КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (34), 2026
ISSN 2663 – 4023
DOI 10.28925/2663-4023.2026.34.1338
УДК 004.056.5

Мельничук Максим Романович
студент кафедри захисту інформації
Національний університет «Львівська Політехніка», Львів, Україна
ORCID: 0009-0009-8606-2029
maksym.melnychuk.kb.2025@lpnu.ua

Опірський Іван Романович
д.т.н., проф., завідувач кафедри захисту інформації
Національний університет «Львівська Політехніка», Львів, Україна
ORCID: 0000-0002-8461-8996
ivan.r.opirskyi@lpnu.ua

АВТОМАТИЗОВАНИЙ АНАЛІЗ МЕТАДАНИХ ЯК ВЕКТОР ВИТОКУ ІНФОРМАЦІЇ

Анотація. У статті досліджується фундаментальна проблема сучасної кібербезпеки — автоматизований збір та аналіз метаданих, які функціонують як «невидиме цифрове ДНК» електронних документів, медіафайлів та мережевих комунікацій. Розвінчується небезпечна «ілюзія безпеки», за якої користувачі та організації зосереджуються переважно на криптографічному захисті видимого контенту, ігноруючи приховані транзитні, описові та структурні артефакти. Робота комплексно аналізує специфічні архітектурні вразливості популярних форматів (Microsoft Office, PDF, JPEG) і на основі реальних інцидентів демонструє, як інтеграція OSINT-інструментів з алгоритмами штучного інтелекту перетворює ці цифрові сліди на зброю для деанонімізації, фізичної розвідки та підготовки високоточних фішингових атак (Spear Phishing). Для практичного підтвердження теоретичної бази у статті наведено результати розширеного експериментального тестування вибірки зі 100 цифрових об'єктів 9 різних форматів. Дослідження виявило критичний рівень фонового витоку чутливої інформації (успішна деанонімізація авторів у 77% випадків). З метою алгоритмізації процесів захисту розроблено та статистично валідовано математичну модель кількісної оцінки загроз — Metadata Information Risk Score (MIRS), точність якої підтверджено метрикою F_1-score на рівні 98%. Окремо доведено, що традиційні засоби базового очищення (Data Scrubbing) є неефективними для глибоких XML-структур офісних документів. Підсумовуючи, для протидії цим загрозам обґрунтовано необхідність багаторівневих оборонних стратегій: від технологій тотальної деконструкції контенту (CDR) та стандартів деідентифікації NIST до фундаментальної імплементації концепції «Privacy by Design» (приватність за замовчуванням).

Ключові слова: метадані, кібербезпека, витік інформації, деанонімізація, OSINT-інструменти, цифрове профілювання, соціальна інженерія, Spear Phishing, Privacy by Design, Content Disarm and Reconstruction (CDR).

ВСТУП

В епоху глобальної цифровізації спостерігається безпрецедентне зростання обсягів цифрових даних, коли у світі щодня генерується понад 2.5 квінтильйона байтів інформації. Разом із цим експоненційним зростанням відбувається стрімке поширення інструментів автоматизації та OSINT-розвідки, що дозволяють масово збирати, агрегувати та аналізувати ці масиви з мінімальним втручанням людини. Проте такі технологічні зміни формують не лише нові можливості для швидкої обробки інформації, але й критичні виклики для захисту конфіденційних даних.

Фундаментальна проблема полягає в тому, що користувачі та корпоративні структури часто стають жертвами когнітивного упередження та небезпечної ілюзії безпеки. Традиційні парадигми кібербезпеки орієнтовані переважно на криптографічний захист корисного навантаження — тобто основного видимого контенту документів чи мережевих комунікацій. Водночас метадані, або «інформація, що описує характеристики даних», масово ігноруються, оскільки вони генеруються апаратним та програмним забезпеченням фоново і непомітно для людини. Користувач схильний помилково вважати інформацію

безпечною, якщо візуально не спостерігає компрометуючих даних на екрані. Проте ці побічні цифрові артефакти функціонують як невидиме цифрове ДНК, здатне розкрити повний контекст операції (географію, час, інструментарій та імена авторів) незалежно від того, чи зашифрований сам файл.

Аналіз останніх досліджень і публікацій. Фундаментальну основу концептуальних засад інформаційної безпеки та управління метаданими закладено у стандартах Національного інституту стандартів і технологій США. Зокрема, документи NIST SP 800-53 [1] та NIST SP 800-188 [20] формують комплексну рамку для деідентифікації масивів даних і визначають необхідність глибокої трансформації квазі-ідентифікаторів. Вони досі залишаються базовими для розробки будь-яких організаційних політик безпеки (OPSEC) та архітектури приватності [21].

Значний масив сучасних досліджень фокусується на загрозах автоматизованого збору транзитних та описових артефактів. У своїх роботах Eckersley P. [3] та Wu Z. зі співавторами [4] системно описують механізми витоку даних через відбитки браузера (Browser Fingerprinting), доводячи, що мікроархітектурні розбіжності обладнання формують унікальний ідентифікатор користувача. Практичні аспекти застосування OSINT-інструментів для профілювання детально розкривають Hou S. та ін. [7], які відстежували поведінкові патерни через метадані публічних торрент-мереж, а також Eskandari S. з колегами [6], що виявили загрози деанонімізації через фоновий трафік браузерних Web3-гаманців. Для візуалізації та кореляції таких розрізнених масивів на практиці традиційно використовуються інтелектуальні платформи на кшталт Maltego [5].

Окремий напрям досліджень присвячено архітектурним вразливостям конкретних файлових форматів. Дослідження Kurz H. та ін. [13] детально аналізує вразливість CVE-2022-25641 [14] та розкриває клас загроз «Тіньові атаки» (Shadow Attacks) у криптографічно підписаних PDF-документах. Автори доводять, що маніпуляція прихованими таблицями перехресних посилань дозволяє змінити відображуваний контент без порушення валідності самого підпису. Водночас аналіз резонансних інцидентів, таких як масове розкриття військових баз через публічні теплові карти Strava [10, 11] та ідентифікація особи за прихованими атрибутами документів [15, 16], підтверджує, що геопросторові та адміністративні метадані еволюціонували від віртуальної проблеми до засобу кінетичного впливу.

Сьогодні також спостерігається стрімка еволюція векторів атак через інтеграцію збору метаданих із генеративним штучним інтелектом. Роботи Roy S. [8] та аналітичні звіти компанії Vectra AI [9] переконливо доводять, що використання корпоративних артефактів для підготовки цілеспрямованих атак (Spear Phishing) за допомогою великих мовних моделей безпрецедентно підвищує їхню переконливість, досягаючи показника успішних кліків на рівні 54%.

У площині протидії цим загрозам література переважно розглядає інструменти санації контенту. Для індивідуальної кібергігієни виділяють утиліти глибокого криміналістичного очищення на кшталт MAT2 [17], тоді як корпоративний сегмент покладається на системи DLP з функцією адаптивного редагування (Adaptive Redaction) [18]. Однак передовою технологічною еволюцією експерти все частіше називають системи деконструкції та реконструкції контенту (CDR) [19], що функціонують за принципом нульової довіри та зумовлюють глобальний перехід індустрії [22] до парадигми «Privacy by Design».

Узагальнюючи наведені дослідження, можна констатувати, що наукова спільнота чітко ідентифікує небезпеку автоматизованого збору метаданих, вразливість популярних форматів та необхідність впровадження процедур деідентифікації. Проте в існуючих працях спостерігається суттєва прогалина: відсутня комплексна алгоритмізована модель кількісної оцінки ризиків для об'єктивної класифікації цифрових об'єктів. Крім того, бракує розширеного експериментального доведення того, наскільки базові засоби очищення (Data Scrubbing) є концептуально неефективними проти глибоких XML-структур офісних документів. Заповнення цих прогалин шляхом розробки математичної моделі MIRS та її валідації на вибірці зі 100 файлів формує мету та ключову новизну даної статті.

Метою цієї статті є комплексне дослідження того, як саме автоматизований збір та кореляційний аналіз розрізнених метаданих створює критичні вразливості для приватності осіб та корпоративної безпеки. Робота покликана не лише продемонструвати перетворення цих невидимих слідів на потужну зброю для деанонімізації та підготовки високоточних кібератак, але й експериментально оцінити реальний рівень розкриття прихованої інформації на розширеній вибірці зі 100 повсякденних документів 9 популярних форматів. Завдяки розробці та статистичній валідації математичної моделі оцінки ризиків Metadata Information Risk Score (MIRS), стаття має на меті перевірити ефективність алгоритмів класифікації загроз (із застосуванням матриці помилок та розрахунком метрик точності). Крім того, дослідження експериментально порівнює результативність традиційних засобів очищення (Data Scrubbing) із новітніми підходами, щоб обґрунтувати алгоритмічний фундамент для автоматизованої санації контенту та практичного впровадження концепції «Privacy by Design» (приватність за замовчуванням) на всіх рівнях цифрової взаємодії.

