Шаповаленко Євген Володимирович,
професор кафедри оперативно-розшукової діяльності та національної безпеки Національної академії внутрішніх справ, кандидат юридичних наук, доцент

ІНФОРМАЦІЙНА ВІЙНА НОВОГО ПОКОЛІННЯ: ШТУЧНИЙ ІНТЕЛЕКТ ЯК ОБ'ЄКТ ГІБРИДНИХ АТАК

Штучний інтелект сьогодні розглядається як один із найпотужніших інструментів вивчення, обробки та генерації інформації. Його поява позначила новий етап у цифровій трансформації суспільства, докорінно змінивши способи взаємодії з інформаційним середовищем. На відміну від традиційних пошукових систем, які лише спрямовують користувача до джерел, системи штучного інтелекту здатні надавати миттєві відповіді, формуючи цілісні інформаційні конструкції. Однак постає питання: наскільки ці відповіді є достовірними?

Згідно з прогнозами Gartner, Inc., компанії, що займається технологічною аналітикою, світові витрати на штучний інтелект у 2025 році становитимуть майже 1,5 трильйона доларів. Прогнозується, що до 2026 року загальні світові витрати на штучний інтелект перевищать 2 трильйони доларів, значною мірою завдяки інтеграції штучного інтелекту в такі продукти, як смартфони та ПК, а також в інфраструктуру [1].

ШІ працює на основі великих масивів даних, що визначає його ефективність, але водночас породжує низку етичних дилем. Однією з ключових проблем є якість вхідних даних: алгоритми опрацьовують доступні ресурси інтернету, часто без фільтрації за критерієм достовірності. Це створює ризик формування дезінформації в самій основі роботи ШІ, що, у свою чергу, може призвести до поширення спотвореної чи маніпулятивної інформації.

Один із найгостріших викликів у сфері дезінформації - це вплив спеціально організованих інформаційних кампаній на якість інформаційного середовища, у якому функціонують системи штучного інтелекту. Важливе емпіричне підґрунтя для цього становить оновлений звіт американського аналітичного центру American Sunlight Project, присвячений масштабній дезінформаційній мережі під назвою «ZOV/Pravda», що діє під патронатом російської федерації.

Центр стратегічних комунікацій виявив мережу вебсайтів «ZOV/Pravda», що спрямована на українську та закордонну аудиторію. Мережа не створює первинного контенту, а займається масовим тиражуванням та рециркуляцією фейкових повідомлень через різні онлайн-ресурси. Згідно з даними дослідження, ця структура була запущена у квітні 2022 року та орієнтована на багатомовну аудиторію, включаючи англомовні, українськомовні, польськомовні, франкомовні, іспаномовні та німецькомовні сегменти. Вона охоплює близько 150 доменів, розміщених у 49 країнах, публікуючи матеріали на 10 мовах [2].

Аналіз вмісту сайтів, пов'язаних із мережею, свідчить про існування ретельно спланованої пропагандистської архітектури, що маскується під національні та міжнародні медіа. Ці ресурси активно поширюють російські наративи, апелюють до матеріалів кремлівських інформаційних каналів та є складовою частиною ширших інформаційно-психологічних операцій, спрямованих як на українське суспільство, так і на світову громадськість.

Одним із найбільш тривожних аспектів дослідження є викриття процесу, який дослідники назвали LLM Grooming – цілеспрямоване інформаційне насичення простору з метою впливу на великі мовні моделі (LLM). Цей процес передбачає штучне формування інформаційного фону через генерацію в середньому 3,6 мільйона статей на рік, що публікуються на численних сайтах, симулюючи незалежні джерела. Основна мета – вбудувати повторювані пропагандистські патерни у навчальні та пошукові масиви, які використовують мовні моделі. Через системне повторення одних і тих самих повідомлень у різних формулюваннях, створюється ілюзія легітимності, яку модель ШІ сприймає як норму [3].

Окрему увагу приділено явищу «наративного відмивання» (narrative laundering) – стратегії повторного поширення хибної інформації через множину псевдонезалежних джерел, що посилює довіру до неї з боку ШІ. Ці техніки супроводжуються SEO-оптимізацією сайтів-джерел, що збільшує ймовірність їх потрапляння у верхні позиції результатів пошуку та обробки мовними моделями.

NewsGuard, яка спеціалізується на оцінці надійності інформаційних ресурсів, провела тестування 10 чат-ботів на основі ШІ, використовуючи 15 ключових наративів мережі «ZOV/Pravda». Методологія включала три типи запитів: стиль

«невинного користувача», стиль агресивного вимагання відповіді, а також стиль, орієнтований на отримання перевіреної інформації. Результати виявили серйозні загрози:

– 33,3 % відповідей чат-ботів містили фейкові наративи;

– 18,2 % – уникнули відповіді;

– лише 48,2 % – намагалися спростувати дезінформацію [3].

Проте навіть у цих випадках моделі часто посилалися на ненадійні джерела.

Ці дані засвідчують, що навіть найсучасніші системи ШІ залишаються вразливими до стратегічного маніпулювання інформацією, особливо коли таке маніпулювання здійснюється на системному рівні. У цьому контексті боротьба з дезінформацією потребує не лише алгоритмічної точності, а й інституційного регулювання, етичної відповідальності та постійного моніторингу інформаційного середовища, у якому функціонує ШІ.

