

and Technology. NIST AI 100-2 E2025. Gaithersburg, MD, 2025. DOI: 10.6028/NIST.AI.100-2e2025.

5. MITRE ATLAS: Adversarial Threat Landscape for Artificial-Intelligence Systems / The MITRE Corporation. URL: <https://atlas.mitre.org/> (дата звернення: 20.05.2026).

6. OWASP Top 10 for Large Language Model Applications 2025 / OWASP Foundation. 2025. URL: <https://genai.owasp.org/resource/owasp-top-10-for-llm-applications-2025> / (дата звернення: 20.05.2026).

7. Kalai A. T., Nachum O., Vempala S. S., Zhang E. Why Language Models Hallucinate. arXiv:2509.04664v1. 2025. URL: <https://arxiv.org/pdf/2509.04664> (дата звернення: 20.05.2026).

8. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1 / National Institute of Standards and Technology. Gaithersburg, MD, 2023. DOI: 10.6028/NIST.AI.100-1.

9. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. L, 2024/1689, 12.07.2024. URL: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj> (дата звернення: 20.05.2026).

10. Загрози та ризики застосування технологій штучного інтелекту. Кібербезпека: освіта, наука, техніка. 2025. URL: <https://csecurity.kubg.edu.ua/index.php/journal/article/view/520> (дата звернення: 20.05.2026).

Роман ШИРШОВ, старший викладач, Національна академія Служби безпеки України

ВИКОРИСТАННЯ ШТУЧНОГО ІНТЕЛЕКТУ В РОЗВІДЦІ ЗА ВІДКРИТИМИ ДЖЕРЕЛАМИ (OSINT): МОЖЛИВОСТІ, ІНСТРУМЕНТИ ТА ОБМЕЖЕННЯ

Вступ. Розвідка за відкритими джерелами (Open-Source Intelligence, OSINT) є основою сучасної аналітичної діяльності правоохоронних і безпекових структур: за оцінками, OSINT забезпечує від 80 до 90% усього розвідувального матеріалу західних служб [1]. Стрімке зростання обсягів цифрових даних – соціальних медіа, супутникових знімків, IoT-сенсорів, корпоративних реєстрів, медіа-архівів – зробило традиційні ручні підходи неспроможними обробляти інформаційний потік у режимі реального часу. Світовий ринок OSINT-рішень оцінюється у 14,85 млрд дол. США (2024) із прогнозом зростання до 49,39 млрд дол. до 2029 року (CAGR близько 28%), і значна частина цього зростання припадає саме на AI-підсилені платформи [2]. Систематичний огляд Browne та ін. (2024), що охопив 163 рецензованих публікації за період після 2019 року, фіксує лавиноподібне зростання застосування ШІ на всіх етапах OSINT-циклу [3]. Метою тез є системний огляд застосувань ШІ у задачах OSINT, аналіз ключових інструментальних напрямів і обмежень, а також формулювання вимог до методології безпечного впровадження.

Етапи OSINT-циклу, в яких задіяний ШІ. Класичний цикл - постановка завдання, збір, обробка, аналіз, поширення - сьогодні насичений AI-компонентами на кожному кроці. **Збір** автоматизовано через скрейпінг з LLM-навігацією, моніторинг соціальних мереж і dark web, обробку нестандартних джерел. **Обробка** – нормалізація, дедуплікація, переклад, розпізнавання іменованих сутностей. **Аналіз** – кластеризація, побудова графів зв'язків, виявлення аномалій, прогнозування. **Поширення** – автоматична генерація звітів, інтерактивних дашбордів і брифінгів. Огляд Browne та ін. показує, що 44% досліджень з AI-OSINT походять зі США та Індії; домінують методи NLP, глибокого навчання і графових неймереж [3].

