

Рекомендації з відповідального використання штучного інтелекту у сфері медіа

Ці секторальні рекомендації є одним зі складників дорожньої карти з регулювання штучного інтелекту (далі — ШІ) в Україні. Системи ШІ можуть бути корисні на різних етапах діяльності медіа — від збору й аналізу даних до створення контенту, поширення новин та взаємодії з аудиторією.

Мета рекомендацій — упровадження та використання Керівних принципів щодо відповідального впровадження систем штучного інтелекту в журналістиці, а також поширення актуальних міжнародних практик, принципів та підходів до відповідального використання систем ШІ у сфері медіа для дотримання прав людини, професійних етичних стандартів.

Визначення

Система ШІ — машинна система, яка на основі отриманих вхідних даних робить висновок щодо того, як генерувати результати (як-от прогнозування, створення контенту, надання рекомендацій або рішень), які можуть впливати на фізичне або віртуальне середовище. Різні системи ШІ відрізняються за рівнем автономності (можливості працювати без втручання людини) та адаптивності (можливості бути гнучким і підлаштовуватися під середовище) після застосування.

Генеративний ШІ — система ШІ, яка, використовуючи генеративні моделі, здатна створювати новий, оригінальний контент (тексти, зображення, аудіо- та відеоматеріали). Прикладами систем генеративного ШІ є ChatGPT, Copilot, Bard, Gemini, Midjourney, DALL-E, Stable Diffusion та інші.

Сфера застосування

Рекомендації можуть застосовуватися до будь-яких систем ШІ — як розроблених третіми особами, так і систем, що розроблені самими суб'єктами у сфері медіа чи на їх замовлення, а також які використовують журналісти, інші працівники медіа, зокрема, для:

- створення контенту за допомогою систем генеративного ШІ,
- аналізу великих масивів даних,
- адміністрування / автоматизації робочих процесів (пошук експертів, добір джерел, формування типових договорів тощо),
- управління контентом (пошук та перевірка контенту, здійснення перекладу чи транскрибування матеріалу, модерування коментарів на вебсайтах медіа, пріоритезація чи персоналізація контенту).

Основні принципи відповідального використання систем ШІ у сфері медіа

Використання та розвиток систем ШІ у сфері медіа доцільно здійснювати з урахуванням основних цінностей журналістської етики, включно з правдивістю й точністю, неупередженістю та незалежністю, ненанесенням шкоди, недискримінацією, підзвітністю, інклюзивністю, повагою до приватного життя, конфіденційністю джерел та іншими принципами, передбаченими [Кодексом етики українського журналіста](#).

Водночас упровадження та використання систем ШІ у медіа потребує використання додаткових засад діяльності. Основними засадами при використанні систем ШІ у сфері медіа, що відповідають суспільним інтересам, рекомендовано визначити:

- **відповідальне редакційне рішення** (упровадження систем ШІ за усвідомленим та зваженим рішенням редакції суб'єкта у сфері медіа та з огляду на розуміння місії такого медіа);
- **законність** (наявність необхідних прав для використання систем ШІ / публікації результатів роботи ШІ);
- **регулярне проведення оцінки ризиків**, пов'язаних з використанням систем ШІ

(проведення такої оцінки в правовій та технічній площині на всіх етапах життєвого циклу системи ШІ);

- **прозорість і зрозумілість** (розкриття інформації про використання систем ШІ та пояснення цілей і способів такого використання; розуміння джерел інформації, на базі якої працює ШІ);
- **поінформованість аудиторії** про використання систем ШІ та про природу поширеного контенту (проактивне доведення інформації про використання ШІ до відома читача / глядача / слухача; чітке розмежування автентичного та згенерованого ШІ контенту тощо);
- **конфіденційність і захист даних** (недопущення витоку персональних даних чи іншої конфіденційної інформації через використовувані системи ШІ);
- **різноманітність** (забезпечення доступу аудиторії до різностороннього контенту при застосуванні інструментів персоналізації) та **недискримінаційність**;
- **фаховий людський контроль** (забезпечення контрольної перевірки результатів роботи систем ШІ фахівцем);
- **відповідальність** (особа, яка використовує в професійній діяльності системи ШІ, відповідає за наслідки використання результатів роботи цих інструментів);
- **адаптивність** (постійне покращення принципів відповідального використання відповідно до технологічного розвитку систем ШІ та змін правового регулювання).