Для досягнення мети сформульовано задачі:

1) проаналізувати специфічні архітектурні вразливості популярних форматів (Microsoft Office, PDF, JPEG) та механізми прихованого витоку транзитних, описових і структурних артефактів;

2) провести розширене експериментальне тестування на вибірці зі 100 цифрових об'єктів 9 різних форматів для оцінки реального рівня фонового розкриття конфіденційної інформації;

3) розробити математичну модель кількісної оцінки загроз — Metadata Information Risk Score (MIRS) — для автоматизованої класифікації файлів за рівнем ризику;

4) статистично валідувати запропоновану модель за допомогою експертної оцінки та розрахунку метрик ефективності (Accuracy, Precision, Recall, F1-score);

5) експериментально порівняти результативність традиційних засобів базового очищення (Data Scrubbing) із новітніми підходами для обґрунтування необхідності застосування технологій деконструкції та реконструкції контенту (CDR).

Новизна дослідження полягає у переході від виключно описового аналізу вразливостей метаданих до їхньої математичної формалізації та алгоритмізації через модель кількісної оцінки MIRS, а також в експериментальному доведенні концептуальної недостатності базових засобів очищення (Data Scrubbing) для глибоких XML-структур офісних документів, що формує науково обґрунтований базис для практичного впровадження систем Content Disarm and Reconstruction (CDR) та концепції «Privacy by Design».

ОСНОВНА ЧАСТИНА

Природа та класифікація метаданих

У класичному інформаційному менеджменті метадані визначаються просто як «інформація, що описує характеристики даних». Однак у контексті сучасної кібербезпеки це поняття набуває значно глибшого змісту. Згідно з концептуальними засадами Національного інституту стандартів і технологій США (NIST SP 800-53), метадані класифікуються на структурні (ті, що описують архітектуру об'єкта) та описові (ті, що безпосередньо описують семантику контенту)[1].

Кібербезпека вимагає від організацій забезпечення беззаперечної достовірності (trustworthiness) метаданих, що включає захист їхньої точності та цілісності. Таким чином, метадані — це системний артефакт, що неминуче генерується під час створення чи транзиту інформації. Цей артефакт здатний розкрити повний контекст операції (географію, час, інструментарій) незалежно від того, чи зашифрований сам контент.

Таблиця 1

Роль та вразливості різних класів метаданих згідно зі стандартами кібербезпеки

Клас метаданих	Функціональне призначення	Приклад атрибутів	Вектор загрози для безпеки
Описові (Descriptive)	Ідентифікація ресурсу для індексації.	Ім'я автора, заголовок, ключові слова.	Деанонімізація особи, розкриття тематики зашифрованого документа.
Структурні (Structural)	Опис архітектури, зв'язків між об'єктами.	Версія формату, структура XML, мапінг.	Ідентифікація версій ПЗ для підготовки специфічних експлойтів.
Адміністративні (Administrative)	Технічне управління, контроль доступу.	Дата створення, RSID, історія правок.	Побудова хронології (Timeline analysis), виявлення корпоративної мережі.

Практична реалізація векторів витоку метаданих суттєво залежить від специфіки цифрової екосистеми та архітектури конкретних типів файлів. Насамперед це стосується корпоративного документообігу та форматів Microsoft Office.

Сучасні формати Office (.docx, .xlsx) технічно є звичайними ZIP-архівами, що містять набір ієрархічно структурованих XML-файлів. Будь-який такий документ можна розпакувати, отримавши прямий доступ до внутрішньої директорії docProps. У ній відкрито зберігаються атрибути: точне ім'я творця, час модифікацій та абсолютні шляхи до локальних дисків при вставці об'єктів OLE (наприклад, C:\Users\Admin\Desktop\SecretProject\image.png). Окремим вектором профілювання є ідентифікатори збереження ревізій — RSID (Revision Save Identifiers). Офісні пакети використовують їх для відстеження

змін під час спільної роботи. Це дозволяє криміналістам розпакувати документ і побудувати детальну картину того, як саме він створювався, які частини тексту були скопійовані, а які додані різними авторами.

На відміну від відкритої ієрархічної структури Office, формат PDF реалізує іншу парадигму збереження даних, що формує власні специфічні вразливості. PDF містить складну об'єктну модель. Традиційні словники метаданих сьогодні замінюються на XMP (Extensible Metadata Platform) — технологію на базі XML, що вбудовується у потік PDF і розкриває точну версію ПЗ генератора. Найкритичнішою вразливістю PDF, яка становить критичну загрозу конфіденційності, є механізм «інкрементальних оновлень» (incremental updates). Під час кожної зміни програма не переписує файл, а додає нові об'єкти в його кінець. Якщо користувач намагається «анонімізувати» документ, наклавши чорний прямокутник на текст або видаливши сторінку, оригінальний текст залишається недоторканим у попередній секції файлу і може бути легко відновлений криміналістичними утилітами.

Якщо в електронних документах основна загроза криється в історії модифікацій та ревізіях тексту, то мультимедійні файли розширюють площину атак через прив'язку до апаратних та геопросторових маркерів. Так, у цифрових зображеннях домінуючим стандартом зберігання метаданих є EXIF, масив якого у форматі JPEG обмежений 64 кілобайтами[2]. EXIF розкриває надзвичайно чутливі дані: точні GPS-координати, налаштування експозиції та навіть унікальний серійний номер камери. Поширеною вразливістю є невідповідність (inconsistency) вбудованих мініатюр. Для прискорення відображення камери вбудовують стиснуті прев'ю (thumbnails) безпосередньо у масив EXIF. Якщо користувач замальовує обличчя на фото у базовому редакторі, програма часто змінює лише основний масив пікселів, залишаючи оригінальну мініатюру EXIF недоторканою. Крім того, сучасні методи аналізу базуються на порівнянні відбитків таблиць квантування (DQT fingerprint) основного зображення та мініатюри. Будь-яка розбіжність є криптографічним підтвердженням того, що фотографія була змінена у редакторі.

Окрім статичних артефактів, що зберігаються безпосередньо у локальних файлах, критичним вектором деанонімізації є транзитні дані комунікацій. Мережеві метадані утворюються динамічно під час кожної взаємодії з інтернетом. Навіть при використанні захищених протоколів сам факт комунікації розкриває транзитні дані. Потужним вектором витоку є відбитки браузера (Browser Fingerprinting). Завдяки стандартам на кшталт WebGL, вебсайти змушують відеокарту користувача відрендерити невидиму графічну фігуру. Мікроархітектурні розбіжності у виконанні математичних операцій з плаваючою комою різними відеокартами дають унікальний апаратний хеш-підпис пристрою, який майже неможливо змінити[4]. Більше того, масштабні дослідження свідчать про ризик соціального профілювання: на основі HTTP-заголовків (наприклад, атрибута Languages та параметрів екрана) алгоритми роблять високоточні висновки щодо статі, віку та етнічної приналежності користувача.

Таблиця 2

Специфікація прихованих витоків інформації за типами файлів та протоколів

Категорія	Формат / Протокол	Ключові елементи деанонімізації	Технічний механізм прихованого витоку
Документи	OpenXML (.docx)	RSID, docProps, шляхи OLE	Збереження імен авторів та шляхів ПК у прихованих ZIP-структурах.
Документи	PDF	XMP потоки, Incremental updates	Інкрементальне збереження залишає видалений текст у застарілих таблицях.
Медіафайли	JPEG (EXIF)	GPS, Серійний номер камери	Неузгодженість між зміненим основним зображенням та вбудованою мініатюрою (DQT).
Мережі	Browser Fingerprint	HTTP-заголовки, WebGL	Мікроархітектурні розбіжності у обчисленнях формують унікальний підпис обладнання.

Комплексний аналіз наведених форматів та протоколів свідчить про те, що реальні масштаби генерації метаданих часто залишаються поза увагою користувача. Концепція «ілюзії анонімності» в кібербезпеці базується на когнітивному упередженні: користувач схильний вважати інформацію безпечною, якщо візуально не спостерігає компрометуючих даних. Проте у цифровому світі видимий контент є лише поверхневим рівнем представлення даних. Ця ілюзія руйнується технічною реальністю агрегації даних. Сучасні алгоритми машинного навчання працюють за принципом об'єднання розрізнених

фрагментів. Зловмисник може проаналізувати XMP-потік документа, відновити чорнові версії тексту з інкрементальних оновлень PDF та вивчити RSID для побудови графа соціальних зв'язків. Навіть використання систем анонімізації мережі (наприклад, VPN) залишає сліди. Відмінність між часовим поясом, який транслюється через HTTP-заголовки, та часовим поясом, жорстко закодованим у метаданих завантаженого зображення, дозволяє обійти мережевий захист і визначити реальне місцеперебування особи[3]. Підсумовуючи, метадані є критичною площиною безпеки — невидимим цифровим ДНК. Тільки інтеграція концепції "Privacy by Design" (приватність за замовчуванням) на етапі розробки систем здатна мінімізувати ці ризики.

Автоматизація збору: Інструменти та методи.

Як було продемонстровано у попередньому розділі, метадані містять багато прихованої інформації, яка генерується автоматично. Раніше для її отримання зловмисникам чи дослідникам доводилося діяти здебільшого вручну: завантажувати кожен документ окремо, відкривати його властивості та шукати імена авторів, історію редагувань чи час створення. Оскільки у сучасному світі щодня генерується понад 2.5 квінтильйона байтів даних, виключно ручний підхід втратив свою релевантність через неможливість масштабування та високі часові затрати. Автоматизація також усуває людський фактор і гарантує послідовність — програма не втомиться і не пропустить жодної важливої деталі. Тепер збір прихованої інформації відбувається автоматизовано. Цей процес є важливою складовою OSINT — розвідки на основі відкритих джерел. Автоматизація дозволяє програмам за лічені хвилини знайти, завантажити та проаналізувати тисячі файлів із різних сайтів компаній. Завдяки цьому метадані перестали бути просто дрібними залишками інформації у файлі — вони перетворилися на зручний матеріал для масового збору даних та створення детальних профілів користувачів.