III у роботі з текстовою інформацією. Засоби обробки природної мови (NLP) вирішують задачі розпізнавання іменованих сутностей (NER) – людей, організацій, локацій, торгових марок, ідентифікаторів транспорту; машинного перекладу для мультимовного OSINT; реферування й тематичного моделювання; аналізу тональності та поляризації дискурсу; розв’язання сутностей (entity resolution) для зведення профілів у різних соціальних мережах. Окремий клас задач – LLM-агенти-дослідники, які автоматично виконують багатокрокові запити: декомпозують питання, виконують пошук, синтезують результат. Експериментальне дослідження Karabiyik та ін. (2024) показало, що сучасні генеративні моделі (Gemini, GPT-4) демонструють значущий OSINT-потенціал, але водночас містять вбудовані фільтри приватності, які обмежують доступ до персональних даних [4].

Агентний OSINT та автоматизація розслідувань. Новою хвилею є агентні системи (agentic AI) на основі LLM, що автономно планують і виконують багатокрокові розслідування: декомпозують завдання, обирають інструменти (пошук, скрейпінг, бази даних, API), ітеративно уточнюють гіпотези й самостійно синтезують звіт із посиланнями на джерела. Архітектури з доповненою пошуком генерацією (retrieval-augmented generation, RAG) прив’язують відповіді моделі до перевірених документів, знижуючи рівень галюцинацій. Такі системи інтегруються з класичними OSINT-платформами – графовими аналізаторами зв’язків, інструментами візуалізації та геопросторовими модулями, – перетворюючи аналітика з виконавця рутинних запитів на верифікатора й інтерпретатора. Browne та ін. окремо відзначають інтеграцію наявних OSINT-інструментів з III як ключову, але досі недостатньо опрацьовану дослідницьку прогалину [3]. Водночас зростання автономності агентів підвищує ризики: помилка на одному кроці поширюється ланцюгом міркувань, а надмірні повноваження агента (excessive agency) створюють нові вектори атак.

Візуальний OSINT (IMINT/GEOINT). Машинне навчання радикально розширило можливості роботи з зображеннями та відео: класифікація об’єктів і логотипів, автоматичне розпізнавання військової техніки, типів літальних апаратів і кораблів; геолокація за зображенням (cross-view matching, прив’язка до супутникових знімків) і хронолокація (визначення часу за тінями); виявлення дублікатів і модифікацій (reverse image search); розпізнавання облич; аналіз супутникових знімків (зміна забудови, рух військ, оцінка пошкоджень). Зростання обсягів мультимедіа в соціальних мережах (SOCMINT) робить ручний перегляд тисяч фото й відео непосильним – AI-моделі дозволяють автоматично відсіювати нерелевантний контент і виокремлювати значущі кадри. У роботі з морськими даними AIS і трекінгом транспорту AI-моделі виявляють патерни підозрілої поведінки – «темні» зупинки сигналу, типові схеми обходу санкцій, перевалки в морі. Цей напрям успішно застосовувано у відомих розслідуваннях Bellingcat – від збиття MH17 до сирійського конфлікту й війни в Україні; у 2024 році опубліковано Bellingcat Online Investigation Toolkit, який системно каталогізує AI- та не-AI-інструменти для практиків [5].

Корпоративна розвідка та фінансовий комплаєнс. Одним із найдинамічніших напрямів прикладного AI-OSINT є корпоративна розвідка, due diligence і протидія відмиванню коштів. AI-платформи автоматично агрегують дані з корпоративних реєстрів, судових баз, санкційних списків (OFAC, ЄС, ООН, РНБО), реєстрів кінцевих бенефіціарних власників (UBO) та adverse-media джерел, будуючи графи власності й контролю для виявлення прихованих зв’язків і кінцевих вигодоодержувачів. Так, платформа S&P Global Entity Insights у партнерстві з Quantifind забезпечує скринінг понад 30 млн записів про сутності – контролерів, топ-менеджерів, членів рад і власників – за санкціями та негативними згадками з понад 30 тис. джерел [6]. Машинне навчання істотно знижує частку хибних збігів при санкційному скринінгу й скорочує час перевірки бенефіціарної структури з десятків годин до кількох; предиктивна аналітика дозволяє переходити від статичної перевірки відповідності до безперервного моніторингу зміни ризик-профілю контрагента. Цей напрям критично важливий для виявлення схем обходу санкцій – зокрема через ланцюги підставних компаній.

