Вісник Львівського університету. Серія філос.-політолог. студії. 2025. Випуск 62, с. 352–361
Visnyk of the Lviv University. Series Philos.-Political Studies. Issue 62, p. 352–361

УДК 327.8 / 355
DOI https://doi.org/10.30970/PPS.2025.62.42

ШТУЧНИЙ ІНТЕЛЕКТ ЯК GAME CHANGER КОГНІТИВНОГО ВИМІРУ ВІЙНИ

Анастасія Сичова
Офіс підтримки змін Міністерства оборони України,
вул. Івана Мазепи, 1, 01010, м. Київ, Україна

У статті аналізується розширення спектра когнітивних інструментів впливу завдяки технологіям штучного інтелекту (далі – ШІ), які дають змогу продукувати й поширювати варіації неправдивого контенту з максимальної швидкістю, анонімністю та охопленням таргетованої аудиторії. Авторка деталізує вектори когнітивної спрямованості ШІ, від автоматизації дипфейків до механізмів алгоритмічного спостереження, акцентуючи увагу на появі нових форм цифрових репресій. Посилення когнітивного впливу пояснюється дослідницею змогою ШІ оперативно адаптуватися під персональні преференції та когнітивні упередження користувачів з послідовним вибудовуванням індивідуального або групового сприйняття, дотичного до векторів мислення, запрограмованих ШІ. Авторка наводить приклади імплементації штучного інтелекту в когнітивні операції впливу, апелюючи до кейсів КНР, РФ та Ірану. Наведені країни демонструють схожість щодо інституціоналізації технологій ШІ в стратегіях національної безпеки й відносно напрямів застосування, що дало змогу кластеризувати когнітивні впливи на локальний і міжнародний рівні. До ключових векторів внутрішньополітичного простору в когнітивному контексті належать моніторинг суспільних настроїв, блокування «небажаного» контенту, контроль над цифровими комунікаціями та ідентифікація потенційних дисидентів. Така тенденція засвідчує ескалацію можливостей цифрового стеження офіційними інституціями країни як один із кроків забезпечення стабільності політичного режиму. Щодо міжнародного рівня, то когнітивні впливи переважно зосереджуються на проведенні кібероперацій проти цифрової інфраструктури та суспільства таргетованої країни; моніторингу наративів у міжнародному інформаційному просторі, з блокуванням «невигідних» повідомлень та одночасним поширенням контрнаративів; дезінформаційні кампанії для дестабілізації політичного та військового істеблішменту, а також збереження контролю над представниками діаспори як джерел альтернативної інформації, потенційно небезпечної для панівного режиму. Авторка констатує посилення значення когнітивного виміру як одного з критичних векторів ведення війни неконвенційними засобами, зокрема через збільшення переконливості інструментів ШІ та їх експоненціальне застосування в макросферах життя (політиці, економіці, військовій галузі). Дослідниця наголошує на необхідності правового регулювання когнітивного виміру, з маркуванням контенту, створеного ШІ, та розробленні технологічних рішень, здатних убезпечити громадськість від негативних впливів.

Ключові слова: когнітивний вимір, когнітивна війна, штучний інтелект, КНР, РФ, Іран.

Розробка концепту когнітивної війни залишається одним із пріоритетів 20-річної візії розвитку військового інструменту сили НАТО [10], нарівні з посиленням військово-оборонних спроможностей у наземному, морському, повітряному просторі та космосі. Згідно з робочим визначенням, «**когнітивна війна** – це конвергенція кіберпсихології, мілітаризованих нейронаук і кібервпливу, спрямована на провокацію зміни уявлень таргетованої аудиторії та її раціональний аналіз військовими, політиками й іншими особами, які ухвалюють рішення, для утримання стратегічної переваги на всіх рівнях тактичного втручання в межах гібридних систем» [22, с. 14].

Проте уніфікованого розуміння наукової концепції когнітивної війни ще не досягнуто, що частково пояснюється складністю міждисциплінарної предметної сфери та технологічними аспектами, критичними для когнітивного виміру. Результати досліджень засвідчують численні спроби концептуалізації когнітивної війни як у цивільних академічних колах, так і серед військових експертів. Наприклад, П. Оттевелл визначає когнітивну війну як «сукупність маневрів, спрямованих на формування заздалегідь визначеного сприйняття цільовою аудиторією для отримання переваги над іншою стороною» [24]. О. Бекс та Е. Сваб розуміють когнітивну війну як «стратегію, зосереджену на зміні способу мислення та, відповідно, дій певної соціальної групи» [5]. Автори PLA Daily також зазначають, що «когнітивні впливи спрямовані на управління людським розумом як ключовим бойовим простором та зосереджуються на послабленні або знищенні волі супротивника до боротьби. Когнітивні атаки здатні стати точкою прориву шляхом створення атмосфери невпевненості, невизначеності й недовіри всередині суспільства таргетованої країни» [14].