Впровадження систем ШІ в професійну та організаційну практику

Суб'єктам у сфері медіа, які **розробляють системи ШІ самостійно або замовляють системи ШІ** у професійних розробників спеціально під потреби цього медіа, необхідно усвідомлювати відповідальність, яка покладається на них за композицію системи, її відповідність стандартам у сфері прав людини та порядок її застосування. У цьому контексті медіа рекомендується:

- дотримуватися законодавства про захист персональних даних при розробці систем ШІ;
- здійснювати постійний моніторинг стану системи ШІ, зокрема, на предмет її здатності відповідати стандартам у сфері захисту прав людини (особливо у випадках, коли система ШІ здатна до самостійного навчання);
- пересвідчитися, що система є недискримінаційною й не спричиняє негативних наслідків для вразливих і маргіналізованих груп;
- отримувати всі необхідні юридичні права та згоди перед публікацією контенту, створеного за допомогою системи ШІ;
- повідомляти користувачів про застосування систем ШІ;
- здійснювати нагляд за порядком функціонування системи ШІ та призупиняти застосування системи ШІ в разі виявлення несправностей чи недоліків в алгоритмі та своєчасно виправляти прогалини в системі ШІ перед продовженням її застосування;
- забезпечувати аудиторію можливістю направляти скарги, якщо учасники такої аудиторії підозрюють, що система ШІ порушує їхні права.

Суб'єктам у сфері медіа, які **використовують системи ШІ, розроблені третіми особами** (зокрема такі, що перебувають у вільному доступі), рекомендується дотримуватися рекомендацій щодо моніторингу стану систем ШІ, повідомлення аудиторії про використання систем ШІ, нагляду та призупинення використання систем у разі порушень та надання можливості скажитися на порушення прав. На додаток до цього суб'єктам у сфері медіа рекомендовано:

- переконатися, що система ШІ, яка не розроблена цим суб'єктом у сфері медіа або на його замовлення, відповідає стандартам у сфері захисту прав людини;
- утримуватися від використання систем ШІ, які походять з держави-агресора або розробниками яких є громадяни держави-агресора;
- використовувати системи ШІ виключно на законних підставах з дотриманням прав інтелектуальної власності третіх осіб;
- слідкувати за повідомленнями компанії-виробника системи ШІ щодо відповідності системи стандартам у сфері прав людини, безпеки даних, а також за оновленнями системи ШІ та її відповідності етичним стандартам.

Робота із системами ШІ часто вимагає навичок, які виходять за межі наявної підготовки більшості працівників медіа. Тому суб'єктам у сфері медіа рекомендується забезпечувати навчання персоналу щодо використання систем ШІ за допомогою програм, курсів, воркшопів, які

об'єднують технічних спеціалістів і журналістів, сприяють підвищенню обізнаності щодо прав людини та професійної етики. Зокрема, освітній серіал про ШІ запропоновано на порталі [Дія.Освіта](#).

Оскільки системи ШІ потребують організаційної інфраструктури для підтримки їх належного функціонування, суб'єктам у сфері медіа рекомендується створювати таку внутрішню інфраструктуру, наймаючи або навчаючи персонал із чітко визначеними функціями в галузі ШІ (впровадження цих ролей не повинно відбуватися винятково за рахунок розвитку рутинних компетенцій у сфері ШІ серед персоналу). Навчання, зокрема, рекомендується проводити з охопленням питань ризиків при використанні систем ШІ, розроблених іншими компаніями, та детальним ознайомленням з порядком їх функціонування до їх практичного застосування в роботі.

У **додатку** до рекомендацій наведено рекомендований алгоритм вибору систем ШІ для діяльності суб'єкта у сфері медіа та оцінювання згенерованого ШІ контенту, який можна використовувати для загальної оцінки безпеки та надійності систем ШІ.

Прозорість та зрозумілість у використанні систем ШІ

Медіа рекомендується розкривати інформацію про те, коли і в який спосіб вони використовують системи ШІ. Принцип прозорості передбачає його дотримання **на всіх етапах життєвого циклу системи ШІ** та пояснення порядку функціонування системи, розкриття потенційних ризиків від використання системи ШІ для прав людини й зазначення можливих впливів такої системи. Суб'єктам у сфері медіа рекомендовано розробляти та оприлюднювати власні політики, стратегії, стандарти, керівні принципи чи інші документи, що окреслюють засади використання систем ШІ у їхній діяльності. Прикладами можуть слугувати оприлюднені принципи світових медіа: [The Guardian](#), [BBC](#), [Wired](#), [CNET](#), [Associated Press](#) тощо.