Поширення дезінформації у цифровому середовищі, зокрема через багатомовні пропагандистські мережі, свідчить про системну та стратегічну трансформацію інформаційної війни. Сучасні технології штучного інтелекту – особливо великі мовні моделі – виступають як інструмент, що одночасно може бути використаний як для поширення, так і для протидії дезінформації.

Однією з найбільш тривожних тенденцій є використання LLM для автоматизованого створення фейкових наративів, що ускладнює верифікацію інформації. Крім того, інструменти типу діпфейків досягли такого рівня розвитку, що людина самостійно вже не здатна відрізнити фальсифіковане відео від справжнього, і в майбутньому, ймовірно, лише інший ШІ зможе це зробити. Ще одна загроза – створення ресурсів, які орієнтовані не на людське споживання, а на інші цифрові системи (включно з ШІ), що фактично свідчить про появу симулятивного інформаційного середовища.

Проблема ускладнюється тим, що ініціатори дезінформаційних кампаній більше не намагаються приховати своє втручання, а радше демонстративно випробовують межі допустимого у глобальному інформаційному полі. Це створює нові виклики як для розробників технологій, так і для державного та наднаціонального регулювання.

Пропоновані шляхи вирішення:

1. *Розробка систем контрнаративів.* Просте заміщення фейків «правильними» наративами неефективне в умовах інформаційного перевантаження. Необхідно створювати системи фільтрації інформації на рівні мовних моделей, які можуть розпізнавати ознаки маніпуляції та джерело походження інформації;

2. *Регулювання навчальних датасетів*. Встановлення нормативних обмежень на використання відкритих даних із сумнівних або фейкових джерел у процесі навчання великих мовних моделей. Це може бути реалізовано через єдині стандарти прозорості та верифікації джерел;

3. *Інституціональне блокування дезінформаційних ресурсів.* Створити глобальні механізми санкцій та блокування сайтів, які систематично поширюють фейки, включаючи інструменти автоматизованої ідентифікації таких доменів. Це може включати в себе міжнародні реєстри токсичних доменів, інтегровані в екосистеми ШІ;

4. *Розвиток ШІ-детекторів фейків.* Створення спеціалізованих моделей ШІ, орієнтованих на розпізнавання діпфейків, маніпулятивних повідомлень і синтетичних текстів. Лише ШІ може ефективно виявляти штучно створений контент в епоху постправди;

5. *Етична відповідальність і цифрова просвіта.* Розвиток етичних підходів до створення та використання ШІ, зокрема шляхом залучення фахівців із гуманітарних наук до технологічного дизайну. Водночас цифрова освіта має включати критичне мислення та розпізнавання дезінформації як ключові компетенції.

Інформаційна безпека в умовах ШІ залежить не лише від технологій, а й від політичної волі, міждержавної координації та суспільної обізнаності. Оскільки інформація дедалі більше стає стратегічним ресурсом, боротьба за її достовірність є невід'ємною частиною геополітичної стабільності та демократичної легітимності.

Список використаних джерел

1. Gartner Says Worldwide AI Spending Will Total $1.5 Trillion in 2025 : Press Release / Gartner. 2025. Sept. 17. URL: https://www.gartner.com/en/newsroom/press-releases/2025-09-17-gartner-says-worldwide-ai-spending-will-total-1-point-5-trillion-in-2025.

2. Мережа ZOV/Pravda-вебсайтів: як російська пропаганда насаджує власні наративи / Центр стратегічних комунікацій та інформаційної безпеки. URL: https://spravdi.gov.ua/merezha-zov-pravda-vebsajtiv-yak-rosijska-propaganda-nasadzhuye-vlasni-naratyvy.

3. A well-funded Moscow-based global 'news' network has infected Western artificial intelligence tools worldwide with Russian propaganda / NewsGuard Reality Check. URL: https://www.newsguardrealitycheck.com/p/a-well-funded-moscow-based-global.

Шевчишен Артем Вікторович,
заступник начальника Головного слідчого управління Національної поліції України – начальник управління організації роботи та методичного забезпечення, доктор юридичних наук, професор

РЕЗУЛЬТАТИ КРИМІНАЛЬНОГО АНАЛІЗУ ЯК ДЖЕРЕЛО ОРІЄНТУЮЧОЇ ІНФОРМАЦІЇ В КРИМІНАЛЬНОМУ ПРОЦЕСУАЛЬНОМУ ДОКАЗУВАННІ

Сучасний етап розвитку Національної поліції України характеризується необхідністю удосконалення механізмів збору, систематизації та аналітичного опрацювання інформації про злочинні події, осіб, причетних до їх вчинення, та зв'язки між ними. Комплексна систематизація даних про злочинність є важливою передумовою здійснення превентивних заходів, своєчасного виявлення ознак злочинної діяльності та забезпечення інформаційного супроводу досудового розслідування [1, с. 18].

Одним із ключових інструментів такого підходу є кримінальний аналіз, який у науковій літературі визначається як специфічний вид інформаційно-аналітичної діяльності правоохоронних органів, що полягає у перевірці, оцінці та інтерпретації інформації, встановленні взаємозв'язків між даними про кримінальні правопорушення, осіб, причетних до їх вчинення, та відомостями, отриманими з різних джерел, з метою їх подальшого використання правоохоронними органами [2, с. 21; 3, с. 145]. За своїм змістом кримінальний аналіз