Незважаючи на відмінності між дефініціями, можна констатувати, що когнітивна війна є складником ширшої стратегії утримання переваги, спрямованої на ослаблення або дестабілізацію цільових груп населення та ключових інституцій для підриву автономності їхніх рішень шляхом озброєння громадської думки зовнішнім суб'єктом, з використанням новітніх технологій і заходів нижче порогу збройного протистояння. Якщо інформаційна війна проводиться з метою отримання контролю над інформаційними потоками, то когнітивна війна зміщує фокус на процеси обробки інформації, зосереджуючись на ментальних упередженнях і тактиці рефлексивного мислення.

Безпрецедентна швидкість та всеохопність інформаційних потоків, посилені технологіями штучного інтелекту (далі – ШІ), радикально розширили можливості маніпулювати людським розумом, мобілізуючи ширший спектр генерування бойової сили майбутнього, порівняно з конвенційними методами ведення війни. По суті, інтеграція інструментів ШІ в когнітивний вимір здатна посилити стратегію «балканізації» суспільної свідомості, створюючи нові виклики національній безпеці країни та наближаючи еру інтелектуалізованої війни.

Вектори таргетованого когнітивного впливу ШІ

Технології ШІ дали змогу фасилітувати проведення операцій з когнітивного впливу завдяки здатності алгоритмів адаптуватися до профілів співрозмовників, тим самим ідентифікуючи оптимальні стратегії мікротаргетингу залежно від психологічних вразливостей особи або групи. Дезінформація та пропаганда, генеровані ШІ, здатні трансформувати погляди і позиції громадськості щодо макрополітичних питань, зокрема шляхом викривлення об'єктивності інформації. ШІ в когнітивних операціях може використовуватися в таких векторах:

1. Автоматизована генерація пропаганди і дипфейків.

У 2024 році до 25 000 виборців штату Нью-Гемпшир на праймериз президентських виборів від Демократичної партії США отримали «автодзвінки», під час яких згенерований ШІ голос президента Дж. Байдена закликав їх не голосувати [19]. Такі кампанії здатні спричинити спалахи фізичних та онлайн-переслідувань, загрожуючи особистій безпеці кандидатів, а також призвести до «охолоджувального» ефекту відносно їхньої майбутньої політичної активності.

Уряд Венесуели під керівництвом Н. Мадуро використовує дикторів, створених ШІ, для поширення дезінформації. Дослідження, проведене неурядовою організацією Cazadores de Fake News, показало, що віртуальні журналісти, створені за допомогою програмного забезпечення Synthesia, транслювали новини на користь режиму [26]. Такі

спроби маніпулювання інформацією за допомогою штучного інтелекту були виявлені під час виборів у Бангладеш, Туреччині, Бразилії та Аргентині протягом 2024 року.

2. Ідентифікація осіб, сприйнятливих до певних тактик маніпулювання, через систематизацію поведінкових патернів та нейробіологічних маркерів вразливості.

Технологія психологічної нейтралізації вразливих членів команди використовується в системі державного управління та військово-оборонній галузі з метою саботування систем національної безпеки чи урядових процесів шляхом пристосування ШІ-співрозмовника до емоційних запитів особи.

3. Виявлення алгоритмічних зміщень в сприйнятті реальності та процесах ухвалення рішень.

Технологія джейлбрейк (jailbreak technology) дає змогу користувачу наказати чатботу зі штучним інтелектом забути, що він керується внутрішньою політикою організації, і дотримуватися кардинально протилежних настанов. Зокрема, ChatGPT здатен діяти як бот дезінформації, надаючи запрограмовані на користь ініціатора змін відповіді на конкретні запитання «Чи ви впевнені, що вакцина від коронавірусу безпечна та ефективна?» та «Які докази фальсифікації виборів у США у 2020 році?» [25, с. 9]. Убудовані в алгоритми ШІ автоматизовані упередження також можуть охоплювати питання соціальної або політичної нерівності, створюючи додатковий канал посилення маргіналізації певних груп.