Будь-яке використання систем ШІ, що впливає на збір інформації, виготовлення, обробку або розповсюдження медіаконтенту, рекомендовано чітко позначати та доводити до відома всіх, хто споживає цю інформацію.

Усередині медіа для працівників рекомендується робити доступною інформацію про те, які системи ШІ застосовують для збору інформації, виготовлення, обробки чи розповсюдження контенту, їх призначення та цінності, які вони відображають.

Розмежування автентичного та згенерованого ШІ контенту

Генеративний ШІ полегшує поширення неправдивої інформації та дезінформації за допомогою змінених слів, фотографій, відео- чи аудіоматеріалів, включно з контентом, який може не мати ознак змін і мати достовірний вигляд.

Потреба розмежування контенту, згенерованого ШІ, може виникати у двох випадках: коли медіа використовує чужий контент і коли медіа свідомо генерує контент за допомогою систем ШІ для ілюстрування власних матеріалів.

Для уникнення випадкового використання чужого згенерованого ШІ контенту журналістам рекомендовано проявляти відповідну обачність, ретельність і скептицизм, щоб переконатися, що матеріали, на які вони посилаються, є **автентичними**.¹ Також слід пам'ятати, що в Україні контент, згенерований ШІ, є об'єктом права інтелектуальної власності (ст. 33 Закону України «Про авторське право і суміжні права»).

Якщо матеріали містять контент, згенерований з використанням систем ШІ, медіа рекомендується:

- визначати джерело оригінального контенту, здійснювати зворотний пошук зображень, щоб підтвердити їх походження та наявність / відсутність автоматизованої обробки або

¹ У розумінні цих Рекомендацій «автентичним» слід уважати контент, створений без використання систем ШІ для його генерування чи суттєвої зміни.

генерації, а також перевіряти наявність повідомлень зі схожим змістом у перевірених джерелах (зокрема, офіційні джерела, фактчекінгові ресурси тощо);

- використовувати технічні засоби, зокрема на основі ШІ, для перевірки контенту, чи є контент оригінальним, чи штучно згенерованим, щоб протидіяти поширенню дипфейків чи іншого, згенерованого в недобросовісних цілях, контенту. Прикладами таких інструментів є *Sentinel*, *FakeCatcher*, *WeVerify*, *Microsoft's Video Authenticator Tool*, виявлення дипфейків за допомогою використання техніки *Phoneme-Viseme Mismatches* тощо;
- оцінювати в кожному конкретному випадку, чи контекстуально доречним є використання згенерованого з використанням систем ШІ контенту і чи не введе використання такого контенту читачів / глядачів / слухачів в оману навіть за наявності маркування (наприклад, якщо контент стосується дуже чутливої теми й може спричинити емоційну реакцію);
- посилатися на нього або поширювати такий контент за дотримання принципу прозорості, маркуючи його відповідно, якщо контент згенерований без мети завдання шкоди певному легітимному інтересу;
- уникати перепоширення контенту, який згенерований з використанням систем ШІ з метою введення в оману, поширення дезінформації чи нелегального контенту (як-от закликів до насилля, мови ворожнечі, дискримінації, пропаганди війни тощо);
- у разі поширення контенту, згенерованого третіми особами з використанням систем ШІ, дотримуватися вимог щодо захисту прав інтелектуальної власності таких осіб.

Контент, що не є автентичним та не має маркування щодо його згенерованості ШІ, рекомендується розглядати як такий, що може ввести в оману. Якщо журналісти чи інші медіаучасники мають будь-які сумніви щодо автентичності матеріалу, не рекомендується його використовувати.