Інструментальна екосистема. Ринок пропонує широкий спектр AI-підсиленних OSINT-платформ: комплексні розвідувальні рішення з інтегрованими GEOINT- і SOCMINT-модулями, інструменти аналізу й візуалізації графів зв'язків (наприклад, Maltego), платформи корпоративної розвідки та перевірки бенефіціарних власників (зокрема Sayari, YouControl), агрегатори цифрових слідів за ідентифікаторами (OSINT Industries). Систематизовані каталоги таких інструментів – як-от Bellingcat Online Investigation Toolkit – спрощують підбір рішень під конкретне завдання [5]. Стійкою тенденцією є вбудовування LLM-асистентів безпосередньо в робочі процеси аналітика: для автоматичного реферування зібраних даних, формування запитів і чернеток звітів. Ефективність екосистеми визначається не окремим інструментом, а зрілістю методології та якістю інтеграції джерел і верифікації результатів.

Виявлення дезінформації та deepfake. Генеративний ШІ створив нову категорію викликів: фабриковані зображення, deepfake-відео, синтетичні голоси, AI-генеровані «свідчення». ENISA фіксує, що на початок 2025 року вже понад 80% спостережуваних соціально-інженерних кампаній спираються на ШІ для генерації контенту; з'явилися «зловмисні» моделі (Xanthorox AI) для масової автоматизації таких атак [7]. Контрзаходи поєднують ШІ-класифікатори згенерованого контенту, аналіз метаданих, верифікацію джерел і ланцюгів поширення. Показовим є український досвід: платформи Osavul і Mantis Analytics, створені у перші місяці повномасштабного вторгнення, застосовують LLM- та NLP-моделі для виявлення нарративів, координованої неавтентичної поведінки (coordinated inauthentic behaviour) і ботомереж у режимі, близькому до реального часу. Модулі CommSecure (виявлення нарративів) і CIB Guard (аналіз координації акаунтів) дозволяють викривати масштабні російські інформаційні операції на ранніх стадіях; ключова перевага - швидкість, за якої правдива інформація встигає досягти аудиторії раніше за ворожий нарратив [8]. Український досвід засвідчує, що ефективна протидія AI-дезінформації потребує поєднання трьох компонентів - запобігання, виявлення і реагування [8].

Виклики та обмеження. Попри прогрес, ШІ в OSINT має низку принципових обмежень. **По-перше** – галюцинації LLM-агентів: вигадані факти, неіснуючі джерела, фальшиві посилання, помилкова атрибуція цитат; як показано в [9], це не дефект конкретної моделі, а внутрішня властивість самого процесу навчання й оцінювання мовних моделей. **По-друге** - adversarial-загрози: отруєння публічних даних, prompt injection через скомпрометовані веб-сторінки, evasion-атаки на класифікатори – формалізовано в таксономії NIST AI 100-2 E2025 [10]. **По-третє** – операційна безпека: запити до публічних LLM-сервісів самі по собі є витком оперативної інформації аналітика. **По-четверте** – правові й етичні рамки: робота з персональними даними, біометрією й обличчями має узгоджуватись з GDPR, EU AI Act (з диференційованими обов'язками для високоризикових систем) [11] та національним законодавством. Систематичний огляд Ghioni та ін. формалізує управлінсько-етичні (GELSI) обмеження AI-OSINT і виявляє суттєві прогалини в нормативному регулюванні [1]. Український контекст ризиків застосування ШІ в безпекових задачах розглянуто у [12].

