

РИЗИКИ ВИКОРИСТАННЯ ТЕХНОЛОГІЙ ШТУЧНОГО ІНТЕЛЕКТУ В МЕДІА

Інтерес до можливостей штучного інтелекту в журналістиці значно зріс завдяки нейромережам, які дозволяють нам малювати словами (DALL-E 2, Midjourney та Stable Diffusion), та відкриттю безкоштовного доступу в Україні до ChatGPT. Чат-бот розроблений на основі мовної моделі штучного інтелекту (ШІ) американською лабораторією досліджень OpenAI. Автоматизовані інтелектуальні системи активно застосовуються у світових медіа, вони допомагають насамперед оптимізувати роботу журналістів, спростити збір та опрацювання великих обсягів інформації. Та водночас варто пам'ятати про недосконалість доступних на сьогодні продуктів ШІ та усвідомлювати всі ризики та наслідки їхнього використання.

О. Васьківська, досліджуючи технології штучного інтелекту в журналістиці сучасності, визначає такі основні ризики штучного інтелекту для ЗМІ:

- зловживання інструментами ШІ (наприклад, дипфейками);
- скорочення робочих місць (ШІ та роботизовані системи потребують набагато менше витрат, ніж медіапрацівники);
- етичні конфлікти, найчастіше пов'язані з використанням неперевіраних і неточних даних [1].

Етичні дилеми стосовно використання ШІ дійсно перебувають в центрі уваги. Зокрема, гостру дискусію викликало повідомлення про трагедію в Дніпрі, проілюстроване фото, згенерованим за допомогою ШІ. Найбільша загроза, що вірусне поширення фото в соцмережах як справжнього може стати ідеальною зброєю для російської пропаганди. Адже це – не реальне зображення того, що відбувалося насправді. А значить, вже й справжні фотографії теж нібито можуть бути якимись «не такими» [2].

Юристи Центру демократії та верховенства права серед основних ризиків називають проблему з авторськими правами. Адже програми, що працюють на технологіях ШІ, вчать на творах людей. А результат генерації тексту, фото чи відео може містити об'єкти, захищені авторським правом іншої особи [7]. Консультант з цифрових технологій Віталій Мороз зауважує, що в суспільних дискусіях ще немає чіткої відповіді на важливе питання – як саме описувати контент, згенерований штучним інтелектом? Чи це оригінальний твір на основі нейромережі? Чи це підробка з амбіцією на чиєсь авторство? На думку ж юристки О. Петрів, небезпека полягає у потенційному ризику збільшення кількості творів, що порушують авторське

право інших осіб. Безумовно, згенерувати новий твір за допомогою технологій штучного інтелекту набагато швидше і легше, але ризики порушення прав інтелектуальної власності є дуже високими [3].

Та найбільшим ризиком, на наш погляд, є збільшення кількості дезінформації в інформаційному просторі. Як зазначають самі розробники ChatGPT, іноді чат-бот може виконувати неприйнятні вказівки і демонструвати ознаки упередженої поведінки, а також надавати правдоподібні, але неправильні або відверто безглузді відповіді. Прикладами таких безглузвих відповідей часто діляться користувачі соцмереж. Зокрема, допис в інстаграмі «Нового каналу» до 105-ї річниці з дня народження українського письменника та громадського діяча Олеса Гончара, згенерований ШІ, містив «маловідомі факти» його життя. Проте читачі виявили, що ці факти не маловідомі, а хибні [4]. А цього можна було б уникнути, провівши елементарний фактчекінг. Тобто будь-яка інформація, створена ChatGPT, потребує перевірки. Особливо, якщо на основі неї ви ухвалюєте важливе рішення. До того ж варто враховувати, що чат-бот однієї версії кардинально відрізняється від іншої і може містити застарілу інформацію. У деяких запитаннях це настільки впливає на відповідь, що вона виявляється хибною.

Інтернет-видання «Детектор медіа» провело експеримент з метою визначити ризики поширення дезінформації за допомогою ChatGPT. Результати експерименту засвідчили, що попри вміння вибудовувати логічні ланцюжки і створювати тексти заданої форми, чат-бот не зіставляє факти різними мовами, відповідно, не відрізняє правдиву інформацію від неправдивої. Тож ризики використання ChatGPT у нинішньому вигляді для поширення фейків є досить високими, адже він може створювати прийнятні тексти із різними посилами й емоційним забарвленням [5].

Занепокоєння OpenAI, що ChatGPT може стати зброєю в чужих руках, підтвердило дослідження аналітичного центру NewsGuard, проведене в січні 2023 року. Аналітики виявили, що чат-бот здатен генерувати тексти, які розвивають наявні конспірологічні теорії та включають в їх контекст реальні події. До того ж цей інструмент має потенціал для автоматизованого поширення (з допомогою ботоферм) великої кількості повідомлень, тему і тональність яких визначатиме людина, а безпосередній текст – генеруватиме ШІ [6].

Тож, відкриваючи для себе нові можливості використання штучного інтелекту в медіа, зважаймо на ризики й відповідально ставмося до продукту, який ми виробляємо. Особливо важливо враховувати безпекові ризики застосування мовних моделей та протидіяти поширенню штучно згенерованого неправдивого контенту.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Васьківська О. Технології штучного інтелекту в журналістиці сучасності. URL: <https://dspace.nau.edu.ua/handle/NAU/57218>
2. Верховна Рада проілюструвала трагедію в Дніпрі зображенням, яке створив штучний інтелект. Чому це шкідливо? URL: <https://techno.nv.ua/ukr/innovations/dnipro-malyuk-neyromerezha-50298142.html>
3. Крижня М. Штучний інтелект у творчості та на війні. Про можливості, ризики та авторське право. URL: <https://suspilne.media/426576-stucnij-intelekt-u-tvorcosti-ta-na-vijni-pro-mozlivosti-riziki-ta-avtorske-pravo/>
4. «Новий канал» через ChatGPT поширив фейкову біографію Олеся Гончара. URL: <https://ms.detector.media/profstandarti/post/31588/2023-04-04-novyj-kanal-cherez-chatgpt-poshyryv-feykovu-biografiyu-olesya-gonchara/>
5. Повстання чатботів. Як працює ChatGPT і чим він небезпечний. URL: <https://ms.detector.media/it-kompanii/post/31027/2023-01-20-povstannya-chatbotiv-yak-pratsyuie-chatgpt-i-chym-vin-nebezpechnyy/>
6. ChatGPT: епоха нових загроз чи нових можливостей? URL: <https://www.ukrinform.ua/rubric-technology/3672046-chatgpt-epoha-novih-zagroz-ci-novih-mozlivostej.html>
7. ChatGPT та медіа: як чат-бот допоможе журналістам та в чому ризики. URL: <https://cedem.org.ua/consultations/gpt-ta-media/>