто підкреслити, що багато детекторів добре працюють лише на тих видах фейків, на яких їх навчали, і різко втрачають ефективність на нових видах підробок. Тому бажано використовувати кілька методів і не покладатися тільки на автоматичний аналіз. Сервіси варто розглядати як допоміжний «тривожний дзвіночок», що підказує, де потрібна підвищена увага, але остаточне рішення повинне спиратися на комплексну перевірку. Найкращий захист від дипфейків — це здоровий скептицизм та готовність перевіряти інформацію в кількох незалежних джерелах, особливо якщо вона здається надто неправдоподібною або емоційно зарядженою.

Розділ 6. Розробка редакційної політики щодо ШІ

- Українські та зарубіжні редакційні політики щодо ШІ
- Добірка рекомендацій з відповідального використання ШІ
- Додаток. Шаблон редакційної політики

Українські та зарубіжні редакційні політики щодо ШІ

Із розвитком нових технологій зростають і етичні ризики. Звіт **Analysis & Policy Observatory (APO)** наголошує на занепокоєнні як журналістів, так і авдиторії щодо можливих маніпуляцій і поширення дезінформації через контент, створений або відредагований ШІ. Лише невелика кількість медіа має чіткі політики щодо використання генеративного ШІ, тоді як і журналісти, і глядачі вважають прозорість щодо застосування ШІ надзвичайно важливою.

«Суспільне» та «Українська правда» стали першими в Україні, хто опублікував свої політики щодо використання ШІ. «**Українська правда**» визначила такі ключові принципи використання штучного інтелекту:

- **Генерація тексту.** «Українська правда» не публікує статті, колонки чи новини, повністю згенеровані ШІ, окрім демонстраційних випадків. ШІ допомагає з ідеями, заголовками та короткими текстами для соцмереж, але все проходить затвердження редакторами. ШІ може допомогти в генеруванні ідей для статей під час мозкового штурму. Для досліджень і аналізу ШІ використовується як додатковий інструмент.
- **Генерація перекладів.** Переклади здебільшого роблять люди, ШІ допомагає лише у підборі відповідних слів. Редагування контенту за допомогою ШІ заборонене.
- **Генерація зображень.** Зображення, створені ШІ, не замінюють фотографії, їх використовують лише для ілюстрацій або соцмереж з відповідним маркуванням. Художники застосовують ШІ для ідей, але створюють оригінальні роботи самостійно.
- **Генерація голосу.** ШІ використовують для озвучення подкастів, підготовлених редакторами. Історичні аудіозаписи не редагують синтезом голосу, щоб не спотворювати факти.

«Українська правда» прагне йти у ногу з технологіями, зберігаючи при цьому етичні та відповідальні журналістські практики. Ми визнаємо, що ШІ має обмеження, може упускати або спотворювати важливу інформацію. Тому наша робота повністю лишається людиноцентричною. Людина і тільки людина буде оцінювати ефективність інструментів штучного інтелекту та приймати рішення щодо публікації будь-якого продукту, створеного або обробленого ШІ», — вказано на сайті видання[33].

Для роботи з ШІ Суспільне керується такими принципами:

- **Прозорість і відповідальність**. Весь контент, створений за допомогою ШІ, контролюється працівниками Суспільного, які відповідають за його відповідність законам і редакційним стандартам.
- **Ефективність і доцільність**. ШІ застосовуватиметься для автоматизації транскрибування, перекладу, озвучення, субтитрування, пошуку інформації, формування чернеток, аналізу даних та перевірки фактів. Усі результати перевіряються працівниками.
- **Обмеження використання ШІ**. Не використовують інструменти ШІ, що можуть спричинити інформаційні бульбашки, упередженість, суперечать журналістським стандартам чи стратегії Суспільного, або неефективні з ресурсного погляду.
- **Пріоритет людей і захист прав**. Робота журналістів, редакторів, дизайнерів, відеографів, фотокореспондентів та інших працівників творчих підрозділів є основою Суспільного, тож інструменти ШІ будуть допоміжними.

33 Українська правда. Редакційна політика щодо AI.

