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

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В ПРОТИДІЇ ДЕЗІНФОРМАЦІЇ

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 даними з 2013 по 2024 рік, залучили у розвиток штучного інтелекту 471 млрд доларів приватних інвестицій. Ця сума перевищує надходження в усіх інших країнах світу, узятих разом (289 млрд доларів). Серед лідерів також Китай (119 млрд дол.), Велика Британія (28 млрд дол.), Канада та Ізраїль (по 15 млрд дол.).

Статистика наведена зі звіту про Індекс ШІ за 2025 рік Стенфордського інституту людиноорієнтованого штучного інтелекту. Щодо даних за останній рік, то за цей період США зберігає беззаперечне лідерство – приватні інвестиції в штучний інтелект зросли до 109,1 млрд дол. - майже в 12 разів більше, ніж у Китаї (9,3 млрд дол.).

Сектор генеративного штучного інтелекту залучив 20% усіх приватних інвестицій в ШІ. Цифри досягли $33,9 млрд у 2024 році, що на 18,7% більше, ніж у 2023 році.

Уряди по всьому світу нарощують свою частку інвестицій у ШІ. Канада анонсує 2,4 мільярда доларів, Китай запустив фонд напівпровідників на 47,5 мільярда доларів, Франція виділила 109 мільярдів євро, Індія планує вкласти 1,25 мільярда доларів тощо [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 «Правда», що діє під патронатом Російської Федерації [2].

Мережа «Правда» 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 [2].

Окрему увагу приділено явищу «наративного відмивання» (narrative laundering) – стратегії повторного поширення хибної інформації через множину псевдо незалежних 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 «Правда». 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.

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 [3].

Поширення дезінформації у цифровому середовищі, зокрема через багатомовні пропагандистські мережі на кшталт «Правда», 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 [2].

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.

Аналізуючи проблематику зазначеного питання пропонуємо наступні шляхи вирішення:

– розробка систем контрнаративів (створювання системи фільтрації інформації на рівні мовних моделей, які можуть розпізнавати ознаки дезінформації та джерело походження інформації);

– регулювання навчальних датасетів (встановлення нормативних обмежень на використання відкритих даних із сумнівних або фейкових джерел у процесі навчання великих мовних моделей);

– інституціональне блокування дезінформаційних ресурсів (створення механізмів санкцій та блокування сайтів, які систематично поширюють фейки, включаючи інструменти автоматизованої ідентифікації таких доменів);

– розвиток ШІ-детекторів фейків (створення спеціалізованих моделей ШІ, орієнтованих на розпізнавання діпфейків, маніпулятивних повідомлень і синтетичних текстів);

– етична відповідальність і цифрова освіта (розвивати цифрову освіту, етичні підходи до створення та використання ШІ).

У підсумку можна зазначити, що інформаційна безпека під час використання ШІ залежить не лише від технологій, а й від політичної волі, міждержавної координації та суспільної обізнаності. Оскільки цифрова інформація дедалі більше стає стратегічним ресурсом, боротьба за її достовірність є невід'ємною частиною геополітичної стабільності та демократичної легітимності.

Список використаних джерел

1. Розвиток штучного інтелекту – США вклали грошей більше, ніж інші країни разом взяті. URL: https://renews.com.ua/tehnologiyi/rozvitok-shtychnogo-intelekty-ssha-vklali-groshei-bilshe-nij-inshi-krayini-razom-vziati/.

2. A well-funded Moscow-based global ‘news’ network has infected Western artificial intelligence tools worldwide with Russian propaganda. URL: https://www.newsguardrealitycheck.com/p/a-well-funded-moscow-based-global.

3. Мережа ZOV/Pravda-вебсайтів: як російська пропаганда насаджує власні наративи. URL: https://spravdi.gov.ua/merezha-zov-pravda-vebsajtiv-yak-rosijska-propaganda-nasadzhuye-vlasni-naratyvy.

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

КРИМІНАЛЬНИЙ АНАЛІЗ ЯК ДОПОМІЖНИЙ ІНСТРУМЕНТ ФОРМУВАННЯ ТА ЗБИРАННЯ ДОКАЗІВ У КРИМІНАЛЬНОМУ ПРОВАДЖЕННІ

У сучасних умовах боротьби зі злочинністю правоохоронні органи зіштовхуються з новими, дедалі складнішими викликами, зумовленими високим рівнем латентності злочинів, зростанням масштабів та організованості злочинних угруповань, а також багаторівневою і часто транснаціональною структурою взаємозв’язків між учасниками протиправної діяльності. Злочини дедалі частіше вчиняються із використанням складних схем конспірації, цифрових технологій, підставних осіб та фіктивних суб’єктів господарювання, що значно ускладнює традиційні методи досудового розслідування.

З огляду на зазначені вище обставини, на сучасному етапі важливо впроваджувати нові підходи у протидії злочинності та розвивати нові навички для оптимального опрацювання інформації оперативними працівниками. Цей процес базується