

РОЛЬ ФАКТЧЕКЕРІВ У БОРОТЬБІ З ДЕЗІНФОРМАЦІЄЮ, СТВОРЕНОЮ ЗА ДОПОМОГОЮ ДИПФЕЙКІВ

Дипфейки становлять значну загрозу для сучасного інформаційного простору, оскільки здатні поширювати дезінформацію з високим ступенем реалістичності, що ускладнює їх виявлення та спростування. На думку редактора та старшого аналітика проєкту «VoxCheck» Кирила Перевошикова, найнебезпечнішими для українців, і, в принципі, для будь-якої аудиторії є дипфейки, які поєднуються з ІІСО та створені для підризу підтримки України на Заході, для підризу довіри до влади, зниження підтримки армії або для розпалювання ворожнечі [1].

Дипфейки стали викликом для всіх ресурсів, що займаються розвінчуванням дезінформації. Водночас, за словами головного редактора фактчекінгового порталу «БезБрехні» Олексі Шарабури, фактчек і раніше передбачав виявлення підробок, тому виявлення підробок з використанням штучного інтелекту стало новим викликом лише частково. Попри те, що технічний складник став дещо складніше, насправді нічого нового – все той же пошук невідповідностей деталей чи специфічних способів поширення логічних помилок тощо.

Кожний фактчекінговий ресурс має свої рецепти виявлення дипфейків. Зокрема, експертка проєкту «StopFake» Олена Чуранова зазначає, що з цією метою використовуються «AI or Not» або DeepFake-o-meter. Фактчекерка наголошує, що насправді всі інструменти, які допомагають вирізнити контент, згенерований штучним інтелектом, є нескладними. Вони працюють за принципом: ти завантажуєш фото, відео або текст – інструмент аналізує, наскільки висока ймовірність того, що цей контент був створений ШІ. Водночас експертка наголошує на тому, що інструменти – це не найголовніше у визначенні фейків [2].

Щодо конкретних інструментів, за спостереженням Олексі Шарабури, вони постійно змінюються: ледь не щотижня виходять оновлення, якийсь інструмент виявлення ШІ-зображення починає давати кращий результат, новий парсер чи краулер з використанням нової мовної моделі стає краще розуміти контекст. Проте всі ці інструменти дають лише ймовірну оцінку того, чи був застосований штучний інтелект при створенні зображення, відео, тексту. Головною роботою залишається оцінка отриманих даних самим фактчекером.

Через неточний результат для перевірки відео та зображень фактчекери використовують якомога більшу кількість різних ШІ-детекторів. За словами Олекси Шарабури, результати бувають курйозними, коли кота з п'ятьма лапами програма не помічає або оцінки різних програм надто розходяться. У результаті можна сформулювати загальне уявлення, та не можна покладатися у фактчеку лише на висновок такої програми. Саме через випадки некоректних результатів окремих програм фактчекер не називає конкретні інструменти, а радить використовувати найпопулярніші інструменти для перевірки «AI or not» з найсвіжішого огляду фактчекерів/осінтерів.

Щодо аналізу підробних відео за допомогою ШІ, з досвіду роботи команди «БезБрехні», ШІ-детектори можуть краще помітити сповзання маски, мікрофрізи, невідповідність деталей чи обставин зйомки. У такому разі теж багато важить контекст, на який треба звертати увагу ШІ, якщо, наприклад, ви просите його проаналізувати відео. Можливо ця людина не могла цього в принципі сказати, бо дотримується інших ідеологічних чи релігійних поглядів або ж не перебувала в цей момент у країні. Тут варто завжди пам'ятати правило спілкування з ШІ: що якісніше запитання, то кращою буде відповідь.

Оскільки інструменти ШІ розвиваються шаленими темпами, фактчекери радять постійно пробувати інструменти для клонування голосу, обличчя, різноманітні маски. Адже розуміння, як саме створюється дипфейк, які недоліки та недосконалості виникають під час створення тим чи іншим інструментом, допомагає швидко розпізнати «виробника» за характерними ознаками.

Водночас як у науковому, так і професійному середовищі наголошують, що певні моменти підробки можна помітити й самостійно. Зокрема, при аналізі потенційно підробленого відео слід звернути увагу на такі ознаки: 1) нечіткість або розмитість обличчя на загальному фоні відео; 2) зміни в тоні шкіри поблизу країв обличчя; 3) подвоєння контурів підборіддя, брів або самого обличчя; 4) розмиття обличчя в місцях перетину з руками або іншими предметами; 5) ознаки обрізаних фрагментів обличчя навколо рота, очей або ший; 6) неприродні рухи м'язів обличчя, відсутність кліпання; 7) несподівані зміни фону або освітлення протягом відео.

Загалом позиція, що перевірка дипфейків має бути комбінованою – не тільки використання інструментів, а й власні спостереження за різними сумнівними деталями на відео – є спільною для фактчекерів.

Отже, фактчекінгові ресурси відіграють значну роль у виявленні та спростуванні дипфейків, використовуючи технічні інструменти та методи критичного аналізу для перевірки автентичності контенту та виявлення ознак маніпуляцій. Ефективна боротьба з дипфейками вимагає від фактчекерів не лише навичок технічного аналізу, а й глибокого розуміння політичного та

соціального контексту, щоб ідентифікувати цілі та мотиви створення та поширення дезінформації.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Багацька Л. Аналітик-фактчекер про те, як розпізнати дипфейки: «Створити фейк легше, дешевше та швидше, ніж викрити його» / Л. Багацька // Антикор [Електронний ресурс]. – Режим доступу: [https://antikor.com.ua/articles/690334-analitik-faktcheker_o_tom_kak_raspoz-](https://antikor.com.ua/articles/690334-analitik-faktcheker_o_tom_kak_raspoz-natj_dipfejki_sozdatj_fejk_legche_desheвле_i_bystree_chem_razoblachitj_ego)

[natj_dipfejki_sozdatj_fejk_legche_desheвле_i_bystree_chem_razoblachitj_ego](https://antikor.com.ua/articles/690334-analitik-faktcheker_o_tom_kak_raspoz-natj_dipfejki_sozdatj_fejk_legche_desheвле_i_bystree_chem_razoblachitj_ego)

2. Міський В. «Усе, що ми бачимо й чуємо, може бути підробкою», – Олена Чуранова зі StopFake у подкасті «Медіуми» / В. Міський, Н. Соколенко // Детектор медіа [Електронний ресурс]. – Режим доступу: <https://detector.media/mediumy/article/238216/2025-02-13-use-shcho-my-bachymo-y-chuiemo-mozhe-buty-pidrobkoyu-olena-churanova-zi-stopfake-u-podkasti-mediumy/>