- **Захист даних і використання контенту**. При зборі персональних даних Суспільне буде інформувати про це користувачів та забезпечувати захист даних в межах законодавства.
- **Дотримання світових стандартів**. Суспільне дотримується законодавства України та ЄС щодо ШІ і адаптуватиме політику до нових викликів[34].

24 канал дотримується таких правил:

- **Редакційний контроль**: Весь контент, створений ШІ, проходить перевірку людиною. ШІ-тексти вважаються лише початковим матеріалом, і відповідальність за кінцевий результат несе людина.
- **Маркування зображень**: ШІ-згенеровані ілюстрації обов'язково позначаються. Їх використовують лише для окремих типів матеріалів у розважальних рубриках.

Штучний інтелект, як зазначено на сайті видання, допомагає редакції під час «обробки інформації, пошуку ідей, формування тексту і його озвучування»[35].

The Guardian розробив три основні принципи, які визначають, як редакція буде або не буде використовувати інструменти GenAI, а саме:

- **Для користі читачів**: ШІ застосовують лише з людським контролем і там, де він справді покращує створення та поширення оригінальної журналістики. Використання ШІ в публікаціях можливе лише за згоди старшого редактора та з прозорим поясненням для читача.
- **Для підтримки місії й працівників**: The Guardian використовує ШІ для підвищення якості роботи — аналізу даних, корекції текстів, маркетингу та оптимізації бізнес-процесів, не замінюючи журналістську роботу.
- **З повагою до прав авторів**: Редакція виступає за прозорі та етичні моделі ШІ, що враховують авторські права, дозволи й чесну винагороду. Контент Guardian не може використовуватись ШІ без дотримання цих принципів.

BBC встановлює три ключові принципи щодо використання генеративного ШІ:

- **В інтересах суспільства**: BBC впроваджує ШІ, щоб посилити свою публічну місію та створювати більше користі для авдиторії. Водночас компанія прагне мінімізувати ризики — зокрема поширення дезінформації, порушення авторських прав і зниження довіри до медіа. BBC співпрацює з технологічними компаніями, регуляторами та іншими медіа, щоб зробити розвиток ШІ безпечним і прозорим.
- **У пріоритеті — талант і креативність**: BBC не замінює журналістів ШІ. Навпаки, підтримує творчу роботу редакцій і авторів. ШІ можуть використовувати лише як інструмент для пришвидшення чи покращення робочих процесів — наприклад, аналізу даних чи пошуку ідей. При цьому BBC завжди враховує права авторів і правовласників.
- **Прозорість і людський контроль**: У BBC наголошують, що рішення завжди ухвалюють люди. Якщо в контенті є фрагменти, створені ШІ, про це чесно повідомлять авдиторії. Жоден матеріал не буде повністю базуватися на ШІ без участі редакторів.

34 Принципи Суспільного щодо використання штучного інтелекту.
35 Політика використання штучного інтелекту на сайті 24 Каналу.

Wired дотримується встановлених принципів щодо використання генеративного штучного інтелекту в текстах та зображеннях:

- Редакція не публікує тексти, повністю або частково створені ШІ, а також відредаговані за допомогою ШІ.
- Видання може використовувати ШІ для пропонування заголовків або текстів для коротких публікацій у соціальних мережах.
- Журналісти можуть використовувати ШІ для генерування ідей для історій.
- ШІ може допомагати як інструмент для пошуку і узагальнення інформації, але журналісти зобов'язані перевіряти всі факти і джерела самостійно, адже ШІ робить помилки і не завжди помічає важливе.
- Редакція може публікувати зображення та відео, створені за допомогою ШІ, лише якщо в роботах є значний творчий внесок митця і вони не порушують авторських прав. Про використання ШІ завжди повідомляється.
- Видання не використовує зображення, згенеровані ШІ, як заміну стокових фотографій, щоб підтримати фотографів, для яких продаж фото — важливе джерело доходу.
- WIRED або митці, яких залучає команда, може використовувати інструменти ШІ для генерування ідей.