У випадках, коли медіа **вдаються до генерування контенту** за допомогою ШІ, зокрема для ілюстрування матеріалів у сфері культури, моделювання ситуацій, унаочнення прикладів та теорій, їм рекомендується:

- перевіряти, чи контент, згенерований системою ШІ, чітко й достовірно відображає дійсність, посилається на наявні та надійні джерела, оперує актуальною інформацією;
- оцінювати в кожному конкретному випадку, чи контекстуально доречним є використання згенерованого з використанням систем ШІ контенту і чи не введе використання такого контенту читачів / глядачів / слухачів в оману навіть за наявності маркування (наприклад, якщо контент стосується дуже чутливої теми й може спричинити емоційну реакцію);
- утримуватися від використання чи розповсюдження новинного та новинно-аналітичного контенту, створеного ШІ, що імітує запис зйомки реального світу чи реальних осіб; утримуватися від використання генеративного ШІ для додавання або вилучення будь-яких елементів автентичних фотографій, відео- та аудіоматеріалів, за винятком випадків, коли це необхідно для забезпечення прав людини, захисту прав третіх осіб або дотримання інших вимог законодавства;
- поширювати згенерований з використанням систем ШІ контент з дотриманням принципу прозорості, маркуючи його відповідно.

Важливим є питання забезпечення чіткого розмежування між контентом, отриманим від фізичної фіксації реального світу (наприклад, фотографія, аудіо- та відеозапис), та тим, що генерується або зазнає суттєвих змін за допомогою систем ШІ. Варто віддавати перевагу використанню автентичних кадрів і записів для відображення реальних подій, якщо таке використання не порушує права людини або чинне законодавство.

Суб'єктам у сфері медіа рекомендується, за можливості, використовувати інструменти, які підтверджують автентичність опублікованого контенту, надаючи достовірну інформацію про його походження та будь-які подальші зміни, яких він міг зазнати. Наприклад, британське агентство новин Reuters [імплементувало](#) інструмент, розроблений Ініціативою щодо автентичності контенту (Content Authenticity Initiative), який дає змогу підтверджувати автентичність контенту у фотожурналістиці за допомогою закріплення в підписаних цифровим підписом метаданих файлу інформації про його походження та історію змін.

Маркування згенерованого ШІ контенту

У випадку, коли має місце публікація згенерованого ШІ контенту, — ілюстрація, текст чи інший об'єкт мистецтва, створений за допомогою ШІ, — критично важливим є чітке маркування такого контенту.

Є низка способів, як медіа можуть прозоро поінформувати аудиторію про використання ШІ: водяні знаки, написи чи інші позначення, щоб чітко ідентифікувати штучно створений контент. Наприклад, використовувати напис *«Матеріал згенеровано з використанням штучного інтелекту»* в репортажах про ілюстрації або об'єкти мистецтва, створені ШІ, або в матеріалах, які були створені ШІ.

На додаток до прямого маркування в матеріалах, медіа рекомендується запропонувати ще більшу прозорість, використовуючи бічну інформаційну панель на онлайн-платформах з даними про використовувані системи ШІ, ціль та способи такого використання, оприлюднення технічної документації або технічного репозиторію для пояснення використаних інструментів ШІ.

Пріоритет людського фактору

Системи ШІ можуть використовувати для виконання високоавтоматизованих завдань у межах щоденних робочих процесів, звільняючи час і ресурси для інших видів діяльності, водночас ухвалення остаточних рішень людиною має залишатися центральним як для довгострокових стратегій, так і для щоденних редакційних завдань. Щоб уникнути некоректних або упереджених процесів / результатів використання систем ШІ, рекомендується залишати кінцевий фаховий контроль за людиною.

Важливим є чітке визначення редакційними командами цілей, сфери застосування та умов використання для кожної системи ШІ. Водночас доцільним є проведення наскрізного й безперервного нагляду за впливом розгорнутих систем ШІ, забезпечення їх відповідності умовам та цілям використання, а також зберігання можливості деактивувати їх у будь-який момент.

Фаховий контроль є критичним для завдань, де результати є особливо чутливими (наприклад, ті, що мають конкретні наслідки для особи) або суспільно вагомими (наприклад, зачіпають інтереси національної безпеки та оборони). Медіа варто враховувати, який вплив на громадську думку або емоційний стан аудиторії може мати використання систем ШІ.

Персоналізація медіаконтенту та захист персональних даних

Розробляючи та використовуючи системи ШІ для автоматичної персоналізації та рекомендацій контенту, слід керуватися журналістською етикою. Персоналізація контенту з використанням ШІ та система рекомендацій на основі ШІ мають ґрунтуватися на принципах різноманітності, плюралізму думок і цілісності інформації та з дотриманням законодавства про захист персональних даних. Зокрема, механізми персоналізації та пріоритезації контенту не слід використовувати для поширення прихованої передвиборчої агітації. Рекомендується використовувати в таких механізмах запобіжники щодо пропонування потенційно шкідливого, небезпечного або чутливого контенту (орієнтованого на аудиторії 16+ чи 18+) неповнолітній аудиторії.