Висновки. ШІ якісно змінив можливості OSINT – від швидкості й масштабу збору до глибини аналізу - і фактично став обов'язковим компонентом сучасної аналітичної практики в безпековій, корпоративній та журналістській сферах. Водночас він не замінює аналітика: критичне мислення, верифікація, контекстне розуміння й операційна обережність залишаються прерогативою людини. Безпечне впровадження ШІ в OSINT потребує: збереження «людини в контурі» для всіх верифікаційних рішень; систематичної перевірки джерел, отриманих через LLM, проти первинних даних; розмежування публічних і приватних AI-сервісів за чутливістю запитів; інтеграції методології NIST AI RMF та таксономій NIST AI 100-2 / MITRE ATLAS у процеси оцінювання інструментарію; формування національних освітніх стандартів з AI-OSINT і протидії синтетичній дезінформації. Перспективними напрямками є розвиток пояснюваних і верифікованих агентних архітектур, методів виявлення синтетичного контенту та національної екосистеми підготовки аналітиків.

Список використаних джерел

1. Ghioni R., Taddeo M., Floridi L. Open source intelligence and AI: a systematic review of the GELSI literature. *AI & Society*. 2023. DOI: 10.1007/s00146-023-01628-x.
2. The OSINT Market: AI Integration Shaping Intelligence in 2025 / SpecialEurasia. March 2025. URL: <https://www.specialeurasia.com/2025/03/24/osint-market-ai-integration/> (дата звернення: 20.05.2026).
3. Browne T. O., Abedin M., Chowdhury M. J. M. A systematic review on research utilising artificial intelligence for open source intelligence (OSINT) applications. *International Journal of Information Security*. 2024. Vol. 23, Issue 4. P. 2911–2938. DOI: 10.1007/s10207-024-00868-2.
4. Karabiyik U., Ozer A., Lima V. Generative AI and Open-Source Intelligence: Exploring Capabilities and Privacy Implications / Purdue University, CERIAS Annual Security Symposium 2024. 2024. URL: <https://www.cerias.purdue.edu/symposium/index.php/posters/year/2024/R00-I3A> (дата звернення: 20.05.2026).
5. Wild J. Bellingcat's Online Investigation Toolkit / Bellingcat. 2024. URL: <https://bellingcat.gitbook.io/toolkit> (дата звернення: 20.05.2026).
6. S&P Global Market Intelligence Enhances Entity Insights Solution with Addition of Sanctions and Adverse Media Screening Data / S&P Global Market Intelligence. 24.01.2024. URL: <https://press.spglobal.com/2024-01-24-S-P-Global-Market-Intelligence-Enhances-Entity-Insights-Solution-with-Addition-of-Sanctions-and-Adverse-Media-Screening-Data> (дата звернення: 20.05.2026).
7. ENISA Threat Landscape 2025 / European Union Agency for Cybersecurity. October 2025. 87 p. URL: <https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025> (дата звернення: 20.05.2026).
8. Marushchak A., Petrov S., Khoperiya A. Countering AI-powered disinformation through national regulation: learning from the case of Ukraine. *Frontiers in Artificial Intelligence*. 2025. Vol. 7, Art. 1474034. DOI: 10.3389/frai.2024.1474034.
9. Kalai A. T., Nachum O., Vempala S. S., Zhang E. Why Language Models Hallucinate. arXiv:2509.04664v1. 2025. URL: <https://arxiv.org/pdf/2509.04664> (дата звернення: 20.05.2026).
10. Vassilev A., Oprea A., Fordyce A., Anderson H., Davies X., Hamin M. Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations / National Institute of Standards and Technology. NIST AI 100-2 E2025. Gaithersburg, MD, 2025. DOI: 10.6028/NIST.AI.100-2e2025.
11. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. L, 2024/1689, 12.07.2024. URL: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj> (дата звернення: 20.05.2026).
12. Загрози та ризики застосування технологій штучного інтелекту. Кібербезпека: освіта, наука, техніка. 2025. URL: <https://csecurity.kubg.edu.ua/index.php/journal/article/view/520> (дата звернення: 20.05.2026).