

Список використаних джерел

1. Жаровська І., Ортинська Н. Інформаційна війна як сучасне глобалізаційне явище. *Вісник Національного університету «Львівська політехніка»*. Серія: Юридичні науки, 2020, № 7(2), с. 56–61.
2. Карнаух С. Концептуалізація поняття «інформаційна війна»: від військових доктрин до лінгвістичних досліджень. *Folia Philologica*, 2024, № 7, с. 11–20.
3. Требін М. П. Феномен інформаційної війни у світі, що глобалізується. *Вісник Національного університету «Юридична академія України імені Ярослава Мудрого»*. Серія: Філософія, філософія права, політологія, соціологія, 2013, № 2, с. 188–198.

Артур Волобоєв

Донецький державний університет внутрішніх справ

ЩОДО ПРОТИДІ АДАПТИВНІЙ ДЕЗІНФОРМАЦІЇ: ОКРЕМИЙ АСПЕКТ

Сьогодні суспільство зіткнулося з нетиповою ситуацією, коли технології штучного інтелекту (далі – ШІ) одночасно є і потужним інструментом створення дезінформації, і ефективним засобом її виявлення. Така подвійна роль ШІ створює абсолютно нові виклики для правової системи, які неможливо вирішити за допомогою традиційних підходів до захисту інформаційного простору. Український дослідник Г. Г. Почепцов переконливо стверджує про те, що дезінформація є складовою будь-якої інформаційної операції. Її мета – не просто спотворити факти, а підмінити систему координат, у межах якої ці факти оцінюються. Так створюються умови, коли аудиторія сама робить потрібні висновки [1, с. 71-72]. В умовах повномасштабної агресії проти України ця проблема виходить за межі суто наукового інтересу, – вона набула фундаментального характеру, оскільки інформаційна складова стала ключовим об'єктом маніпулювання у гібридній війні.

Вважаємо, що адаптивна дезінформація – це еволюційно новий тип неправдивого контенту, здатний до самонавчання та динамічних змін залежно від того, як на нього реагує цільова аудиторія. Традиційний фейк, який М. О. Кіца визначає як різновид дезінформації, тобто навмисне подання неправдивої інформації [2], є статичним явищем. Натомість ШІ породжує інший феномен – він не просто копіює реальність, а створює паралельні інформаційні всесвіти, які успішно маскуються під справжні. Технології Deepfake, що базуються на глибокому навчанні нейромереж, генерують відео, аудіо та текст такої якості, що звичайна людина не може відрізнити підробку від оригіналу. Це підтверджується даними українських дослідників, які вказують на те, що лише від 3 до 11 % наших співгромадян можуть відрізнити правду від брехні в інформаційному просторі. [3, с. 2]. Це означає, що суспільство залишається надзвичайно вразливим перед технологічно досконалыми маніпуляціями.

З огляду на це виникає ситуація, при якій ШІ-системи, що розробляються для виявлення підробленого контенту, автоматично навчають генеративні алгоритми обходити методи детекції. Тобто виникає так звана «технологічна гонка озброєнь» між творцями фейків та їхніми викривачами. Кожне покращення детекторів провокує появу складніших способів виробництва дезінформації, а нові техніки виявлення, у свою чергу, стимулюють розробку ще витонченіших підробок. Український досвід демонструє особливий характер цієї проблеми. Ворожа пропаганда широко застосовує технології ШІ для фальсифікації відео з українськими військовими, а потім поширює їх через анонімні телеграм-канали, які маскуються під українські інформаційні джерела.

Правова доктрина стикається тут із кількома питаннями, які потребують розв'язання. Насамперед, перша проблема – це визначення відповідальності за контент, створений автономними системами ШІ. Типове уявлення про правосуб'єктність, яке визнає суб'єктами права лише фізичних та юридичних осіб, не може охопити ситуації, коли шкода заподіюється алгоритмічно, без прямого умислу конкретної людини. Друга проблема стосується доказування умислу у випадках поширення ШІ-генерованого контенту. Чинне законодавство, зокрема стаття 259 «Завідомо неправдиве повідомлення про загрозу безпеці громадян, знищення чи пошкодження об'єктів власності» Кримінального кодексу України [4], вимагає встановлення прямого умислу особи, що практично ускладнюється в ситуаціях автоматизованої генерації та розповсюдження дезінформації.