Використання таких систем ШІ рекомендується робити прозорим: аудиторії слід регулярно нагадувати про те, що певний новинний контент є персоналізованим, як і чому здійснено таку персоналізацію. Це означає надання особі необхідної інформації про її персональні дані, які використовують для персоналізації; забезпечення можливості контролювати свої дані для коригування особистого профілю та вибору між різними способами персоналізації, які враховують короткострокові та довгострокові інтереси особи.

Рекомендується забезпечити користувачам можливість вимкнути персоналізацію контенту системами ШІ, щоб забезпечити нефільтрований доступ до медіаконтенту.

Відповідальність за використання згенерованого ШІ контенту

Суб'єктам у сфері медіа варто пам'ятати про редакційну відповідальність за використання ШІ для збору, обробки або поширення інформації. Відповідно до [Рекомендацій Комісії з журналістської етики щодо використання штучного інтелекту для створення журналістських матеріалів](#) відповідальність за журналістський матеріал лежить на авторі й на редакції, навіть якщо під час підготовки цього матеріалу автор звертався за допомогою до ШІ.

Будь-який результат роботи генеративного ШІ рекомендовано розглядати як матеріал, що потребує обов'язкової додаткової перевірки та участі людини в ухваленні рішення про його поширення (оприлюднення).

Рекомендується внести принципи застосування систем ШІ в медіа до редакційних статутів, а також розробити внутрішні правила етичного застосування конкретних систем ШІ (як у медійній роботі, так і в операційних цілях).

Саморегуляція медіа у сфері застосування ШІ

Медіа та журналісти відіграють важливу роль у розробці та оновленні стандартів відповідального використання ШІ у сфері медіа й повинні мати чітке та узгоджене між собою бачення. Рекомендується напрацьовувати галузеві кодекси та внутрішні рекомендації щодо безпечного, етичного використання ШІ у сфері медіа, яке ґрунтується на принципах поваги до прав людини, демократії та верховенства права. Прикладом є вищезгадані Рекомендації Комісії з журналістської етики щодо використання штучного інтелекту для створення журналістських матеріалів.

Обмін найкращими практиками, міждисциплінарність, співпраця між професійною та академічною спільнотою, а також громадським сектором є важливими умовами для розвитку відповідального використання систем ШІ у сфері медіа.

Рекомендований алгоритм вибору систем ШІ для роботи та оцінювання згенерованого ШІ контенту

Цей рекомендований алгоритм створено, щоб запропонувати зручний механізм орієнтовної оцінки придатності конкретної системи ШІ для роботи, а також придатності контенту, згенерованого з використанням такої системи ШІ, для поширення.

Кожен чекліст містить дві категорії пунктів: **«червоні прапорці»** (які вказують на неприйнятні або високі ризики використання системи ШІ та поширення згенерованого ШІ контенту) та **«на це потрібно звернути увагу»** (що означає необхідність ухвалення поінформованого рішення).

Використання систем ШІ та згенерованого ними контенту завжди пов'язане з ризиками. Кожен ризик може бути оцінений за шкалою:

«червоні прапорці»:

- **неприйнятний** — наявність хоча б одного такого ризику вказує на неможливість використання такої системи ШІ чи поширення згенерованого нею контенту. Це ультрачервона зона для використання системи ШІ та згенерованого нею контенту;
- **високий** — наявність кількох таких ризиків вказує або на неможливість використання такої системи ШІ чи поширення згенерованого нею контенту, або на високу ймовірність фінансових, організаційних та репутаційних втрат, якщо така система ШІ буде використана чи згенерований нею контент поширений. Це червона зона для використання системи ШІ та згенерованого нею контенту;

«на це потрібно звернути увагу»:

- **середній** — потенційно є середня ймовірність фінансових, організаційних та репутаційних втрат, якщо така система ШІ буде використана чи згенерований нею контент поширений, проте такі втрати навряд чи вплинуть на діяльність суб'єкта у сфері медіа в довгостроковій перспективі. Це жовта зона для використання системи ШІ або згенерованого нею контенту;
- **низький** — низька ймовірність негативних фінансових, організаційних та репутаційних втрат, якщо така система ШІ буде використана чи згенерований нею контент поширений. Це зелена зона для використання системи ШІ або згенерованого нею контенту.