Для автоматичного збору «даних про дані» кіберфахівці та зловмисники використовують спеціальні програми. Кожна з них найкраще підходить для певного типу файлів і виконує свою частину роботи. Наприклад, одні інструменти шукають приховані записи у звичайних офісних документах, а інші — геолокаційні координати у фотографіях. На практиці зазначені інструментальні засоби інтегруються у комплексні автоматизовані пайплайни. Наприклад, спеціаліст може налаштувати Metagoofil для завантаження сотень документів із сайту, а потім миттєво пропустити їх через ExifTool для масового аналізу інформації.

Таблиця 3

Характеристика базових утиліт для автоматизованого збору метаданих

Назва інструменту	Цільові формати файлів	Основний функціонал у розвідці (OSINT)
ExifTool	Зображення, відео, PDF (понад 200 форматів)	Швидко витягує з файлів EXIF-дані: модель камери, налаштування, час зйомки та точні GPS-координати.
Metagoofil	Публічні документи (PDF, Word, Excel, PowerPoint)	Автоматично шукає файли на заданому сайті, завантажує їх і екстрагує ідентифікаційні дані користувачів, версії програм та шляхи збереження на комп'ютерах користувачів.
FOCA	Документи, файли вебсайтів, мережеві архіви	Аналізує завантажені документи, щоб знайти приховану інформацію про внутрішню мережу компанії (імена користувачів, моделі принтерів, операційні системи).

Окрім використання готових програм, дуже популярним є створення власних автоматизованих скриптів. Найчастіше їх пишуть на мові програмування Python, оскільки для неї існує багато готових і зручних бібліотек. Наприклад, бібліотеку PyPDF2 використовують для аналізу текстових документів, а Pillow — для зображень. Також студенти та дослідники часто звертаються до модуля exifread для роботи з фотографіями та pdfminer.six, який дозволяє витягувати приховані дати створення навіть із пошкоджених PDF-документів. Сьогодні достатньо написати невеликий скрипт, який буде масово та непомітно збирати інформацію з відкритих сторінок компаній у соцмережах або на їхніх офіційних сайтах.

Проте первинна агрегація розрізнених масивів метаданих є лише початковим етапом розвідки. Набагато складніше й важливіше правильно поєднати ці розрізнені дані між собою. Для кореляції знайдених патернів та виявлення прихованих залежностей базового скриптингу стає недостатньо. На цьому етапі доцільним є впровадження методів штучного інтелекту (ШІ) та алгоритмів машинного навчання. Штучний інтелект допомагає знаходити логічні зв'язки там, де людина через великий обсяг інформації їх просто не помітить. Наприклад, спеціальні алгоритми можуть автоматично зіставити

геолокацію (GPS-координати), знайдену в метаданих кількох фотографій, із часом публікації коротких дописів у соціальних мережах, таких як Twitter. Це дозволяє системі створити точну карту переміщень людини, дізнатися, де вона живе, працює та з якими людьми зустрічається. Наприклад, алгоритми просторової кластеризації, такі як DBSCAN, здатні згрупувати навіть розрізнені координати з метаданих та створити детальну траєкторію щоденного руху людини. Навіть якщо деякі соціальні мережі (наприклад, Telegram чи Facebook) намагаються захистити користувачів і видаляють EXIF-дані під час завантаження нових фотографій, алгоритми можуть аналізувати старі публікації або файли, надіслані через інші, менш захищені платформи, щоб відновити загальну картину. Для того щоб наочно побачити всі ці зв'язки, фахівці використовують інтелектуальні платформи, такі як Maltego. Вона будує зручні графічні схеми (мережі зв'язків)[5], які показують, як пов'язані між собою різні люди, приховані автори документів та пристрої. Завдяки інтеграції з понад 120 джерелами даних, такі платформи допомагають слідчим миттєво побачити цілісну структуру зв'язків у великих організаціях. Таким чином, поєднання автоматизованих скриптів та штучного інтелекту робить збір метаданих дуже швидким і небезпечним процесом. Це доводить, що сьогодні загрозу для приватності становить не лише основний контент файлів, а й ті невидимі сліди, які генеруються кінцевим обладнанням автоматично.

Вектори витоку інформації та профілювання.

Логічним наслідком автоматизованого профілювання є перехід від простої агрегації даних до їх практичної експлуатації. Зібрані цифрові сліди є лише сировиною, проте їх структурування та використання для реалізації конкретних векторів атак становить справжню загрозу для конфіденційності та корпоративної безпеки. У цьому розділі детально розглядається, як саме непрямі цифрові сліди перетворюються на потужний інструмент у руках зловмисників: від розкриття особистості користувачів до підготовки цілеспрямованих атак на організації.

У сучасному цифровому середовищі абсолютна анонімність є здебільшого ілюзією, оскільки деанонімізація — процес зворотного зв'язування псевдонімного профілю в інтернеті з реальною фізичною особою — стає все більш автоматизованою. Цей процес рідко базується на одному прямому витоку; найчастіше він спирається на глибинний кореляційний аналіз непрямих метаданих, зібраних із різних джерел. Кожен користувач залишає за собою унікальний цифровий відбиток, який може складатися з версії операційної системи, параметрів браузера, геолокації та інтенсивності генерації мережевого трафіку. Порівнюючи ці статистичні характеристики з відомими профілями, алгоритми здатні з високою точністю ідентифікувати особу. Окрім традиційного мережевого трафіку, серйозну загрозу становлять метадані сучасних децентралізованих систем та піринговых мереж. Дослідження показують, що навіть метадані публічних торрент-трекерів можуть розкрити поведінкові патерни[7], які вказують на конкретну особу чи її інтереси. Аналогічна ситуація спостерігається і в сфері криптовалют. Популярні браузерні гаманці (наприклад, для доступу до Web3) часто генерують фоновий трафік та виставляють інтерфейси провайдера[6], що призводить до витоку структурних зв'язків між псевдонімними адресами. Це дозволяє стороннім трекерам пов'язувати історію перегляду вебсторінок користувача з його фінансовим станом, руйнуючи будь-які спроби зберегти анонімність в інтернеті.

Метадані публічних документів є одним із найбільш вразливих місць у системі захисту будь-якої організації. Компанії щодня публікують безліч електронних файлів у форматах PDF, Word чи Excel, часто не усвідомлюючи, що ці документи містять приховану інформацію про їхню внутрішню інфраструктуру. Цей етап збору інформації з відкритих джерел називається «футпринтингом» (footprinting) і дозволяє атакувальникам детально дослідити організацію ще до початку активної фази кібератаки. Аналіз корпоративних метаданих зазвичай спрямований на вирішення двох ключових завдань шпигунства. По-перше, це визначення внутрішньої структури компанії. Використовуючи спеціалізовані автоматизовані утиліти, такі як FOCA або Metagoofil, зловмисники масово витягують із завантажених документів імена авторів, їхні корпоративні електронні адреси та історію спільних редагувань. Завдяки цій інформації можна скласти точну схему підрозділів компанії, зрозуміти ієрархію підлеглості та визначити ключових співробітників, на яких найвигідніше спрямувати подальші атаки. По-друге, метадані дозволяють безконтактно виявити використовуване програмне забезпечення та його вдосконалення. У прихованих властивостях документів часто зберігаються абсолютні шляхи до файлів на локальних комп'ютерах користувачів (наприклад, структура папок на диску C:), назви операційних систем, моделі мережевих принтерів та точні версії програм, у яких ці файли створювалися. Якщо аналіз PDF-файлу показує, що документ був згенерований застарілою версією редактора з відомими вразливостями, зловмисник отримує готовий вектор для атаки. Це дозволяє підготувати специфічний експлойт, уникаючи гучного сканування внутрішньої мережі, яке могло б привернути увагу систем безпеки.

Якщо у корпоративному секторі витік метаданих переважно загрожує комерційній таємниці, то для військових та урядових структур він має прямий вплив на фізичну безпеку та життя персоналу. Основну

небезпеку тут становить порушення базових правил операційної безпеки (OPSEC) під час використання цифрових пристроїв. Найбільш критичними є геолокаційні координати (GPS) та мітки часу, які автоматично генеруються розумними гаджетами або вбудовуються у формат EXIF під час створення цифрових фотографій. Класичним прикладом такого витоку став глобальний інцидент із застосунком для фітнес-трекерів Strava. Компанія-розробник опублікувала загальнодоступну теплову карту (heatmap), що відображала сукупну активність усіх користувачів платформи. Оскільки військовослужбовці у різних куточках світу не вимикали свої трекери під час ранкових пробіжок чи патрулювань, ця непряма інформація де-факто розсекретила точне розташування прихованих військових баз, внутрішню структуру режимних об'єктів та маршрути постачання. Крім того, публікація фотографій із зон виконання завдань без попереднього очищення EXIF-даних дозволяє противнику за лічені секунди визначити координати місця зйомки. Аналізуючи часові мітки у поєднанні з геолокацією, супротивник може відстежувати маршрути переміщення важливих осіб, вивчати режим роботи урядових об'єктів та планувати диверсійні атаки.