Associated Press встановлює стандарти для використання генеративного ШІ:

- Журналісти відповідають за точність інформації.
- Генеративний ШІ не застосовується для змін фото, відео чи аудіо; заборонено додавати або вилучати елементи.
- Не дозволяється поширювати ШІ-зображення, які можуть ввести в оману. Але якщо таке зображення є темою новини, його можна показувати, обов'язково вказуючи у підписі, що воно створене ШІ.
- Працівникам заборонено вводити конфіденційну інформацію в ШІ-інструментах.
- Журналісти мають ретельно перевіряти контент, особливо отриманий ззовні, на відсутність ШІ-генерованого матеріалу. CNET встановлює чіткі принципи щодо використання генеративного штучного інтелекту (GenAI).

CNET не використовує генеративний ШІ для створення текстів, зображень чи відео, окрім випадків, коли показують, як працює ШІ.

- Весь контент має бути правдивим, унікальним і відредагованим людьми. Якщо у матеріалі є текст, створений ШІ, на це обов'язково вказують.
- Автори отримують визнання за свою роботу, а працівників навчають ретельно перевіряти факти й правильно посилатися на джерела.
- ШІ використовують для впорядкування великої кількості інформації або для спрощення дослідницьких і організаційних завдань, але не для написання повних статей.
- Згенеровані ШІ зображення чи відео не публікують, крім демонстраційних прикладів.
- Політика постійно переглядається, а зміни у використанні ШІ повідомляються публічно.

Добірка рекомендацій з відповідального використання ШІ

Щоб ефективно впровадити етичні принципи використання штучного інтелекту в редакційній роботі, варто не лише усвідомлювати виклики і ризики, а й мати чіткий план дій у вигляді редакційної політики. Для її підготовки варто скористатися перевіреними рекомендаціями і практичними кроками від провідних міжнародних та українських організацій.

У січні 2024 р. Міністерство цифрової трансформації України у співпраці з ГО «Лабораторія цифрової безпеки», представниками Експертного комітету з питань розвитку штучного інтелекту, профільних державних органів та громадських організацій, Національною радою з питань телебачення та радіомовлення, Міністерством культури та стратегічних комунікацій і Центром демократії та верховенства права **розробило рекомендації** з відповідального використання штучного інтелекту в медіа. Мета цих рекомендацій — допомогти українським журналістам познайомитися з тим, як у світі вже використовують ШІ у медіа — зокрема, таких, як **The Guardian**, **BBC**, **Wired**, **Associated Press**, **CNET** та інших. Усі вони використовують ШІ лише як допоміжний інструмент: для аналітики, досліджень, генерації ідей, але не для повноцінного створення журналістських матеріалів. Ці медіа наполягають на людському контролі, прозорості, захисті авторських прав і недопущенні маніпуляцій. Журналістська етика, точність і відповідальність залишаються ключовими цінностями, а ШІ має підсилювати, а не замінювати професійну роботу журналістів.

В Україні також створена **дорожня карта з регулювання штучного інтелекту**. Представники профільного міністерства зазначають, що цей стратегічний документ сприятиме готовності українського бізнесу до впровадження майбутнього законодавства, схожого на європейський AI Act (правовий акт ЄС у сфері штучного інтелекту, що почав діяти з серпня 2024 року). Окрім того, ініціатива покликана підвищити цифрову грамотність суспільства у питаннях ідентифікації та запобігання ризикам, асоційованим із застосуванням ШІ-технологій.

Комісія з журналістської етики підготувала **Рекомендації з журналістської етики щодо використання штучного інтелекту для створення журналістських матеріалів**, в яких наголошує на нових етичних викликах, що з'явилися або можуть з'явитися у зв'язку з поширенням сучасних моделей штучного інтелекту. Зокрема, в зоні ризику перебувають персональні дані героїв редакційних матеріалів, якими журналістам не радять ділити під час спілкування з ШІ. Відповідно до цих **Рекомендацій** відповідальність за журналістський матеріал лежить на авторі й на редакції, навіть якщо під час підготовки цього матеріалу автор звертався по допомогу до ШІ. Також Комісія з журналістської етики радить маркувати тексти чи графічні зображення, створені за допомогою ШІ, поясненнями на зразок:

«Цей матеріал частково написаний з використанням штучного інтелекту, після чого всі факти перевірені журналістом». Або: «Текст цього матеріалу написаний штучним інтелектом і відредагований журналістом редакції». Або: «Фото / зображення згенеровано ШІ з метою захисту та безпеки джерела інформації»[36].