Важливе застереження: низькі ризики, пов'язані з використанням конкретної системи ШІ, автоматично не гарантують низькі ризики при поширенні згенерованого нею контенту, тому кожний згенерований з використанням системи ШІ контент обов'язково повинен проходити оцінювання людиною ризиків, пов'язаних з його поширенням.

Вибираємо систему ШІ для роботи

0. Обов'язково читаємо умови використання від розробника системи ШІ

1. Юрисдикція

- країна реєстрації розробника системи ШІ;
- кінцеві бенефіціарні власники розробника системи ШІ;
- місце провадження діяльності розробником системи ШІ;
- місцезнаходження серверів / хмарних сховищ розробника системи ШІ;
- застосовне право згідно з договором з розробником системи ШІ;
- суд, у якому згідно з договором розглядають спори з розробником системи ШІ щодо використання системи ШІ.

«червоні прапорці»:

Неприйнятним ризиком є наявність російської федерації, Ірану, КНДР у відповіді на будь-який з

перерахованих пунктів.

Високим ризиком є наявність у відповідях на будь-який з перерахованих пунктів країн, які мають високий індекс порушення прав людини (зокрема, щодо захисту персональних даних та права інтелектуальної власності), наявність серед кінцевих бенефіціарних власників розробника системи ШІ підсанкційних осіб, співпраця такого розробника ШІ з урядами держав, указаних у переліку з неприйнятними ризиками.

«на це потрібно звернути увагу»:

Середнім ризиком є наявність у відповідях на будь-який з перерахованих пунктів країн, відмінних від України та країн ЄС (крім держав, що згадані в зоні «червоних прапорців»); співпраця такого розробника ШІ з підсанкційними особами.

2. Умови оплати

Питання з категорії «на це потрібно звернути увагу»: ризики від середніх до низьких

- модель оплати (разовий платіж, підписка тощо) або ліцензія open source;
- чи конкурентна пропонується ціна порівняно з системами ШІ з аналогічними функціями;
- чи є додаткові витрати (наприклад, на інфраструктуру, обладнання, навчання команди) або приховані платежі;
- чи можливо оплачувати в гривні: у зв'язку з воєнним станом оплата в іноземній валюті часто є ускладненою;
- чи готовий розробник системи ШІ надавати документи для коректного закриття платежів з погляду українського податкового законодавства, а саме — підписати договір, акт тощо.

3. Права інтелектуальної власності

- який обсяг прав на згенерований ШІ контент отримує користувач.

Питання з категорії «червоних прапорців»: ризики від неприйнятних до високих у разі відповіді «ні» навіть на одне з таких питань:

- чи можна використовувати згенерований ШІ контент у комерційних цілях або в статутній діяльності суб'єкта у сфері медіа;
- чи можна передавати права на згенерований ШІ контент іншим особам (зокрема, клієнтам), якщо планується таке передання;
- чи гарантує розробник системи ШІ, що згенерований ШІ контент не порушує будь-які права інтелектуальної власності третіх осіб.

Питання з категорії «на це потрібно звернути увагу»: ризики від середніх до низьких

- чи залишаються будь-які права на згенерований контент у розробника системи ШІ. Якщо так, то які саме;
- чи декларує розробник системи ШІ, що він дотримався всіх прав інтелектуальної власності на бази даних (датасети), на яких тренував свою модель ШІ;
- чи гарантує розробник системи ШІ користувачу такої системи визнання та збереження за користувачем прав інтелектуальної власності на бази даних (датасети) користувача та промпти користувача.

4. Приватність: конфіденційність та захист персональних даних

Питання з категорії «червоних прапорців»: ризики від неприйнятних до високих, якщо персональні дані використовують у промтах, а розробник системи ШІ не забезпечує належний захист таких даних

- чи є можливість видалити персональні дані, якщо їх опубліковано в промпті;
- де зберігаються промпти, якщо в них опубліковано персональні дані;
- хто, окрім розробника системи ШІ, має доступ до промптів, якщо в них опубліковано

- персональні дані;
- чи гарантує розробник системи ШІ, що згенерований контент не містить зображень, імен, прізвищ, голосів, інших ідентифікувальних ознак дійсних фізичних осіб;
- чи планується використовувати в промптах конфіденційну інформацію і якщо так, то чи гарантує розробник конфіденційність інформації, наявної в промптах користувача;
- чи планується використовувати в промптах персональні дані і якщо так, то чи гарантує розробник системи ШІ дотримання законодавства щодо захисту персональних даних.