Найбільш розповсюдженим і руйнівним наслідком витоку метаданих є їхнє використання у методах соціальної інженерії, зокрема для організації цілеспрямованого фішингу (Spear Phishing). Згідно з галузевою статистикою, понад 90% успішних цілеспрямованих кібератак на організації розпочинаються саме з такого шахрайського електронного листа. На відміну від звичайного масового спаму, ефективність Spear Phishing полягає у глибокій персоналізації повідомлення, яка стає можливою виключно завдяки попередньо зібраним даним розвідки. Використовуючи метадані (наприклад, імена колег, точні назви внутрішніх проектів, специфічне програмне забезпечення), зловмисник створює електронний лист, який виглядає абсолютно легітимним та викликає довіру. Якщо співробітник отримує повідомлення нібито від свого керівника з проханням терміново переглянути документ, у назві якого фігурує реальний проект компанії, ймовірність того, що він відкриє шкідливий файл, критично зростає. Сьогодні цей вектор загрози переживає стрімку еволюцію через інтеграцію великих мовних моделей (LLM) та генеративного штучного інтелекту[8]. Автоматизація дозволяє зловмисникам значно знизити вартість підготовки атаки (менше 5 доларів на одну ціль), одночасно масштабуючи її та підвищуючи переконливість. Штучний інтелект здатний миттєво обробляти метадані, вивчати стилістику спілкування жертви та формувати бездоганні тексти, що адаптуються під конкретного отримувача. Експериментальні дослідження підтверджують, що згенеровані штучним інтелектом фішингові листи на основі зібраних метаданих є надзвичайно ефективними: їхній показник успішних кліків (click-through rate) досягає 54%, тоді як традиційні, не персоналізовані атаки демонструють результативність лише на рівні близько 12%[9]. Понад 83% сучасних фішингових розсилань тією чи іншою мірою вже використовують згенерований контент, який легко уникає виявлення традиційними системами безпеки, що базуються на статичних сигнатурах.

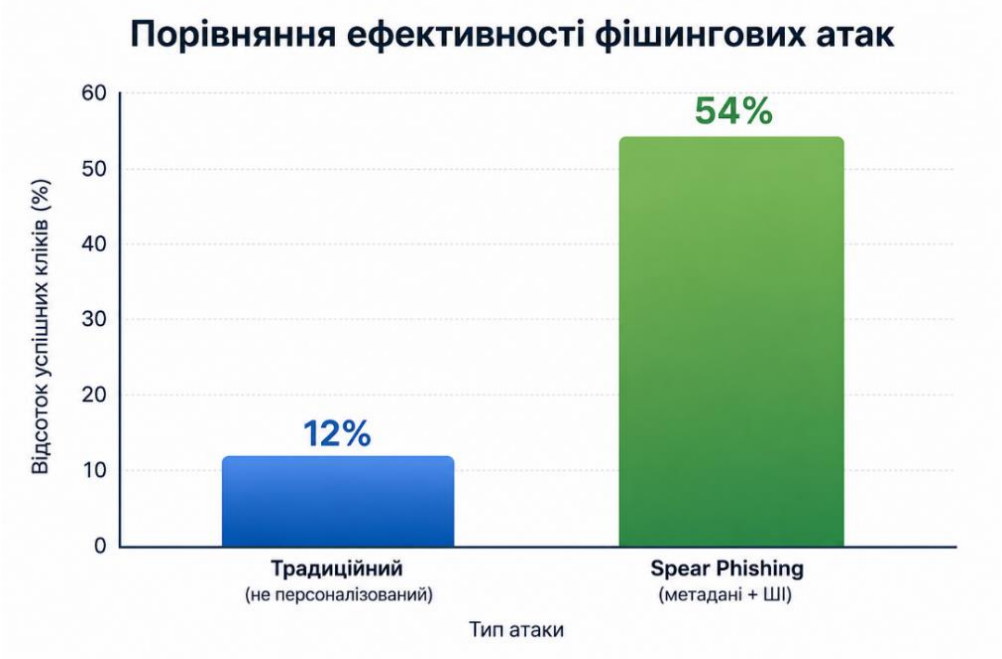

Рис. 1. Порівняння результативності (CTR) традиційних фішингових атак та цілеспрямованого Spear Phishing[9]

Таблиця 4

Зіставлення виявлених метаданих із векторами кібератак та їхніми наслідками

Тип виявлених метаданих	Основний вектор атаки	Наслідки для безпеки об'єкта
Мережеві маркери, історія гаманців Web3	Кореляційний аналіз, Деанонімізація	Розкриття реальної особи анонімного користувача, прив'язка фінансових активів до історії перегляду.
Геолокація, часові мітки (EXIF, трекери)	Фізична розвідка, Порушення OPSEC	Розкриття координат військових баз, маршрутів патрулювання, режиму дня керівництва.
Версії ПЗ, абсолютні шляхи збереження	Експлуатація вразливостей (Zero-day)	Використання специфічних експлойтів для виявленого застарілого корпоративного ПЗ без сканування мережі.
Імена авторів, структура каталогів	Spear Phishing, Соціальна інженерія	Маніпуляція довірою співробітників, підвищення ймовірності кліку (до 54%), крадіжка облікових даних

Підсумовуючи, можна стверджувати, що недбале ставлення до метаданих перетворює їх на потужну зброю, яка стирає межу між віртуальною загрозою та реальними наслідками. Інформація, яка генерується фоново і залишається поза увагою користувача, стає надійним фундаментом для підготовки та реалізації складних ланцюгів атак.

АНАЛІЗ РЕАЛЬНИХ ІНЦИДЕНТІВ (КЕЙС-СТАДІ)

Теоретичні моделі збору та агрегації метаданих знаходять своє найбільш руйнівне підтвердження під час аналізу реальних інцидентів. Сучасна практика кібербезпеки доводить: критичні витоки інформації найчастіше відбуваються не через злам складних криптографічних систем, а через ігнорування користувачами побічних цифрових артефактів. Аналіз наведених нижче кейсів демонструє, як метадані здатні розсекречувати фізичні об'єкти, компрометувати політичні процеси та деанонімізувати джерела, перетворюючись на зброю реального впливу.

Глобальний витік через Strava Heatmap (2018)

Масове використання IoT-пристроїв (фітнес-трекерів, смарт-годинників) генерує безперервний потік геопросторових метаданих, які автоматично синхронізуються з хмарними платформами. Агрегація цих даних створює стратегічні загрози на макро- та мікрорівнях. У 2018 році розробник фітнес-застосунку Strava опублікував глобальну "теплову карту", згенеровану на основі мільярда активностей користувачів. Оскільки політика публічності працювала "за замовчуванням", дослідники змогли виявити на карті локації, де активність місцевого населення була нульовою, але іноземні військові регулярно використовували трекери. Аналіз цього масиву метаданих дозволив із високою точністю ідентифікувати секретні американські бази в Сирії, Афганістані (провінція Гільменд), маршрути патрулів у Сомалі, а також внутрішнє планування Зони 51 у США[10].

Кінетичний вплив: Справа Ржицького (2023)

Найбільш загрозливий ступінь небезпеки виникає, коли метадані стають інструментом фізичного усунення. У липні 2023 року в Краснодарі був застрелений колишній командир російського підводного човна Станіслав Ржицький[11]. Зловмисникам не потрібно було застосовувати складне зовнішнє спостереження — вони здійснили аналіз "патернів життя" (Pattern of Life Analysis) на основі його відкритого профілю у Strava. Агрегація історичних GPS-треків та часових міток (timestamps) дозволила алгоритмічно вирахувати, що жертва щодня бігає одним і тим самим маршрутом, починаючи о шостій ранку. Це дозволило виконавцю замаху обрати ідеальну "сліпу зону" без камер нагляду, перетворивши живу людину на передбачувану мішень.

Геополітична криза: Звіт Мехліса (2005)

Другий критичний вектор лежить у площині документообігу. Формати Microsoft Office та PDF створюють у користувачів небезпечну ілюзію: якщо текст видалено з екрана, він зникає назавжди. Проте їхня архітектура базується на інкрементальному збереженні, залишаючи повну хронологію мислення автора. У 2005 році комісія ООН під керівництвом Детлева Мехліса опублікувала звіт щодо розслідування вбивства прем'єр-міністра Лівану Рафіка Харірі. З дипломатичних міркувань перед публікацією зі звіту видалили імена сирійських посадовців. Однак документ був завантажений на сайт ООН у форматі Word із

невимкненою функцією відстеження змін. Аналітики легко відновили метадані історії правок, що розкрили імена брата президента Сирії та керівників спецслужб[12], ледь не спровокувавши міжнародний конфлікт.

Тіньові атаки (Shadow Attacks) у PDF

Сучасні дослідження виявили фундаментальну вразливість самої специфікації формату PDF (CVE-2022-25641)[14], яка дозволяє експлуатувати метадані навіть у криптографічно підписаних документах. Цей клас загроз отримав назву "Тіньові атаки"[13]. Зловмисник створює документ із двома шарами: видимим (легітимним) та прихованим (тіньовим). Коли жертва накладає цифровий підпис, він фіксує весь масив байтів. Згодом зловмисник маніпулює таблицями перехресних посилань (cross-reference tables), змушуючи програму ігнорувати початковий шар і відображати тіньовий контент (наприклад, інші умови договору чи реквізити). При цьому цифровий підпис залишається абсолютно дійсним, оскільки сам зміст підписаної частини не змінювався. Це доводить, що візуальний шар документа та його бінарна структура — різні сутності.