Європейський досвід правового регулювання свідчить про обережність у цих питаннях. Резолюція Європарламенту 2015/2103(INL) від 16 лютого 2017 року чітко констатує, що на сучасному етапі відповідальність має покладатися на людину, оскільки роботи не можуть нести відповідальність *per se* за дії, що завдали шкоди третім особам [5, с. 50]. Концепція розвитку ШІ в Україні, схвалена Розпорядженням Кабінету Міністрів України від 2 грудня 2020 року № 1556-р, окреслює базові принципи розвитку галузі [5, с. 47], однак конкретних інструментів протидії використанню ШІ для виробництва дезінформації вона не містить.

На нашу думку, ефективна протидія вимагає комплексної системи, яка поєднує технологічні, правові та освітні компоненти.

Так, технологічний компонент має передбачає створення національної платформи моніторингу та верифікації контенту з використанням власних ШІ-детекторів, адаптованих до особливостей українського інформаційного середовища. Правовий компонент має включати модернізацію законодавства про інформацію з урахуванням феномену синтетичного контенту, запровадження обов'язкового маркування ШІ-генерованих матеріалів та встановлення критеріїв їх ідентифікації. Освітній компонент стосується формування медіаграмотності як своєрідного «інтелектуального імунітету» проти маніпуляцій.

Окремо слід звернути увагу на проблему упередженості алгоритмів та дискримінації при використанні ШІ. Дискримінація може потрапити у систему

ШІ двома шляхами: через визначення організацією цільових змінних і класів або через навчання системи на дискримінаційних, часто застарілих даних [5, с. 50]. Законодавство, яке регулюватиме алгоритмічне прийняття рішень, має бути галузевим специфічним та розроблятися на основі принципів регулювання відповідної галузі права. Тому при уніфікації та консолідації українського законодавства важливо передбачити норми про відповідальність за упередженість алгоритмів, які використовуються для автоматизованого прийняття рішень.

Отже, можна стверджувати, що протидія адаптивній дезінформації, яку генерує і детектує ШІ, становить складну міждисциплінарну проблему на перетині технології, права та безпеки. Подвійна природа ШІ як інструменту маніпуляції і водночас засобу захисту вимагає формування нової концепції (стратегії) інформаційної безпеки, яку доцільно передбачити не на засадах абсолютного контролю, а на забезпеченні «технологічної стійкості» суспільства.

Список використаних джерел

1. Почепцов Г. Г. Сучасні інформаційні війни / Г. Г. Поцепцов. Київ : Вид. дім «Києво-Могилянська академія», 2015. 497 с.
2. Кіца М. О. Особливості та методи виявлення фейкової інформації в українських ЗМІ. *Вісник Національного університету «Львівська політехніка»*. Вип. № 1 (883). 2017. С. 30-35. DOI: <https://doi.org/10.23939/sjs2017.01.030>.
3. Єфіменко І. В. Протидія дезінформації в соціальних мережах. *Академія візії*. Вип. 44. 2025. С. 1-8. DOI: <https://doi.org/10.5281/zenodo.15826202>.
4. Кримінальний кодекс України. Кодекс, Закон від 05.04.2001 № 2341-III. *Вебпортал Верховної Ради України*. URL: <https://zakon.rada.gov.ua/laws/show/2341-14#Text>.
5. Каткова Т. Г. Штучний інтелект в Україні: правові аспекти. *Цивілістика. Право і суспільство*. Вип. № 6. 2020. С. 46-55. DOI: <https://doi.org/10.32842/2078-3736/2020.6.1.8>.

Світлана Єфіменко

комунальний заклад «Кіровоградський обласний інститут післядипломної педагогічної освіти імені Василя Сухомлинського»

РОЛЬ ФАКТЧЕКІНГОВИХ ІНІЦІАТИВ У ФОРМУВАННІ СТІЙКОСТІ УКРАЇНСЬКОГО ІНФОРМАЦІЙНОГО ПРОСТОРУ

У сучасних умовах гібридної війни і відповідно постійних інформаційно-психологічних операцій (ІПО) з боку ворога, проблема стійкості національного інформаційного простору набуває критичного значення.

Ключовим елементом у протидії дезінформації є діяльність фактчекінгових ініціатив (від англ. fact-checking – перевірка фактів). Вони виконують функцію «інформаційного імунітету», забезпечуючи верифікацію даних, спростування фейків та підвищення медіаграмотності населення.