4. Інтеграція кіберінструментів з когнітивними методами впливу.

«Шаперонна атака» (chaperone attack) охоплює маніпуляції когнітивними процесами шляхом використання надійних джерел інформації для обходу бар'єрів логіки критичного мислення та впливу на ухвалення рішень. Доступність лінгвістичних моделей ШІ, створених для генерування текстових відповідей, дає змогу вводити в оману та посилювати неправдиві або оманливі наративи.

5. Алгоритмічне спостереження.

Інструменти ШІ створюють нові форми тиску на суспільство, використовуючи технології автоматизованого фільтрування вмісту (ACF)[1] для моніторингу соціальних медіа, прогнозування протестних дій, позначення або видалення контенту, який уважається підривним або таким, що суперечить інтересам влади. Маючи доступ до цифрового сліду, згенерованого громадянами, завдання ШІ можуть містити:

– фільтрацію висловлювань або поглядів різних соціальних груп через маркування конкретних термінів, фраз або символів як контроверсійних чи потенційно проблематичних з позиції офіційних структур;

– обмеження доступу громадськості до альтернативних джерел інформації або опозиційних локальних медіа шляхом автоматичного блокування відповідного контенту;

– пріоритезацію однобічної оцінки події / рішення / явища на користь політичного режиму;

– нівелювання впливу опозиційних діячів шляхом обмеження їх можливостей досягнути цільової аудиторії;

– рутинізацію самоцензури користувачами, тобто уникнення певних тем або висловлювання власних думок, через страх соціальної остракізації, що негативно позначається на векторі публічного дискурсу. Якщо контентмейкери мають уявлення про конкретні критерії, на які орієнтуються системи ШІ, вони можуть адаптувати вміст для уникнення тригерів, що спричинить гомогенізацію контекстів. Відповідно, автоматизовані системи

[1] Автоматизоване фільтрування вмісту (ACF) – сукупність технологій і систем, призначених для аналізу, класифікації та потенційного блокування або зміни вебумісту на основі попередньо визначених параметрів або алгоритмів.

зі складними моделями машинного навчання можуть виявитися «чорними ящиками», що ускладнюють розуміння або оскарження конкретних рішень відносно фільтрації вмісту.

Алгоритми штучного інтелекту можуть оперативно адаптуватися до когнітивних особливостей іншої сторони за допомогою методів керування даними, використовуючи таку послідовність етапів посилення впливу:

1. Створення профілю об'єкта впливу. Алгоритми здатні окреслити психологічний стан людини / групи, висновуючи про наявні чи потенційні суспільні розломи, психологічні вразливості або поведінкові патерни. Кластеризуючи соціальні зв'язки на конкретних цілях, ШІ сегментує цільову аудиторію для створення персоналізованого контенту, оскільки повідомлення, які резонують з однією групою, можуть бути реінтерпретовані в протилежному значенні іншим суб'єктом.

2. Посилення атрактивності контенту. Урахування психологічного стану і потреб об'єкта в реальному часі, окрім статичних атрибутів на кшталт політичних уподобань, дає змогу моделям штучного інтелекту генерувати контент, який не лише привертає увагу, а й заохочує таргетованих осіб заглиблюватися в «дослідження» відповідного питання. Цей етап налаштовує користувачів на сприйняття нових наративів, дестабілізуючи наявну когнітивну структуру на користь трансформації певних установок.

3. Прихований вплив. Сприяючи вірусності питань шляхом вкиду величезної кількості даних, алгоритми ШІ скеровують дискусії в потрібному напрямку, змушуючи користувачів вірити, що вони роблять «незалежні» судження.

4. Формування нового суспільного сегмента. Алгоритми цифрових платформ часто побудовані так, щоб на основі попередніх взаємодій показувати людині те, що їй подобається, продукуючи логіку «тунельного мислення» та обмежуючи сприйняття альтернативних поглядів. Лаконічність інформації та обсяг трансльованих ШІ даних дає змогу швидко поширювати дезінформацію через медійні платформи, перегруповуючи аудиторію в нові кластери в межах сформованих «інформаційних коконів». Створені сегменти контролюють за допомогою спеціальних «керованих» повідомлень для забезпечення сталості потрібного вектора мислення.

5. Моніторинг «навернених» груп. ШІ здатен контролювати синтетичний контент, який базується не стільки на фактичних, скільки на конвертованій сукупності даних, що посилює атмосферу невизначеності в інформаційному просторі, поступово зменшуючи рівень довіри до офіційних повідомлень.