Падіння вбивці BTK (2005)

Спроби приховати свою ідентичність за допомогою мережевої анонімності (VPN, Tor) часто розбиваються об локальні метадані, прикріплені до файлів, або апаратні закладки в периферійних пристроях. Американський серійний вбивця Денніс Рейдер (BTK) понад 30 років уникав правосуддя, дотримуючись суворої фізичної конспірації. Однак його критичною помилкою стало нехтування базовою цифровою гігієною: у 2005 році він надіслав на телебачення дискету з листом. Комп'ютерний криміналіст виявив на носії магнітні залишки видаленого документа. Аналіз внутрішніх адміністративних властивостей (docProps) цього файлу показав, що в полі «Остання зміна» стояло ім'я "Dennis", а в полі «Організація» — "Christ Lutheran Church". Звичайний пошук в інтернеті миттєво вказав на Рейдера[15], президента ради цієї церкви, що призвело до його арешту.

Фізична стеганографія: Ріаліті Віннер (2017)

Метадані існують не лише у цифровому середовищі. У 2017 році 25-річна контрактниця АНБ США Ріаліті Віннер роздрукувала секретний звіт щодо втручання у вибори та надіслала його журналістам. Видання опублікувало PDF-скан оригінальної роздруківки. Спецслужби ідентифікували Віннер за лічені дні, використавши Код ідентифікації машини (MIC) — "жовті крапки відстеження"[17], які сучасні кольорові лазерні принтери непомітно наносять на кожен аркуш. Цей мікроскопічний матричний код містив точний серійний номер принтера, а також дату, годину та хвилину друку. Зіставивши ці метадані з внутрішніми логами сервера, слідчі звузили коло підозрюваних з усього відомства до однієї особи.

Наведені кейси дозволяють сформувати кілька критичних інсайтів:

1. Асиметрія загроз: Якщо раніше для пошуку військових об'єктів були потрібні багатомільйонні ресурси, сьогодні це досягається агрегацією дешевих метаданих з IoT-пристроїв.

2. Фонова генерація: Системи (від офісного ПЗ до принтерів) проектуються так, щоб генерувати надлишкові дані непомітно для людини. Користувачі стають жертвами ілюзії безпеки, спираючись лише на те, що бачать на екрані чи папері.

3. Кінетичний вектор: Метадані еволюціонували від інструменту цифрового шпигунства до засобу фізичного усунення об'єктів.

З огляду на ці тенденції, захист інформації не може обмежуватися лише антивірусами. Необхідне впровадження комплексних інструментів та організаційних політик (Scrubbing та OPSEC), які детально розглядаються у наступному розділі.

Методи протидії та мінімізації ризиків

З огляду на масштабні інциденти та критичні вразливості, висвітлені у попередніх розділах, стає очевидним, що традиційні парадигми інформаційної безпеки, орієнтовані переважно на криптографічний захист корисного навантаження (основного контенту) та антивірусне сканування, є концептуально недостатніми. Метадані функціонують як цифровий еквівалент конверта: навіть якщо сам лист надійно зашифрований, інформація про відправника, отримувача, маршрут, час та інструменти створення залишається відкритою для автоматизованого збору та аналізу. Ефективна протидія цій загрозі вимагає впровадження багатоешелонованої оборонної стратегії, яка охоплює технічні інструменти очищення даних, інтеграцію складних мережевих систем фільтрації, суворі організаційні політики та фундаментальні зміни на етапі проєктування програмного забезпечення.

На технічному рівні базовим інструментом захисту конфіденційності є процес автоматизованого очищення метаданих (Data Scrubbing). Цей підхід передбачає використання спеціалізованих програмних засобів для систематичного виявлення та видалення прихованих властивостей файлу, історії правок, геолокаційних маркерів та інших цифрових артефактів перед передачею даних у публічний або неконтрольований простір. Для реалізації цього завдання на рівні індивідуальних користувачів та

дослідників розроблено низку утиліт, кожна з яких має свою специфіку та сферу застосування. Одним із найбільш потужних рішень є MAT2 (Metadata Anonymisation Toolkit) — інструмент із відкритим вихідним кодом[17], написаний на мові Python 3. Архітектура MAT2 дозволяє здійснювати глибоке криміналістичне очищення широкого спектра форматів, включаючи зображення, аудіо (через бібліотеку mutagen), PDF-документи (за допомогою cairo та poppler) та формати Microsoft Office. Важливою особливістю MAT2 є те, що утиліта не просто видаляє виявлені теги, а створює повністю очищену репліку файлу, застосовуючи техніки примусового перестиснення зображень або растеризації тексту в PDF-документах для гарантованого знищення прихованих шарів та водяних знаків. Для роботи виключно з медіафайлами галузевим стандартом залишається утиліта ExifTool, яка підтримує понад 200 форматів і дозволяє масово та рекурсивно очищати цілі директорії від геолокаційних координат (GPS), серійних номерів камер та міток часу за допомогою команд автоматизованого парсингу.

Водночас усе більшої популярності набувають локальні вебдодатки (наприклад, PrivMeta), які виконують очищення структурних метаданих (таких як core.xml та app.xml у документах формату .docx) безпосередньо у браузері на пристрої користувача. Такий підхід виключає необхідність передачі чутливих файлів на зовнішні сервери, що нівелює ризики перехоплення інформації під час самого процесу деанонімізації.

Проте використання ізольованих утиліт є недостатнім для забезпечення корпоративної безпеки через вплив людського фактора та неможливість масштабування. У великих організаціях процеси очищення метаданих інтегруються безпосередньо в мережеву інфраструктуру за допомогою систем запобігання втраті даних (Data Loss Prevention — DLP). Сучасні DLP-системи (наприклад, від розробників Clearswift, Forcepoint або Symantec) функціонують на базі протоколу ICAP (Internet Content Adaptation Protocol), що дозволяє їм перехоплювати та модифікувати вебтрафік і поштові комунікації в режимі реального часу. Завдяки технології адаптивного редагування (Adaptive Redaction)[18], замість повного блокування документа через порушення політики безпеки, DLP-шлюз динамічно здійснює структурну санацію: вилучає приховані коментарі, видаляє таблиці перехресних посилань або редагує конфіденційні метадані авторів, дозволяючи легітимному контенту безперешкодно досягти адресата. Цей механізм забезпечує баланс між суворими вимогами безпеки та безперервністю бізнес-процесів організації.

Найбільш досконалою технічною еволюцією у сфері захисту від витоків через прихований контент є технологія деконструкції та реконструкції контенту (Content Disarm and Reconstruction — CDR)[19]. На відміну від антивірусних систем або пісочниць (Sandboxing), які намагаються виявити шкідливий код за допомогою сигнатур чи поведінкового аналізу, CDR базується на фундаментальній парадигмі нульової довіри (Zero-Trust) до будь-якого вхідного або вихідного файлу. Увесь процес обробки документа технологією CDR відбувається за кілька мілісекунд і складається з суворої послідовності кроків. Спочатку система ідентифікує реальний формат файлу, ігноруючи його розширення, що дозволяє запобігти атакам із маскуванням (file masquerading). Далі файл декомпонується на базові елементи: наприклад, документ Word розпаковується як ZIP-архів, після чого виокремлюються текстові XML-структури, вбудовані OLE-об'єкти, макроси VBA та блоки метаданих. На етапі оцінки політик усі неавторизовані елементи, включаючи скрипти, нетипові метадані та зовнішні посилання, безповоротно знищуються. На фінальному етапі система генерує абсолютно новий, «математично чистий» файл із використанням легітимних шаблонів, візуально ідентичний оригіналу, але позбавлений будь-яких прихованих векторів атак.

Таблиця 5

Порівняльний аналіз технічних методів санації інформаційних об'єктів

Характеристика	Базове очищення (Data Scrubbing)	Адаптивне редагування (DLP / ICAP)	Content Disarm and Reconstruction (CDR)
Архітектурний підхід	Пошук та видалення конкретних тегів метаданих за запитом користувача.	Мережевий моніторинг трафіку та динамічне вилучення даних згідно з налаштованими політиками.	Тотальна деконструкція файлу, знищення всього неавторизованого контенту та генерація нового об'єкта.
Стійкість до Zero-Day атак	Низька. Залежить від здатності інструменту розпізнавати конкретні формати та словники метаданих.	Середня. Ефективність обмежується якістю регулярних виразів та словників у системі DLP.	Дуже висока. Система апріорі відкидає будь-який контент, що не відповідає еталонній специфікації формату.

Вплив на функціональність файлу	Мінімальний. Іноді призводить до пошкодження складних цифрових підписів.	Помірний. Документ зберігає оригінальну структуру, втрачаючи лише заблоковані елементи.	Високий. Усі активні елементи (макроси, складні скрипти, вбудовані об'єкти) видаляються безповоротно.
Основна сфера застосування	Індивідуальна кібергігієна, журналістські розслідування, підготовка файлів до публікації.	Корпоративні поштові сервери (MTA), вебпроксі, файлові сховища.	Кросс-доменні шлюзи, урядові мережі, фінансовий сектор, захист від цілеспрямованих атак (APT).