Благодійний фонд компанії **Thomson Reuters** «Thomson Reuters Foundation» (TRF) створив **посібник** зі створення відповідальної політики використання ШІ в редакціях. Посібник підкреслює критичну важливість людського нагляду, оскільки використання генеративного ШІ без належного контролю може призвести до публікації помилок, що завдають шкоди репутації та довірі читачів. Документ також звертає увагу на фінансові ризики: втрата довіри читачів може мати негайні й тривалі фінансові наслідки.

---

36 Рекомендації Комісії з журналістської етики щодо використання штучного інтелекту для створення журналістських матеріалів.

Міжнародна організація «Репортери без кордонів» (Reporters Without Borders) ініціювала створення **Паризької хартії з ШІ та журналістики** у листопаді 2023 р., яка стала першим міжнародним етичним стандартом для ШІ в журналістиці. Хартію підтримали 16 партнерських організацій і комісія з 32 експертів із 20 країн.

Також варто відзначити **Глобальні принципи штучного інтелекту** — стандарт, що узгоджений видавцями новинних медіа у всьому світі й визначає механізми захисту прав та інтересів виробників контенту, медіа і споживачів.

Практичні кроки для розроблення редакційної політики

При формуванні власних стандартів використання ШІ редакція має врахувати кілька ключових аспектів. Насамперед варто чітко визначити, для яких завдань дозволено застосовувати штучний інтелект, а для яких — категорично заборонено. Це може включати дозвіл на використання ШІ для перевірки орфографії та стилістичного редагування, але заборону на автоматичну генерацію цитат або фактичної інформації.

Окремо слід прописати процедури верифікації контенту, створеного або обробленого за допомогою ШІ. Кожен факт, статистичні дані чи твердження мають проходити додаткову перевірку з незалежних джерел. Редакція також повинна встановити чіткі правила щодо розкриття використання ШІ читачам — від обов'язкового маркування до детального опису того, як саме технологія була задіяна в підготовці матеріалу.

Впровадження етичних стандартів використання ШІ неможливе без належної підготовки журналістської команди. Редакції варто організувати регулярні тренінги з цифрової грамотності, які охоплюватимуть не лише технічні аспекти роботи з ШІ-інструментами, а й етичні питання їх застосування.

Особливу увагу слід приділити навчанню розпізнавання синтетичного контенту та дипфейків. Журналісти мають уміти ідентифікувати ознаки штучно створених зображень, відео чи аудіозаписів, щоб уникнути поширення недостовірної інформації. Корисним буде створення внутрішньої бази знань з прикладами якісного та неякісного використання ШІ в журналістиці.

ДОДАТОК

Шаблон редакційної політики

Політика [НАЗВА МЕДІА] щодо використання штучного інтелекту

1. ЗАГАЛЬНІ ЗАСАДИ

- Мета та сфера застосування
- Основні терміни і визначення
- Принципи відповідального використання

2. ЕТИЧНІ ПРИНЦИПИ

- Журналістська доброчесність
- Транспарентність і підзвітність
- Повага до авдиторії

3. ОПЕРАЦІЙНІ СТАНДАРТИ

- Дозволені сфери використання
- Обмежені сфери використання
- Заборонені практики

4. КОНТРОЛЬ ЯКОСТІ

- Процедури верифікації
- Редакційний нагляд
- Документування процесів

5. КОМУНІКАЦІЯ З АВДИТОРІЄЮ

- Стандарти маркування
- Форми розкриття інформації
- Оброблення запитів та скарг

6. РОЗВИТОК КОМПЕТЕНЦІЙ

- Навчальні програми
- Відповідальні особи
- Оцінка ефективності

7. УПРАВЛІННЯ РИЗИКАМИ

- Моніторинг ризиків
- Процедури реагування
- Вдосконалення процесів

8. ОНОВЛЕННЯ ПОЛІТИКИ

- Регулярність перегляду
- Процедура внесення змін
- Контроль за дотриманням

Дата затвердження: _____
Наступний перегляд: _____
Затверджено: _____

Робочий зошит самодіагностики для редакцій:

від аудиту до редакційної політики

Цей інструмент самодіагностики розроблено для медіа, які прагнуть відповідально використовувати технології штучного інтелекту у своїй роботі. Зошит допоможе вашій редакції структуровано підійти до оцінювання поточних практик, виявлення потенціалу та ризиків, а також формування власної редакційної політики щодо ШІ.