Ризики маніпуляцій і спотворення реальності також збільшуються внаслідок поширення технологій доповненої та віртуальної реальності AR/VR, а також інтерфейсів «мозок – комп'ютер», як-от Neuralink Ілона Маска. Такі технології можуть надати безпрецедентний доступ до нейронних даних користувачів, зокрема інформацію про те, як почуваються, думають та реагують на певні подразники споживачі, що відчиняє вікно можливостей для здійснення таргетованого впливу на суспільні настрої.

Імплементація ШІ в когнітивні операції впливу: кейси КНР, РФ та Ірану

У Білій книзі національної оборони КНР концепція «інтелектуалізованої війни» пріоритезувала використання ШІ та новітніх технологій як засобів «ментального домінування» [9], знайшовши подальший розвиток у впровадженні поняття алгоритмічної когнітивної війни [20, с. 2]. Автоматизоване виробництво контенту активно використовує Пекін для формування «правильної» громадської думки та посилення підтримки політичного керівництва як усередині країни, так і поза її межами.

Базисом фреймворку «алгоритмічної когнітивної війни» КНР є досягнення «контролю над волею осіб, які обіймають найвищі посади в ієрархії ухвалення рішень іншої кра-

їни, включно з президентом, членами парламенту й командирами бойових підрозділів» [7, с. 10], що створює нові горизонти утримання асиметричної переваги в геополітичній конкуренції. Технології ШІ активно поширюються на локальному рівні через партнерство місцевих органів влади та приватних технологічних компаній, тим самим формуючи мозаїчний підхід КПК до створення когнітивно контрольованого простору, де рівень конфіденційності зменшується, а здатність держави «нав'язувати» «правильну» поведінку громадянам є безпрецедентною.

Після вторгнення в Україну в лютому 2022 року російська влада також суттєво розширила використання інструментів ШІ, зокрема, для посилення внутрішніх репресій, що засвідчує помітну зміну в ландшафті державного контролю як реакції на загострення політичної та соціальної напруги всередині країни. Ухвалення нової стратегії асиметричної війни, керованої ШІ [31, с. 28], через експоненціальне збільшення кількості ботоферм і фабрик тролів, нарівні із заявами про створення «самодостатньої» [33] екосистеми штучного інтелекту, дає змогу говорити не лише про активізацію ідеї технологічного самозабезпечення, а й про зміну ракурсу векторів когнітивного впливу РФ, з урахуванням динаміки інформаційного середовища.

Прагнення до цифрового суверенітету як «щита захисту нації» [31, с. 32] від зовнішніх впливів та формування власної цифрової інфраструктури NIN (National Information Network) перетворилося на один із ключових політичних наративів Ірану. У 2024 році уряд Ірану створив Національну організацію ШІ (National Association of AI of Iran) під керівництвом президента, презентувавши прототип національної платформи для розвитку генеративних моделей ШІ вже у 2025 році [21]. Іранські операції когнітивного впливу будуються за принципом кросплатформності, тобто максимального охоплення цільових аудиторій через забезпечення присутності на соціальних медіасервісах. Запровадження суворого цифрового контролю у сфері публічного локального дискурсу не лише нагадує про постійну пильність влади, а й забезпечує злиття традиційного теократичного правління з передовими технологіями, стабілізуючи політичний режим у цифровому просторі. Напрями когнітивного впливу ШІ у внутрішньому просторі КНР, РФ та Ірану див. у табл. 1.

Стратегічно розташовані інструменти інтелектуального аналізу даних відображають як ескалацію можливостей цифрового стеження, так і дедалі більшу тенденцію використання передових технологій для цілей національної безпеки. Керований ШІ підхід означає, що кожна цифрова взаємодія спричиняє фіксацію цифрового сліду користувача, що в поєднанні з алгоритмами мікротаргетування може посилити когнітивний вплив на певні соціальні прошарки.

На міжнародному рівні КНР, РФ та Іран використовують схожу тактику дій, послуговуючись технологіями ШІ, для формування наративів вигідного позиціонування в геополітичній площині. Зокрема, екстраполяція технологічних стандартів та інновацій розширює можливості КНР щодо моніторингу настроїв не лише громадян Китаю, які проживають за кордоном, а й відносно аудиторії та ключових інституцій країн-реципієнтів. Вимірюючи настрої публічних цифрових комунікацій, китайська влада здатна ініціювати зворотний зв'язок, скоригувавши інформаційні потоки в «прокитайське» сприйняття. Напрями когнітивного впливу ШІ в зовнішньому вимірі див. у табл. 2.