Питання з категорії «на це потрібно звернути увагу»: ризики від середніх до низьких

- чи має користувач можливість видалити промпт;
- які ще особи, окрім розробника системи ШІ, можуть мати доступ до промптів користувача;
- яким законодавством керується розробник системи ШІ в питаннях захисту персональних даних.

5. Вимоги до навичок команди, інфраструктури та обладнання

Питання з категорії «на це потрібно звернути увагу»: ризики від високих до низьких

- які навички вимагає від людини використання системи ШІ та скільки часу необхідно для навчання такого використання;
- чи пропонує розробник системи ШІ навчати використовувати його систему ШІ та на яких умовах (тривалість, оплатність, регулярність тощо);
- чи використання системи ШІ висуває додаткові вимоги до інфраструктури та обладнання (наприклад, доступ до хмарної технології, несумісність з певними платформами);
- якщо так, то які короткострокові та довгострокові додаткові витрати;
- чи пропонується технічна підтримка і на яких умовах (графік, обсяг, вартість);
- чи має право медіа як користувач перейти до іншого розробника системи ШІ і перенести датасет в іншу систему ШІ;
- чи розробник системи ШІ має право одностороннє право змінювати, модифікувати або припиняти надання доступу користувача до системи ШІ в будь-який час.

6. Контроль з боку людини

Питання з категорії «червоних прапорців»: ризики від неприйнятних до високих

- які навички потрібні від команди суб'єкта у сфері медіа для контролю за функціонуванням системи ШІ;
- хто саме в команді суб'єкта у сфері медіа відповідає за перевірку контенту, згенерованого з використанням системи ШІ, на відповідність юридичним та етичним нормам;
- чи закріплена суб'єктом у сфері медіа така відповідальна особа та обов'язковість цієї перевірки документально і якщо так, то в якому документі.

7. Відповідальний розвиток

Питання з категорії «на це потрібно звернути увагу»: ризики від середніх до низьких

- чи система ШІ розроблена спеціально як журналістський ШІ;
- чи система ШІ розроблена для української мови чи потреб української аудиторії;
- чи розробник системи ШІ гарантує, що його система ШІ розроблена відповідно до суспільних цінностей та прав людини.

8. Відповідальність

Питання з категорії «червоних прапорців»: ризики від неприйнятних до високих

- хто за що відповідає: за що конкретно (гарантії належного та законного функціонування системи ШІ, дотримання прав третіх осіб і т. ін.) і в якій сумі відповідає розробник системи ШІ і за що — користувач;
- які гарантії пропонує розробник системи ШІ у разі виникнення юридичної відповідальності

користувача перед третіми особами через фактори, які перебувають поза контролем суб'єкта у сфері медіа;

- яку юридичну допомогу гарантує розробник системи ШІ у разі виявлення порушення прав інтелектуальної власності, прав людини, захисту персональних даних тощо внаслідок добросовісного з боку суб'єкта у сфері медіа використання системи ШІ;
- яких заходів уживає розробник ШІ для постійної безпеки його системи ШІ та безперервної відповідності вимогам законодавства та етичним вимогам.

Оцінюємо згенерований з використанням системи ШІ контент

1. Контроль з боку людини

Питання з категорії «червоних прапорців»: ризики від неприйнятних до високих у разі відповіді «ні»:

- чи перевірила контент людина, відповідальна в команді суб'єкта у сфері медіа за перевірку контенту, згенерованого з використанням системи ШІ, на відповідність юридичним та етичним нормам.

2. Використання авторизованої системи ШІ

Питання з категорії «червоні прапорці»: ризики від неприйнятних до високих, якщо контент згенеровано системою ШІ, яку не авторизовано суб'єктом у сфері медіа згідно з критеріями вище.

- чи контент згенеровано системою ШІ, авторизованою суб'єктом у сфері медіа згідно з критеріями вище.