Як користуватися цим інструментом:

1. Заповніть усі розділи робочого зошита, відповідаючи на питання максимально чесно.
2. Проведіть внутрішнє обговорення результатів у команді.
3. Використовуйте отримані висновки для розробки вашої редакційної політики.

БЛОК 1: АУДИТ ПОТОЧНОГО СТАНУ

Картування ШІ-інструментів у редакції

Завдання: Створіть повний перелік усіх ШІ-технологій, які використовує ваша команда. Це допоможе зрозуміти, які інструменти ви найчастіше використовуєте і на які варто оформити підписку в разі потреби.

	Тип інструменту	Використовуємо (так/ні)	Назва/продукт	Частота використання	Мета використання
Генератори тексту					
Інструменти оброблення зображень					
Інструменти редагування відео					
Системи аналізу даних					
Інструменти фактчекінгу					
Системи модерації коментарів					
Інструменти персоналізації контенту					
Автоматизація соціальних медіа					
Інші (вкажіть)					

Оцінювання інтеграції у редакційні процеси

Відзначте етапи, де ШІ є частиною робочого процесу:

- ☐ Планування контент-стратегії
- ☐ Пошук ідей і тем для контенту
- ☐ Дослідження та збір матеріалів
- ☐ Підготовка контенту
- ☐ Редагування
- ☐ Оптимізація для пошукових систем (SEO)
- ☐ Виробництво мультимедіа
- ☐ Оптимізація для пошукових систем
- ☐ Перевірка достовірності
- ☐ Аналіз ефективності
- ☐ Інше (вкажіть): _____

Питання для рефлексії:

1. Який відсоток вашого контенту проходить через ШІ-обробку?
2. Які процеси повністю автоматизовані, а які потребують людського контролю?
3. Де ви відчуваєте найбільшу користь від ШІ?

Аналіз транспарентності

Оцініть ваш поточний рівень відкритості щодо використання ШІ:

Інформування авдиторії:

- ☐ Постійне маркування ШІ-контенту
- ☐ Ситуативне повідомлення
- ☐ Загальна декларація на сайті
- ☐ Відсутність інформування

Способи позначення:

- ☐ Спеціальні примітки до публікацій
- ☐ Інформація в метаданих публікації
- ☐ Візуальні індикатори (іконка, водяний знак тощо)
- ☐ Не маркуємо окремо
- ☐ Інше (вкажіть):_____

БЛОК 2: ТЕХНОЛОГІЧНА ГОТОВНІСТЬ

Інфраструктурний аудит

Оцініть технічну спроможність вашої організації:

Рівень готовності:

- ☐ Повна готовність (маємо всю необхідну інфраструктуру)
- ☐ Базова готовність (маємо базову інфраструктуру з деякими обмеженнями)
- ☐ Часткова готовність (потрібні істотні інвестиції)
- ☐ Початковий рівень (технічні можливості не дають змоги використовувати ШІ)

Основні виклики:

- ☐ Складності інтеграції з наявними системами
- ☐ Недостатня пропускна здатність мережі
- ☐ Проблеми сумісності з наявним програмним забезпеченням
- ☐ Відсутність технічної підтримки
- ☐ Безпекові обмеження
- ☐ Фінансові обмеження
- ☐ Інше (вкажіть): _____

Економічна модель

Яку частку бюджету ваша редакція виділяє на інструменти ШІ?

- ☐ Менше 5 % від IT-бюджету
- ☐ 5–20 % від IT-бюджету
- ☐ Більше 20 % від IT-бюджету
- ☐ Не виділяємо окремого бюджету

Який економічний ефект від використання ШІ ви спостерігаєте?