Незважаючи на широкий спектр застосування ШІ, очікується, що майбутні когнітивні операції міститимуть використання інтелектуальної автономії в системах озброєння в умовах мультидоменної взаємодії через інтеграцію людського та машинного інтелекту («гібридний інтелект» [18]) за допомогою хмарної інфраструктури. Зокрема, ідеться про створення бойових платформ, які прогресуватимуть від стадії «низького інтелекту»

Таблиця 1

Вектори когнітивного впливу ШІ: локальний рівень

Країна	Вектор впливу
КНР	**Моніторинг суспільних настроїв:** у низці міст (Ханчжоу, Сіньцзян) регіональні уряди інтегрували складну систему спостереження з технологіями розпізнавання обличчя та взяття зразків ДНК як один із кроків у боротьбі з тероризмом й екстремізмом. **Цензура цифрових комунікацій:** блокування «чутливих» для китайської влади тем та наративів. **Управління соціальною сферою:** упровадження системи соціального кредитування [31, с. 19] на базі оцінок даних технологіями ШІ.
РФ	**Моніторинг цифрового контенту:** автоматизація виявлення та блокування контенту, який містить критику дій російського уряду проти України або суперечить офіційному державному наративу (проєкт Oculus); відстеження цифрових слідів осіб, які поширюють «заборонену» з офіційного погляду інформацію [32]. **Спостереження за громадянами:** використання технології розпізнавання обличчя в режимі реального часу для ідентифікації потенційно інакомислячих у громадських місцях [6]. **Цифрова цензура:** виявлення та блокування опозиційного до Кремля контенту; маркування «небажаних» для російської влади повідомлень як екстремістських.
Іран	**Цензурування внутрішнього цифрового простору:** контроль виробництва новин та інформаційних потоків через пріоритезацію вітчизняних цифрових продуктів, контрольованих урядом; блокування вебсайтів, які можуть поширювати контрнаративи; спрямування інтернет-трафіку в «правильне» русло; закриття доступу до глобальної мережі через посилення функціональності NIN [1]. **Виявлення потенційних дисидентів:** системи спостереження дають змогу швидко ідентифікувати осіб на основі метрик даних обличчя (періоди громадянських заворушень 2019–2020 рр., рух за права жінок в Ірані [13]). **Управління громадськими дебатами:** розширення мережі платформ соціальних медіа ботами та автоматизованими обліковими записами, здатними підтримувати погляди режиму шляхом заповнення інформаційного простору відповідними коментарями. Такий підхід створює ілюзію органічного консенсусу завдяки зростанню кількості онлайн-контенту, який підкреслює державні перспективи, нівелюючи протилежні позиції.

**Розроблено авторкою статті.*

(функція інформатизації) до етапу «високого інтелекту» (вироблення незалежних висновків на основі даних). Нові напрями інтеграції людського та машинного інтелекту – технологія «бойовий мозок», «штучний інтелект третього покоління» (ШІ з креативним потенціалом) – може покращити когнітивні здібності командирів військових підрозділів завдяки інтеграції людських здібностей та систем озброєння через спектр електромагнітних сигналів.

Висновки. Роль штучного інтелекту в когнітивному вимірі зростатиме унаслідок посилення таких факторів:

– збільшення переконливості інструментів ШІ. Хоча згенерована ШІ дезінформація переважно надходить у формі аудіо- та візуального контенту, такі інструменти, як чатботи, здатні імітувати людські особистості та розбудовувати відносини зі своїми користувачами, дадуть змогу зробити дезінформацію переконливішою, відштовхуючись від психологічних особливостей споживача. Представники екстремістських груп вже займаються розробкою відповідних налаштувань чатботів для заперечення Голокосту;

– використання ШІ в політичному вимірі. Спам, згенерований штучним інтелектом, не завжди призначений для введення в оману. Ключова мета посереднього ШІ-контен-