Незважаючи на високу ефективність, технологія CDR не позбавлена недоліків. Основний виклик полягає у збереженні функціональності специфічних бізнес-документів (наприклад, фінансових моделей у Excel), які критично залежать від легітимних макросів. У таких випадках жорстка конфігурація CDR може порушити бізнес-процеси. Крім того, дослідники виявляють складні крайові випадки (edge cases), пов'язані з файлами-поліглотами. Наприклад, якщо до виконуваного файлу (PE) з чинним цифровим підписом Authenticode приєднано шкідливий ZIP-архів, наївне видалення архіву призведе до зміни контрольної суми та анулювання підпису, що зруйнує систему довіри операційної системи. Тому сучасні алгоритми реконструкції повинні бути обізнаними про контрольні суми (checksum-aware) і структурно зберігати блоки підписів під час санації метаданих.

Будь-які технічні рішення є лише інструментами, які не здатні забезпечити комплексний захист без належної методологічної та організаційної бази. Впровадження суворих правил операційної безпеки (OPSEC) є критичним компонентом захисту від витоків метаданих. Основоположним документом, що регулює цей процес на державному та корпоративному рівнях, є стандарт Національного інституту стандартів і технологій США NIST SP 800-188 («De-Identifying Government Datasets: Techniques and Governance»)[20]. Цей стандарт формує комплексну рамку для деідентифікації масивів даних, з метою мінімізації інформаційних ризиків для осіб та організацій під час публікації або обміну цифровими об'єктами.

Методологія NIST вимагає відмовитися від спрощеного підходу, який обмежується лише видаленням прямих ідентифікаторів (імен, номерів телефонів). Натомість організаціям приписується здійснювати глибоку трансформацію квазі-ідентифікаторів — непрямих даних, таких як геолокація, часові мітки або структурні характеристики пристроїв, сукупність яких дозволяє провести кореляційний аналіз і деанонімізувати особу. Стандарт визначає необхідність досягнення математично обґрунтованої приватності, зокрема використання моделей k-анонімності та алгоритмів диференціальної приватності. Згідно з цими концепціями, організація повинна гарантувати, що будь-який суб'єкт у наборі даних не може бути відрізнений від щонайменше k-1 інших суб'єктів, а додавання штучного шуму (пертурбація даних) унеможливлює точну ідентифікацію без втрати загальної статистичної цінності документа.

Для забезпечення безперервного контролю над виконанням цих вимог, стандарти безпеки (зокрема концепція NIST Privacy Framework) передбачають імплементацію таких організаційних механізмів[21]:

1)**Створення Рад з перегляду розкриття інформації (Disclosure Review Boards - DRB):** Спеціалізовані внутрішні комітети, які несуть відповідальність за оцінку ризиків деідентифікації перед будь-якою публікацією даних або передачею їх стороннім підрядникам. DRB повинні критично оцінювати стандартні налаштування програмного забезпечення, які часто автоматично вбудовують ідентифікуючу інформацію у файли або їхні метадані.

2)**Тестування "Мотивованим зловмисником" (Motivated Intruder Test):** Організаційні процедури вимагають регулярного проведення симуляцій, під час яких компетентні зовнішні спеціалісти (пентестери) намагаються реідентифікувати знеособлені дані або відновити приховані метадані з опублікованих документів, перевіряючи таким чином надійність застосованих методів санації.

3)**Управління життєвим циклом даних та інвентаризація:** Відповідно до вимог NIST SP 800-53, організації повинні забезпечити суворий контроль потоків інформації. Політики мають класифікувати дані на публічні, ті, що мають обмежений доступ, та ті, що заборонені до передачі (no access data), з обов'язковим застосуванням процедур медіасанації (NIST SP 800-88) після завершення використання файлів.

Не менш важливою складовою OPSEC є навчання персоналу основам кібергігієни. Співробітники повинні усвідомлювати, що вектором витоку інформації може стати навіть буденна комунікація. Яскравим

прикладом є еволюція фішингових атак із використанням метаданих календарних файлів (формат .ics). Зловмисники масово експлуатують довіру користувачів до розкладів, вбудовуючи шкідливі посилання та base64-закодовані інструкції у поля метаданих календаря (наприклад, у поле ATTACH). Навчання повинно передбачати інструктаж щодо заборони автоматичного додавання зовнішніх подій до календарів та критичного ставлення до будь-яких системних файлів, які раніше вважалися цілком безпечними. Забезпечення багатофакторної автентифікації та систем умовного доступу додає ще один рівень захисту у випадку, якщо співробітник все ж таки піддасться методам соціальної інженерії на основі скомпрометованих метаданих.

Таблиця 6

Організаційні рівні управління метаданими згідно з NIST Framework

Рівень управління	Практична реалізація в організації	Відповідність вимогам
Інвентаризація та мапування (ID.IM-P)	Точне визначення систем, які генерують та обробляють метадані. Облік власників та операторів даних.	Здатність відстежити походження витоку та визначити відповідальність підрядників.
Управління даними (CT.DM-P)	Видалення даних за політикою. Використання стандартизованих форматів для передачі. Управління метаданими з дозволами на обробку.	Мінімізація надлишкового накопичення цифрових артефактів (Data Minimization).
Управління життєвим циклом (CT.PO-P)	Інтеграція процесів деідентифікації в цикл розробки систем. Розділення коду аналізу від процесу деідентифікації.	Формування відтворюваних робочих процесів. Мінімізація прямої шкоди від розкриття.

Останньою, але фундаментально найважливішою ланкою у стратегії мінімізації ризиків є зміна філософії проєктування програмного забезпечення. Історично розробники зосереджувалися на функціональності, розглядаючи метадані як корисний інструмент для індексації та інтероперабельності систем. Проте жорстка регуляторна політика, зокрема впровадження GDPR у Європі та CCPA у США, кардинально змінила ринковий ландшафт. Усвідомлення того, що середній штраф за порушення вимог мінімізації даних перевищує 4,2 мільйона євро, змусило технологічні компанії інвестувати значні ресурси у створення рішень для автоматичної санації контенту. За прогнозами аналітиків, глобальний ринок інструментів для видалення метаданих, який у 2025 році оцінювався у 3,8 мільярда доларів, зросте до 9,1 мільярда доларів до 2034 року[22].

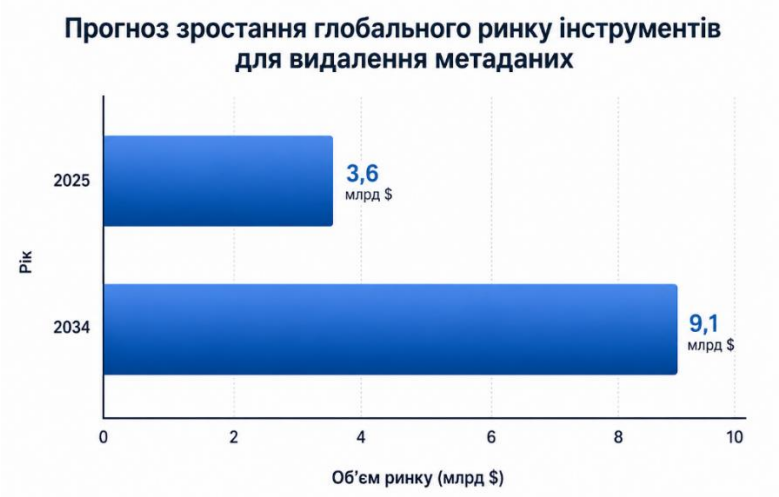

Рис. 2. Прогноз зростання об'єму глобального ринку інструментів для видалення метаданих на період 2025–2034 рр[22]

Це зростання відображає перехід індустрії до концепції "Privacy by Design" (приватність за замовчуванням). Ця парадигма вимагає, щоб захист конфіденційності та очищення від цифрових слідів були не додатковими (opt-in) опціями, які користувач повинен шукати в налаштуваннях, а базовою архітектурною властивістю будь-якої системи. Практична реалізація цього підходу вже спостерігається у масовому сегменті: провідні платформи комунікації (наприклад, Telegram, Signal) та соціальні мережі (Facebook) впровадили алгоритми, які автоматично і безповоротно знищують EXIF-дані (точні

координати, моделі пристроїв) під час завантаження медіафайлів на їхні сервери. Такий підхід захищає мільйони необізнаних користувачів від масового парсингу та OSINT-профілювання. На корпоративному рівні розробники рішень для документообігу (наприклад, Adobe та Microsoft) також розширюють вбудований функціонал інспекторів документів (Document Inspector), дозволяючи адміністраторам примусово застосовувати політики очищення перед кожним збереженням або надсиланням файлу.

Новим, ще не до кінця вивченим викликом для парадигми "Privacy by Design" є стрімке впровадження великих мовних моделей (LLM) та архітектур Retrieval-Augmented Generation (RAG). Системи RAG, які векторизують корпоративні документи для розширення знань штучного інтелекту, створюють нові поверхні для атак. Якщо вихідні документи не проходять глибоку санацію, векторні бази даних поглинають і назавжди вбудовують персональну інформацію (PII) та структурні метадані у свої моделі. Це призводить до ризику інверсійних атак (embedding inversion attacks), коли зловмисник через специфічні запити може змусити модель розкрити приховані дані про авторів або історію проєкту. Для протидії цьому, передові рішення вимагають попередньої обробки документів із використанням гібридних алгоритмів (регулярних виразів та обробки природної мови) для знеособлення як тексту, так і його метаданих ще до моменту векторизації. Штучний інтелект стає інструментом подвійного призначення: він одночасно використовується зловмисниками для масової кореляції даних і генерації персоналізованого фішингу, та інтегрується розробниками в шлюзи безпеки для предиктивного виявлення аномалій у метаданих і автоматичної санації складних структур. Відповідно, стійкість майбутніх інформаційних систем залежатиме від того, наскільки глибоко принципи диференціальної приватності та автоматичної санації будуть інтегровані в базовий код штучного інтелекту та комунікаційних платформ.