- ☐ Суттєве скорочення витрат
- ☐ Помірне скорочення витрат
- ☐ Нейтральний ефект
- ☐ Зростання витрат
- ☐ Занадто рано для висновків

БЛОК 3: ЛЮДСЬКИЙ КАПІТАЛ ТА КОМПЕТЕНЦІЇ

Оцінювання компетенцій команди

	Рівень володіння ШІ	Кількість співробітників, % від команди
Експертний		
Просунутий		
Середній		
Початковий		
Мінімальний		
Не має такого досвіду		

Стратегія розвитку компетенцій

Які заходи ви проводите для підвищення компетенцій у сфері ШІ? (позначте всі відповідні)

- ☐ Внутрішні тренінги та воркшопи
- ☐ Зовнішні освітні курси
- ☐ Самостійне навчання із заохоченням
- ☐ Обмін досвідом між колегами
- ☐ Консультації з експертами
- ☐ Не проводимо спеціальних заходів

Чи є у вашій редакції співробітники, відповідальні за впровадження та використання ШІ?

- ☐ Виділена ШІ-команда/спеціаліст
- ☐ Розподілені обов'язки між наявними ролями
- ☐ Неформальна координація
- ☐ Відсутність чіткого розподілу відповідальності

БЛОК 4: ЕТИКА ТА ДОТРИМАННЯ СТАНДАРТІВ

Етичний фреймворк

Чи має ваша редакція сформульовані етичні принципи використання ШІ?

- ☐ У редакції детально розроблена етична політика щодо ШІ
- ☐ Дотримуємося загальних етичних принципів
- ☐ Побутує лише неформальне розуміння норм
- ☐ Відсутні етичні стандарти щодо використання ШІ

Ключові етичні принципи у вашій практиці:

- ☐ Прозорість для авдиторії
- ☐ Журналістська автономність
- ☐ Редакційна незалежність
- ☐ Протидія дезінформації
- ☐ Конфіденційність даних
- ☐ Недискримінація та інклюзивність
- ☐ Повага до авторських прав
- ☐ Персональна відповідальність за результат

Правові аспекти

Наскільки добре ваша редакція обізнана з правовими нормами щодо використання ШІ?

- ☐ Обізнані зі всіма вимогами
- ☐ Розуміємо основні принципи
- ☐ Поверхово знайомі
- ☐ Відсутня правова обізнаність у редакційної команди

Як ви забезпечуєте дотримання авторських прав при використанні ШІ?

- ☐ Маємо чіткі процедури перевірки контенту
- ☐ Використовуємо лише ліцензовані інструменти та контент
- ☐ Здійснюємо періодичний аудит дотримання норм
- ☐ Спеціальні процедури відсутні

БЛОК 5: УПРАВЛІННЯ РИЗИКАМИ

Матриця ризиків

Оцініть рівень ризику за шкалою від 1 до 5, де 1 – мінімальний ризик, 5 – критичний ризик

	Тип ризику	Оцінка (1-5)	Наявність процедур мінімізації (так/ні)
Поширення дезінформації			
Порушення авторських прав			
Порушення конфіденційності даних			
Відтворення упереджень та стереотипів			
Втрата довіри авдиторії			
Зниження якості журналістики			
Залежність від технологій			
Юридичні ризики			
Етичні дилеми			
Інше (вкажіть)			

Система контролю ризиків

Механізми зниження ризиків:

- ☐ Наявність процедур мінімізації (так/ні)
- ☐ Обов'язкова людська перевірка ШІ-контенту
- ☐ Перехресна верифікація інформації
- ☐ Обмеження сфер застосування ШІ
- ☐ Тестування інструментів ШІ перед впровадженням
- ☐ Регулярний моніторинг якості
- ☐ Навчання персоналу щодо ризиків
- ☐ Документування всіх процесів

Частота перегляду процедур:

- ☐ Щомісяця
- ☐ Щокварталу
- ☐ Раз на пів року
- ☐ Щороку
- ☐ За потребою
- ☐ Не переглядаємо