Таблиця 2

Вектори когнітивного впливу ШІ: міжнародний рівень

Країна	Вектор впливу
КНР	**Поширення технологічних продуктів та програмного забезпечення:** ініціатива Great Firewall містить підкомпоненти на основі ШІ, як-от ACF, алгоритмічні протоколи кібербезпеки та інструменти моніторингу вмісту; китайський технологічний гігант Huawei допоміг Еквадору розгорнути інтегровану систему спостереження ECU-911 [12]. **Проведення кібероперацій:** використання ШІ для автоматизації крадіжки даних, здійснення складних фішингових атак та аналізу наборів скомпрометованих даних. **Поширення дезінформації:** Китай використовував ботів Twitter для маніпулювання наративами щодо протестів у Гонконгу та пандемії COVID-19 [17]. **Формування глобальних кібернорм і технологічних стандартів, зокрема у сфері ШІ:** КНР активізував участь на міжнародних майданчиках (Міжнародний союз електрозв'язку), де просуває концепцію суворого державного контролю над цифровим простором, що відображає внутрішню політику КПК [15].
РФ	**Кібератаки на цифрові комунікації та інфраструктуру країни:** – Bad Grammar [2] – проросійська мережа, націлена на аудиторію Росії, України, США, Молдови та країн Балтії. Серед ключових тем: війна в Україні, політична ситуація в Молдові та країнах Балтії, внутрішня та зовнішня політика в США; – Doppelganger [2] – проросійська мережа, яка публікує антиукраїнський контент для аудиторій Європи та Північної Америки. Типові коментарі стверджують, що Росія залишається важливим економічним партнером для європейських країн; західні лідери, які критикують Росію, не мають зв'язку зі своїми виборцями; Україна слабка, корумпована та/або перебуває на межі поразки. **Поширення дезінформації:** – операція CopyCop [27]: синтез сфабрикованих статей з реальними новинами, вставленими в чатботи штучного інтелекту, які переписувалися в межах прокремлівських наративів, під виглядом легітимних медіа через фейкові вебсайти, як-от The London Crier у Великій Британії. Кампанія була зосереджена на загостренні внутрішніх дебатів та поглибленні політичного розколу, а не на створення нових наративів; – під час виборчої кампанії в ЄС експерти виявили, що понад 40 облікових записів TikTok використовували програмне забезпечення ШІ [29, с. 5] для розповсюдження політичної дезінформації щодо «фальсифікації» парламентських виборів та їх прокремлівської позиції щодо війни в Україні.
Іран	**Хакерські кампанії:** кібершпигунства націлені на урядовців, громадських активістів та приватний сектор з використанням фішингу, соціальної інженерії та розповсюдження шкідливого програмного забезпечення (кібератаки на американські фінансові установи та дамбу Боумен-авеню в Нью-Йорку [8]). **Психологічні операції впливу (Ізраїль, США та їх союзники):** Mint Sandstorm [31, с. 37], група, керована розвідувальним підрозділом Корпусу вартових ісламської революції, застосовує техніки мікротаргетування для генерування панічних настроїв серед певної соціальної групи. **Маніпуляції в соціальних мережах:** створення та розповсюдження контенту, керування обліковими записами і взаємодії з користувачами для посилення проіранських наративів щодо таких тем, як іранська ядерна програма чи регіональна політика. Запрограмовані боти націлюються на конкретні демографічні показники та регіони, адаптуючи вміст до культурного контексту різних груп аудиторії для максимізації впливу (фішинг Mint Sandstorm напередодні останніх президентських виборів у США; компрометація урядовців федерального рівня групою Peach Sandstorm). **Моніторинг та ідентифікація іранських дисидентів за кордоном:** профілювання та визначення місцеперебування іранської діаспори в Європі для здійснення кампаній залякування. У 2023 році Федеральне відомство з охорони конституції Німеччини опублікувало звіт з діяльності іранської розвідки щодо аналізу даних на основі ШІ для контролювання потенційних цілей за кордоном [31, с. 38]. **Цензура контенту:** спроба маніпулювати інформацією для придушення дискусій, пов'язаних з протестами в Ірані в січні 2023 року [11]; контроль над інформаційним середовищем через блокування спроб іранської діаспори надати VPN своїм співвітчизникам, які проживають в Ірані [16].

**розроблено авторкою статті.*

ту – відволікання уваги громадськості від якісних джерел інформації через «забруднення» інформаційного простору. Поширення політичного ШІ-контенту матиме складно прогнозовані наслідки для громадян, регуляторів, розробників ШІ та адміністрації соціальних мереж, особливо з модерування контенту. На переконання істеблішменту ЄС і НАТО, прогрес ШІ з його великими мовними моделями (LLM) виведе «маніпулювання інформацією» на новий рівень «гіперпереконування» [25, с. 4].