ЕКСПЕРИМЕНТАЛЬНА ОЦІНКА РІВНЯ РОЗКРИТТЯ МЕТАДАНИХ ТА АЛГОРИТМІЗАЦІЯ ПРОЦЕСУ САНАЦІЇ

Для практичного підтвердження теоретичної бази щодо неконтрольованого витоку інформації та обґрунтування необхідності багаторівневих оборонних стратегій, було проведено експериментальне дослідження. Метою експерименту стала розробка та апробація автоматизованої моделі аналізу метаданих, що дозволяє не лише масово виявляти цифрові сліди, але й класифікувати їх за рівнем інформаційного ризику для подальшої алгоритмічної санації контенту.

Математична формалізація моделі ризику (Metadata Information Risk Score)

З метою автоматизації процесу ухвалення рішень щодо очищення даних, розроблено кількісний показник — Metadata Information Risk Score (MIRS). Він математично визначає сумарний рівень загрози, який генерує конкретний цифровий об'єкт на основі знайдених у ньому транзитних, описових та структурних артефактів. Загальна формула розрахунку ризику для окремого файлу має такий вигляд:

$$MIRS = (\sum_{i=1}^{n}(W_i \times S_i) \times K_{corr} \quad (1)$$

де:

n — загальна кількість екстрагованих метаданих у досліджуваному файлі.

W_i — ваговий коефіцієнт класу метаданих. Базуючись на стандартах інформаційної безпеки, транзитним та структурним артефактам, які дозволяють підготовку експлойтів, присвоюється найвища вага (W = 3), описовим даним (імена авторів) — середня (W = 2), адміністративним (дати модифікації) — базова (W = 1).

S_i — індекс чутливості конкретного атрибута (за шкалою від 1 до 10). Наявність точних GPS-координат оцінюється як критична вразливість (S = 10). Імена авторів та шляхи збереження на локальних дисках, що формують вектор для соціальної інженерії та Spear Phishing, оцінюються у 7-8 балів. Загальна технічна інформація отримує 1-2 бали.

K_{corr} — коефіцієнт кореляційної загрози. Він становить 1.5 у випадках, коли алгоритм виявляє нелінійну синергію тегів. З точки зору теорії ймовірностей, одночасна наявність імені особи та її точного часу активності експоненційно підвищує ймовірність успішної деанонімізації. В інших випадках K_{corr} = 1.0.

На основі розрахованого значення MIRS цифрові об'єкти автоматично класифікуються на чотири рівні ризику:

1. MIRS < 15 (Low): Загроза мінімальна, файл містить лише системний шум. Втручання не потрібне.
2. 15 ≤ MIRS < 40 (Medium): Присутні адміністративні артефакти. Вимагається застосування базового автоматизованого очищення (Data Scrubbing).

3. 40 ≤ MIRS < 70 (High): Виявлено чутливі дані. Загроза соціального профілювання, що вимагає глибокої санації.

4. MIRS ≥ 70 (Critical): Наявність апаратних ідентифікаторів чи GPS. Обов'язкова ізоляція та застосування технології деконструкції контенту (CDR).

Методологія та результати експериментального тестування

З метою забезпечення високої репрезентативності результатів та мінімізації статистичної похибки, експериментальну вибірку було розширено до 100 цифрових об'єктів. Вибірка охоплює 9 найбільш поширених форматів корпоративного, приватного та веб-орієнтованого обміну даними: формати Microsoft Office (DOCX — 7 файлів, XLSX — 8, PPTX — 14), PDF-документи (9 файлів), медіафайли (PNG — 14, JPEG — 15), вебформати (HTML — 9), текстові документи (RTF — 16) та файли табличних баз даних (CSV — 8). Використання формату CSV слугувало контрольною групою (Baseline), оскільки його архітектура технічно не підтримує вбудованих структурних метаданих.

Процес екстракції артефактів здійснювався за допомогою розробленого програмного пайплайну на базі мови Python, який інкапсулює функціонал OSINT-утиліти ExifTool.

Таблиця 7

Зведена статистика виявлення метаданих за форматами файлів (До санації)

Формат файлу	Кількість	Ідентифікація автора (Author/Creator)	Часові мітки (Create Date)	Дані про ПЗ (Software/Producer)
RTF	16	100% (16/16)	0% (0/16)	0% (0/16)
DOCX	7	100% (7/7)	100% (7/7)	0% (0/7)
PPTX	14	100% (14/14)	100% (14/14)	0% (0/14)
XLSX	8	100% (8/8)	100% (8/8)	0% (0/8)
PDF	9	100% (9/9)	100% (9/9)	100% (9/9)
PNG	14	100% (14/14)	0% (0/14)	0% (0/14)
HTML	9	100% (9/9)	0% (0/9)	0% (0/9)
JPEG	15	0% (0/15)	0% (0/15)	0% (0/15)
CSV (Baseline)	8	0% (0/8)	0% (0/8)	0% (0/8)
Загалом	100	77% (77/100)	38% (38/100)	9% (9/100)

Результати автоматизованого аналізу виявили критично високий рівень інтеграції чутливої інформації, згенерованої ПЗ за замовчуванням. У 77% випадків інструмент успішно вилучив імена творців документів або назви облікових записів. 38% файлів розкрили точний час свого створення. У 100% проаналізованих PDF-документів чітко ідентифіковано використану бібліотеку генерації, що відкриває вектор для експлуатації вразливостей.

Експертна оцінка та статистична валідація моделі MIRS

Для перевірки аналітичної здатності запропонованої моделі MIRS було проведено порівняльний аналіз із залученням експертної оцінки (Ground Truth). Експертна оцінка передбачала ручне маркування: 80 об'єктів було визначено як "такі, що потребують санації" (Positive) через їхню здатність приховувати артефакти, тоді як 20 файлів (переважно Baseline CSV) визнано безпечними (Negative).

Автоматизована модель MIRS опрацювала вибірку: файли з MIRS ≥ 15 маркувалися як вразливі, а MIRS < 15 — як безпечні. У результаті було сформовано матрицю помилок (Confusion Matrix):

- True Positives (TP) = 77: Вразливі файли успішно виявлені алгоритмом.
- True Negatives (TN) = 20: Безпечні файли алгоритм коректно проігнорував.
- Помилка 1-го роду (False Positives, FP) = 0: Хибних спрацьовувань не зафіксовано.
- Помилка 2-го роду (False Negatives, FN) = 3: Три файли JPEG експертно розцінювались як паттерн ризику через загальну специфікацію формату, проте алгоритм класифікував їх як Low, оскільки фізично теги в конкретній вибірці були відсутні.

На основі матриці помилок розраховано ключові метрики ефективності класифікатора:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} = \frac{77+20}{100} = 0.97(97\%) \quad (2)$$

$$Precision = \frac{TP}{TP+FP} = \frac{77}{77+0} = 1.0(100\%) \quad (3)$$

$$Recall = \frac{TP}{TP+FN} = \frac{77}{77+3} \approx 0.962(96.2\%) \quad (4)$$

$$F_1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{1.0 \times 0.962}{1.0 + 0.962} \approx 0.98(98\%) \quad (5)$$

Показник F_1-score на рівні 98% доводить математичну точність вагових коефіцієнтів моделі MIRS та підтверджує, що вона здатна ефективно визначати рівень ризику.

Аналіз ефективності автоматизованої санації

Для оцінки стійкості цифрових об'єктів до засобів очищення, уся експериментальна вибірка (100 файлів) пройшла процес автоматизованої базової санації (Data Scrubbing) із застосуванням утиліти глибокого видалення тегів.

Таблиця 8

Зведена статистика виявлення метаданих за форматами файлів (Після санації)

Формат файлу	Кількість	Ідентифікація автора (Author/Creator)	Часові мітки (Create Date)	Дані про ПЗ (Software/Producer)
RTF	16	100% (16/16)	0% (0/16)	0% (0/16)
DOCX	7	100% (7/7)	100% (7/7)	0% (0/7)
PPTX	14	100% (14/14)	100% (14/14)	0% (0/14)
XLSX	8	100% (8/8)	100% (8/8)	0% (0/8)
PDF	9	0% (0/9)	0% (0/9)	0% (0/9)
PNG	14	0% (0/14)	0% (0/14)	0% (0/14)
HTML	9	100% (9/9)	0% (0/9)	0% (0/9)
JPEG	15	0% (0/15)	0% (0/15)	0% (0/15)
CSV (Baseline)	8	0% (0/8)	0% (0/8)	0% (0/8)
Загалом	100	54% (54/100)	29% (29/100)	0% (0/100)

Після завершення очищення було проведено повторний розрахунок MIRS. Експеримент виявив критичну асиметрію в ефективності традиційного очищення:

1. **Повна санація (MIRS = 0):** Базовий скрабінг ідеально спрацював на об'єктах PDF та медіафайлах (PNG, JPEG). Для 100% цих файлів ризик знизився з High до Low (0 балів), підтверджуючи дієвість підходу для ізольованих медіаформатів.