БЛОК 6: РЕДАКЦІЙНА СТРАТЕГІЯ ШІ

Основоположні принципи

Сформулюйте 5 ключових принципів використання ШІ у вашій редакції:

1. _____

2. _____

3. _____

4. _____

5. _____

Сфери застосування та обмеження

Редакційна функція	Рівень використання		Специфічні умови
Генерація ідей	[ ] Активно	[ ] Обмежено	[ ] Заборонено
Дослідження та аналіз	[ ] Активно	[ ] Обмежено	[ ] Заборонено
Створення чернеток	[ ] Активно	[ ] Обмежено	[ ] Заборонено
Редагування	[ ] Активно	[ ] Обмежено	[ ] Заборонено
Фактчекінг	[ ] Активно	[ ] Обмежено	[ ] Заборонено
Мультимедіа-контент	[ ] Активно	[ ] Обмежено	[ ] Заборонено
Персоналізація	[ ] Активно	[ ] Обмежено	[ ] Заборонено
Аналітика ефективності	[ ] Активно	[ ] Обмежено	[ ] Заборонено

Стандарти якості

Процедури забезпечення якості ШІ-контенту:

1.
2.
3.
4.
5.

Комунікація з авдиторією

Принципи транспарентності:

1.
2.
3.
4.
5.

БЛОК 7: САМООЦІНКА ТА РЕКОМЕНДАЦІЇ

Загальний рівень готовності

Ваша готовність до використання ШІ:

- ☐ Високий рівень (понад 75 %)
- ☐ Середній рівень (50–75 %)
- ☐ Помірний рівень (25–49 %)
- ☐ Початковий рівень (менше 25 %)

Ваші конкурентні переваги у використанні ШІ:

1.
2.
3.

Пріоритетні напрями розвитку

Що потребує першочергової уваги:

1.
2.
3.

Конкретні кроки для підвищення ефективності використання ШІ:

1.
2.
3.
4.
5.

Словник термінів для редакційної команди

Авторське право (copyright) — сукупність прав, які захищають твори літератури, науки, мистецтва.

Автоматизація — використання інструментів ШІ або цифрових сервісів для виконання рутинних завдань без втручання людини.

Велика мовна модель (LLM, Large Language Model) — нейронна модель, натренована на масивах текстових даних, яка здатна генерувати зв'язні відповіді, підтримувати діалог, створювати пояснення.

Генеративний ШІ — підвид ШІ, що створює новий контент: текст, зображення, відео, аудіо або код. Побудований на складних моделях машинного навчання, здатен моделювати стиль, структуру та логіку людських повідомлень.

Дезінформація — свідомо неправдива інформація, яка поширюється з метою ввести в оману, вплинути на громадську думку або завдати шкоди. Журналістська модерація / редакційна перевірка — процес перевірки, уточнення і редагування контенту, створеного за допомогою ШІ, перед публікацією.

Малінформація — правдива інформація, яка поширюється з метою завдати шкоди, наприклад, публікація приватної інформації або вирвана з контексту цитата.

Місінформація — хибна або неточна інформація, поширена без наміру ввести в оману, наприклад, через необізнаність або поспіх.

Модерація / редакційна перевірка — процес перевірки, уточнення і редагування контенту, створеного за допомогою ШІ, перед публікацією.

Позначення ШІ-контенту (AI labeling) — практика маркування контенту, створеного або зміненого за допомогою ШІ, для забезпечення прозорості та уникнення маніпуляцій.

Промпт — запит або інструкція, яку користувач вводить до ШІ-системи для отримання результату. Від якості промпту залежить результат: що чіткіше завдання, то точніша відповідь.

Публічне надбання (public domain) — твори, які не охороняються авторським правом і можуть вільно використовуватися без дозволу або ліцензії.

Суміжні права — права, що належать виконавцям, продюсерам, мовникам тощо, які не створили твору, але мають правові інтереси щодо його поширення або адаптації.

Упередження ШІ (AI bias) — систематичні похибки або викривлення у результатах, що виникають через недосконалі або нерепрезентативні навчальні дані. Можуть призводити до дискримінації або відтворення стереотипів.