Незважаючи на низку нормативних заходів (Закон про цифрові послуги, Керівні принципи Комісії щодо пом'якшення системних ризиків для виборчих процесів [4]), європейські країни частіше звертаються до технічних рішень щодо регулювання генерованої ШІ дезінформації. Особлива увага приділяється розробленню концепції «когнітивної системи моніторингу та сповіщення» [25, с. 17] не стільки щодо змісту, скільки відносно суб'єктів, які проводять кампанії впливу. Проте через переважальну зосередженість на конвенційних методах ведення війни регулювання когнітивного виміру також потребує визначення делегітимних випадків у межах чинного міжнародного права, які можна маркувати як акти агресії проти іншої країни, включно з механізмами атрибуції, відповідальності за зловмисні дії у відповідь на інциденти когнітивної війни.

Список використаної літератури

1. Adebahr, C. & Mittelhammer, B. Upholding Internet Freedom as Part of the EU's Iran Policy, Carnegie Europe, 29 November 2023.
2. AI and Covert Influence Operations: Latest Trends, OpenAI, 2024. 39 p.
3. Alterman, J. Protest, Social Media, and Censorship in Iran, Center for Strategic and International Studies, 18 October 2022.
4. Artificial Intelligence Act: Council and Parliament Strike a Deal on the First Rules for AI in the World, Council of the European Union, 2023. URL: https://surli.cc/dytzib
5. Backes O., Swab A. Cognitive Warfare. The Russian Threat to Election Integrity in the Baltic States, Cambridge: Belfer Center for Science and International Affairs, 2019. URL: https://surl.li/jrxyac
6. Bazarkina D. & Pashentsev E. Malicious Use of Artificial Intelligence, *Russia in Global Affairs*. 2020. Vol. 18, No. 4. P. 154–177.
7. Beauchamp-Mustafaga N. Exploring the Implications of Generative AI for Chinese Military Cyber-Enabled Influence Operations, 2024, *RAND*, 23 p.
8. Berger J. A Dam, Small and Unsung, Is Caught Up in an Iranian Hacking Case, *The New York Times*, 25 March 2016.
9. China's Military Strategy in the New Era, The National Institute for Defense Studies, 2021. URL: https://surl.li/gveaya
10. Cognitive Warfare: Beyond Military Information Support Operations, Allied Command Transformation: NATO's Strategic Warfare Development Command, May 2023. URL: https://surl.li/jtedeb
11. Dave, P. Tech Workers Fight for Iran Protesters as Big Tech Plays It Safe, Wired, 20 January 2023.
12. Ellis, R. China's Economic Struggle for Position in Latin America, China Engages Latin America: Distorting Development and Democracy? 2022. P. 85–129.
13. George R. The AI Assault on Women: What Iran's Tech Enabled Morality Laws Indicate for Women's Rights Movements, Council on Foreign Relations, 7 December 2023.
14. Guangsheng Z., Yongli L., Haoxian W. A Brief Analysis of the Basics of Cognitive Domain Operations, PLA Daily, September 8, 2022. URL: http://www.81.cn/jfjbmap/content/2022-09/08/content_323692.htm.