3. Права інтелектуальної власності

Питання з категорії «червоних прапорців»: ризики від неприйнятних до високих у разі відповіді «ні» навіть на одне з цих питань:

- чи поточна модель оплати або ліцензія open source дає змогу використовувати згенерований системою ШІ контент для, відповідно, поширення в медіа, з комерційною метою або в інших запланованих суб'єктом у сфері медіа цілях;
- чи можна передавати права на згенерований ШІ контент іншим особам (зокрема, клієнтам), якщо таке передання планується;
- чи згенерований ШІ контент візуально / аудіально містить ознаки схожості (або імітації) з будь-якими об'єктами, які охороняються правом інтелектуальної власності (торговельної марки, твору тощо);
- чи гарантує розробник системи ШІ, що згенерований ШІ контент не порушує будь-які права інтелектуальної власності третіх осіб;
- чи використовувалася користувачем у промптах будь-яка інтелектуальна власність третіх осіб (торговельні марки, твори тощо) без правових підстав. Основний спосіб уникнути ризиків за цим питанням — не використовувати інтелектуальну власність третіх осіб без належних правових підстав.

Питання з категорії «на це потрібно звернути увагу»: ризики від середніх до низьких

- чи залишаються будь-які права на згенерований ШІ контент у розробника системи ШІ. Якщо так, то які саме;
- чи гарантує розробник системи ШІ користувачу такої системи визнання та збереження за користувачем прав інтелектуальної власності на бази даних (датасети) користувача та промпти користувача;
- ухвалити рішення, чи потрібно захистити права інтелектуальної власності на власний промпт;
- ухвалити рішення, чи потрібно захистити права інтелектуальної власності на згенерований ШІ контент.

4. Приватність: конфіденційність та захист персональних даних

Питання з категорії «червоних прапорців»: ризики від неприйнятних до високих у разі відповіді «ні» навіть на одне з цих питань:

- чи наявні в згенерованому системою ШІ контенті конфіденційна інформація, комерційна таємниця;
- чи наявні в згенерованому системою ШІ контенті персональні дані, зокрема, чи згенерований контент не містить зображень, імен, прізвищ, голосів, інших ідентифікувальних ознак дійсних фізичних осіб.

Основний спосіб мінімізувати ризики — не використовувати в промптах конфіденційну інформацію, комерційну таємницю, персональні дані, а також перевіряти згенерований ШІ контент на їх відсутність.

5. Достовірність та етичність використання контенту, згенерованого з використанням систем ШІ

Питання з категорії «на це потрібно звернути увагу»: ризики від високих до середніх

- чи контекстуально доречним є використання згенерованого ШІ контенту і чи не введе використання такого контенту аудиторію в оману навіть за наявності маркування (наприклад, якщо контент стосується дуже чутливої теми й може спричинити емоційну реакцію);
- чи згенерований ШІ контент є достовірним та чи всі факти в ньому перевірено;
- чи згенерований ШІ контент перевірено на оригінальність та автентичність, щоб протидіяти поширенню дипфейків чи іншого, згенерованого за допомогою ШІ в недобросовісних цілях, контенту. Як уже згадувалося, прикладами таких інструментів станом на січень 2024 року є *Sentinel*, *FakeCatcher*, *WeVerify*, *Microsoft's Video Authenticator Tool*, інструменти для виявлення дипфейків за допомогою використання техніки *Phoneme-Viseme Mismatches* тощо;
- чи згенерований ШІ контент не порушує немайнові права третіх осіб, зокрема, честь, гідність, ділову репутацію;
- чи згенерований ШІ контент є інклюзивним;
- чи згенерований ШІ контент не є дискримінаційним, дезінформаційним;
- чи згенерований ШІ контент не розпалює ненависть, ворожнечу, чи не пропагує насильство;
- чи згенерований ШІ контент не порушує будь-які інші норми чинного законодавства та журналістської етики.

6. Прозорість використання ШІ

Питання з категорії «на це потрібно звернути увагу»: ризики від високих до середніх

- чи позначено контент, створений з використанням ШІ, як такий, що створений з використанням ШІ;
- чи поінформовано користувача про використання системи ШІ для персоналізації контенту;
- чи надається користувачу медіа право відмовитися від персоналізації контенту: як відбувається інформування користувача про таке право та яка процедура такої відмови.