Фактологічна галюцинація — вигаданий або хибний факт, згенерований ШІ-моделлю, який має достовірний вигляд, але не підтверджений у реальності. Один із головних ризиків при використанні LLM у журналістиці.

Фактчекінг — процес перевірки достовірності фактів, заяв або публікацій. Один із ключових методів протидії дезінформації в журналістиці. Включає верифікацію джерел, аналіз контексту, перевірку зображень, цитат і статистики.

Штучний інтелект (ШІ) — система, здатна автономно виконувати когнітивні функції (аналіз, прогнозування, генерацію, ухвалення рішень), які раніше були притаманні лише людині. За визначенням OECD, ШІ — це система, що впливає на середовище, створюючи рекомендації, прогнози або рішення на основі отриманих даних.

Юридична відповідальність за ШІ-контент — відповідальність за наслідки публікації матеріалу, створеного або зміненого ШІ, яку зазвичай несе користувач або редакція, що ухвалює рішення про публікацію.

Корисні джерела

- **Гайд для організацій громадянського суспільства ОГС про інструменти ШІ.**
- **Закон України «Про авторське право і суміжні права» (стаття 33).**
- **Інститут масової інформації (IMI). Білий список IMI: за друге півріччя 2024 року увійшли 13 медіа. imi.org.ua. 2024.**
- **Мапа рекомендованих медіа.**
- **Ольга Петрів. Штучний інтелект та Artificial Intelligence Act: час для юридичних рамок.**
- **Рекомендації з відповідального використання штучного інтелекту у сфері медіа.**
- **Рекомендації щодо відповідального використання ШІ: питання права ІВ.**
- **Рекомендації щодо ШІ в закладах вищої освіти.**
- **Список безпечних каналів отримання інформації (станом на 11.03.2024).**
- **ШІ: можливості та виклики для медіа.**
- **AI Act. Shaping Europe's digital future.**
- **Microsoft. How to fact-check AI. Microsoft.com.**
- **GIJN (Global Investigative Journalism Network). Top Investigative Journalism Tools 2024. gijn.org. 2024.**
- **OSINT Team. AI Images Detection: A Practical Guide // osintteam.com. 2024.**
- **Digger Tools. Онлайн-інструменти для медіааналізу // digger.tools.**
- **SemanticForce. Digital Risk Protection // semanticforce.ai.**
- **Who.is. WHOIS-сервіс для доменних імен та IP-адрес // who.is.**
- **YouScan. Платформа моніторингу медіа та соцмереж // youscan.io.**

Штучний інтелект у медіа:

гайд для відповідального впровадження

Керівниця проєкту:

Ольга ПРОКОПЕНКО, менеджерка Медіапроєкту Програми підтримки ОБСЄ для України.

Авторський колектив:

Катерина ГОРСЬКА, докторка наук із соціальних комунікацій, професорка Навчально-наукового інституту журналістики КНУ імені Тараса Шевченка.

Тетяна ДУДНІК, керівниця напряму комунікацій національного проєкту з медіаграмотності «Фільтр» Міністерства культури та стратегічних комунікацій України.

Ольга КРАВЧЕНКО, керівниця національного проєкту з медіаграмотності «Фільтр» Міністерства культури та стратегічних комунікацій України.

Анна НЕСТЕРЕНКО, менеджерка з координації національного проєкту з медіаграмотності «Фільтр» Міністерства культури та стратегічних комунікацій України.

Альона РОМАНЮК, фактчекерка, головна редакторка фактчекінгового проєкту «НотаЄнота», викладачка Навчально-наукового інституту журналістики КНУ імені Тараса Шевченка.

Тарас ШЕВЧЕНКО, директор з розвитку Центру демократії та верховенства права.

Видано за сприяння Програми підтримки ОБСЄ для України. У цій публікації висловлено виключно погляди авторів. Вони не обов'язково збігаються з офіційною позицією ОБСЄ.

Дизайн та верстка — Сидоренко О., Синько А.

Літературний редактор — Ірина Мариненко

Формат 210×147.

Гарнітура Acrobat Обл. вид. арк. 2,3

Умов. друк. арк. 5