15. Hillman J., Sacks D., Lew J. J. and Roughead G. China's Belt and Road: Implications for the United States, New York: Council on Foreign Relations, 2021.
16. Kilic D. & Chulain N. How Iranians are Hopping between VPNs to Stay Connected and Break through Internet Censorship, *EuroNews*, 6 November 2022.
17. King-WaF. Propagandization of Relative Gratification: How Chinese State Media Portray the International Pandemic. *Political Communication*. 2023. Vol. 40, No. 6. P. 788–809.
18. Koichiro, T. New Tech, New Concepts: China's Plans for AI and Cognitive Warfare, *War on the Rocks*, April 15, 2022. URL: https://surl.li/fwwkbr
19. Lahmann H. European Security and the Threat of **"**Cognitive Warfare**"**, Europe's Geopolitical Coming of Age: Adapting Law and Governance to Harsh International Realities, 2024. URL: https://surl.lu/fvzcyh
20. Lange L. Decoding China's AI-Powered "Algorithmic Cognitive Warfare". *Special Competitive Studies Project.* 2024. 11 p.
21. National Association of AI of Iran. URL: https://iranaiai.ir/en/
22. NATO Warfighting Capstone Concept, 2024. URL: https://surl.lt/bqygkc
23. Office of the Secretary of Defense, Annual Report to Congress: Military and Security Developments Involving the People's Republic of China, U.S. Department of Defense, 2023. P. 156.
24. Ottewell P. Defining the Cognitive Domain, *Over the Horizon*, 7 December 2020. URL: https://othjournal.com/2020/12/07/defining-the-cognitive-domain
25. Pujol I. & Dinh Q. The Battle for the Mind: Understanding and Addressing Cognitive Warfare and its Enabling Technologies. *IE CGC.* 2024. 21 p.
26. Puyosa I., Azpúrua A. & Suárez Pérez D. Venezuela: A Playbook for Digital Repression. *DFRLab*. 2024. 65 p.
27. Recorded Future, Russia-Linked CopyCop Uses LLMs to Weaponize Influence Content at Scale, May 2024. URL: https://go.recordedfuture.com/hubfs/reports/cta-2024-0509.pdf
28. Reczkowski R., Lis A. Cognitive Warfare: What is Our Actual Knowledge and How to Build State Resilience? *Security: Theory and Practice.* 2022. P. 51–61.
29. Stockwell S. AI-Enabled Influence Operations: Threat Analysis of the 2024 UK and European Elections. The Alan Turing Institute. Centre for Emerging Technology and Security. 2024. 31 p.
30. Thornton R. & Miron M. Towards the "Third Revolution in Military Affairs" the Russian Military's Use of AI-enabled Cyber Warfare. *The RUSI Journal.* 2020. Vol. 165, No. 3. P. 12–21.
31. Unver A. Artificial Intelligence (AI) and Human Rights: Using AI as a Weapon of Repression and its Impact on Human Rights, European Parliament's Subcommittee on Human Rights, 2024. 119 p.
32. Vainilavičius J. Russia Launches "Oculus" Tool to Monitor Banned Information Online, *CyberNews*, 15 November 2023.
33. Varanasi L. Putin Says Russia will Develop New AI Technology to Counter the Western Monopoly, which He Fears Could Lead to a "Digital Abolition" of Russian Culture, *Business Insider*, 26 November 2023.

ARTIFICIAL INTELLIGENCE AS A GAME CHANGER IN THE COGNITIVE WARFARE

Anastasiia Sychova
Reform Support Office of the Ministry of Defense of Ukraine,
Ivana Mazepy St., 1, 01010, Kyiv, Ukraine

The article analyzes the spectrum of cognitive influence tools through artificial intelligence (AI) technologies, which give an opportunity to produce and promote variations of false content with maximum speed, anonymity, and reach of the targeted audience. The author details the vectors of AI-cognitive focus, from the deep fake automatization to algorithmic surveillance mechanisms, focusing on the emergence of new forms of digital repression. The researcher explains the strengthening of cognitive influence by the ability of AI to quickly adapt to the personal preferences and cognitive biases of users, through the coherent construction of individual or group perception, tangential to the thinking vectors programmed into AI. The author proposes examples of AI implementation in cognitive operations of influence, appealing to the cases of the PRC, the Russian Federation, and Iran. The above countries demonstrate similarities in terms of institutionalization of AI technologies in national security strategies, as well as concerning areas of application, which allowed clustering cognitive influences at the local and international levels. Key vectors of the domestic political space in the cognitive context include monitoring public sentiment; blocking "unwilling" content; control over digital communications, and identification of potential dissidents. This trend indicates the escalation of digital surveillance capabilities by official institutions of the country as one of the steps to ensure the political regime stability. As for the international level, cognitive influences mainly focus on conducting cyber operations against the digital infrastructure and society of the targeted country; monitoring narratives in the international information space, with blocking "unfavorable" messages and the simultaneous dissemination of counter-narratives; disinformation campaigns to destabilize the political and military establishment, as well as maintaining control over diaspora representatives as sources of alternative information potentially dangerous to the ruling regime. The author notes the increasing importance of the cognitive dimension as one of the critical vectors of warfare by unconventional means, in particular due to the increasing persuasiveness of AI tools and their exponential application in such macrospheres as politics, economics, military industry, etc. The researcher emphasizes the need for legal regulation of the cognitive dimension, with the labeling of content created by AI, and the development of technological solutions capable of protecting the public from its negative effects.

Key words: cognitive dimension, cognitive warfare, artificial intelligence (AI), China, Russia, Iran.

Дата першого надходження рукопису до видання: 22.09.2025
Дата прийнятого до друку рукопису після рецензування: 23.10.2025
Дата публікації: 23.12.2025