2. **Стійкість до санації (MIRS > 40):** Утиліта базового очищення виявилася нездатною деконструювати глибокі XML-структури офісних пакетів (DOCX, PPTX, XLSX) та файли розмітки (HTML, RTF). Після санації 100% цих форматів (54 файли) продовжували транслювати імена авторів та часові мітки, зберігши рівень ризику Medium/High.

Даний результат є фундаментальним підтвердженням теоретичної гіпотези, висунутої у розділі 5: локальні інструменти очищення є концептуально недостатніми для забезпечення корпоративної безпеки. Експеримент доводить, що для нейтралізації прихованого контенту в складних форматах (Microsoft Office) використання систем базового очищення має бути замінено на технологію тотальної деконструкції та реконструкції контенту (CDR), яка функціонує за парадигмою нульової довіри (Zero-Trust).

Алгоритмізація процесу: технічна реалізація MIRS-пайплайну

Для практичного впровадження моделі MIRS у виробничі інформаційні системи (поштові шлюзи, DLP-рішення, CI/CD-конвеєри розробки) самої математичної формули (1) недостатньо — потрібен чіткий алгоритм послідовної обробки файлу від моменту його надходження до ухвалення автоматизованого рішення про подальшу санацію. Наведена нижче блок-схема формалізує цей процес у вигляді функціонального пайплайну, придатного до програмної реалізації мовою Python з використанням обгортки над ExifTool та бібліотек python-docx, PyPDF2.

ISSN 2663 – 4023

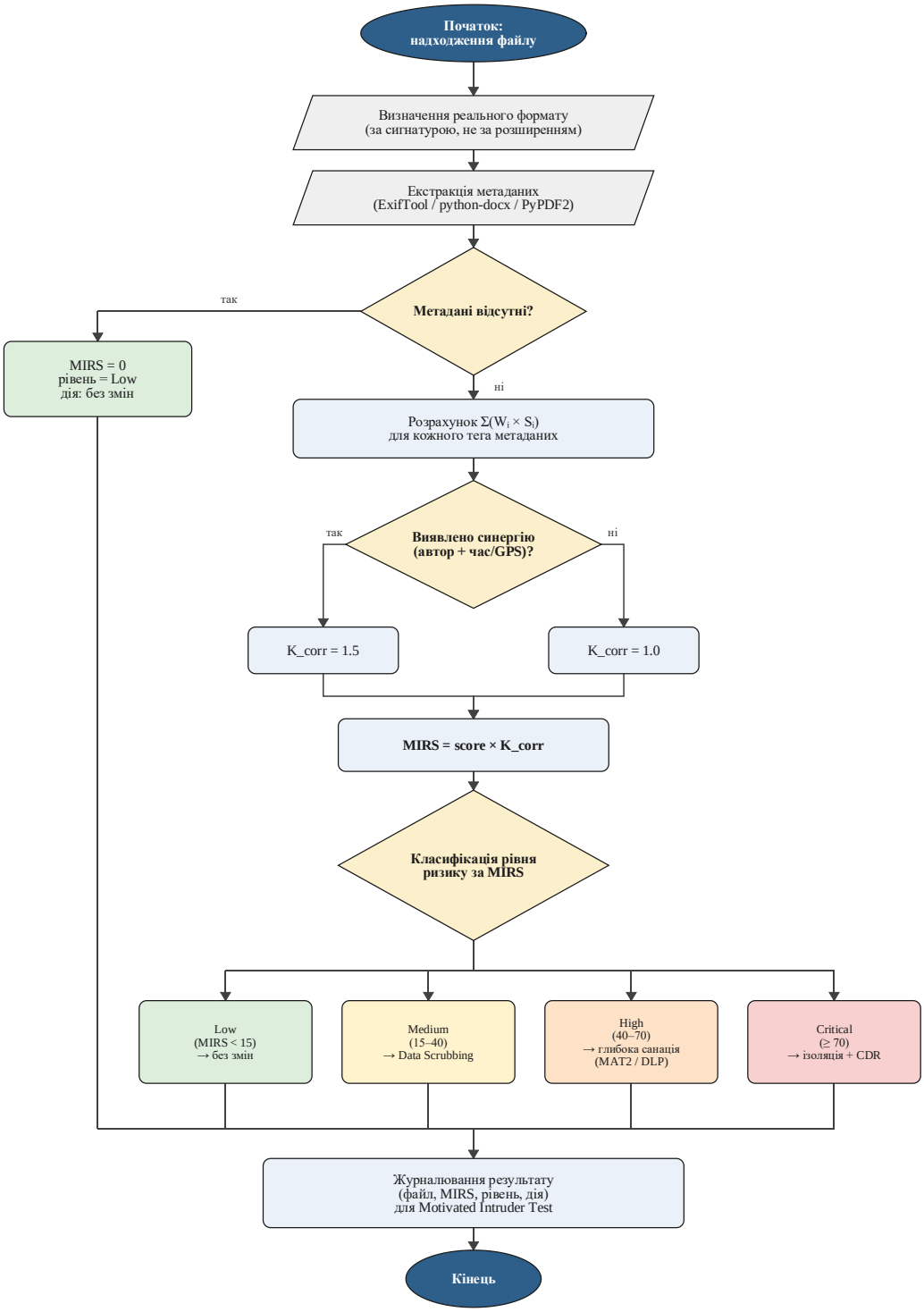

Рис. 3. Блок-схема алгоритму автоматизованої класифікації файлу за моделлю MIRS

Алгоритм має лінійну обчислювальну складність O(n) відносно кількості екстрагованих тегів метаданих n у файлі, що дозволяє обробляти окремий об'єкт за одиниці мілісекунд і масштабувати рішення шляхом паралелізації на рівні пулу воркерів. Ключовою архітектурною перевагою такого підходу є те, що він не замінює описані вище технічні засоби (Data Scrubbing, DLP, CDR), а оркеструє їх: рівень ризику MIRS виступає єдиною точкою прийняття рішення, яка автоматично скеровує файл до відповідного за глибиною засобу санації, уникаючи як надлишкової обробки безпечних об'єктів, так і недостатньої санації критичних.

На практиці такий пайплайн доцільно інтегрувати як асинхронний ICAP-callback у складі поштового шлюзу чи DLP-системи, або як pre-commit/pre-publish хук у корпоративних системах документообігу та CI/CD-конвеєрах, де кожен файл перед публікацією проходить автоматичну перевірку MIRS. Журналювання результатів кожної обробки (файл, значення MIRS, обраний рівень ризику, застосована дія) додатково формує основу для регулярного проведення тестування "Мотивованим зловмисником" (Motivated Intruder Test), описаного в попередньому розділі, оскільки дозволяє ретроспективно оцінити коректність ухвалених автоматизованих рішень.

Обмеження дослідження

Попри отримані результати, дослідження має ряд методологічних обмежень, які варто враховувати при інтерпретації висновків та плануванні подальших робіт.

По-перше, експериментальна вибірка обсягом 100 цифрових об'єктів 9 форматів, хоча й достатня для первинної статистичної валідації моделі MIRS (F1-score = 98%), не є репрезентативною для промислових масштабів корпоративного документообігу, де щодня обробляються тисячі файлів. Узагальнення отриманих метрик точності на ширші, гетерогенніші датасети потребує додаткової валідації.

По-друге, вагові коефіцієнти W_i (для класів метаданих) та коефіцієнти чутливості S_i (для окремих тегів) у формулі MIRS визначені експертно, на основі класифікації NIST SP 800-53, а не шляхом великомасштабного емпіричного калібрування чи навчання на розмічених даних. Це залишає елемент суб'єктивності в моделі та відкриває напрям для подальшої роботи над її автоматизованим налаштуванням (наприклад, методами машинного навчання з підкріпленням на основі зворотного зв'язку від аудиторів безпеки).

По-третє, значення корекційного коефіцієнта синергії K_corr = 1.5 обране як фіксована константа без проведення чутливого аналізу (sensitivity analysis) щодо впливу різних значень цього параметра на якість класифікації та стійкість матриці помилок.

По-четверте, модель протестована на конкретному інструментальному стеку (ExifTool, python-docx, PyPDF2) для 9 поширених форматів; менш розповсюджені або нові формати (HEIC, WebP, файли хмарних колаборaційних платформ на кшталт Google Docs чи Microsoft 365 з власними внутрішніми моделями метаданих) залишилися поза межами експериментальної перевірки.

По-п'яте, ефективність технології CDR обґрунтована в роботі переважно теоретично та опосередковано — через експериментально доведену нездатність традиційного Data Scrubbing очистити глибокі XML-структури. Пряме емпіричне порівняння CDR і Data Scrubbing на тій самій вибірці файлів у межах цього дослідження не проводилося і є перспективним напрямом подальшої роботи.

Нарешті, дослідження зосереджується переважно на статичних файлових метаданих; мережеві та транзитні метадані (зокрема, відбитки браузера) розглянуті лише в оглядовій частині, без власного експериментального тестування, а нормативна база спирається здебільшого на стандарти NIST (США), тоді як юрисдикційні особливості (GDPR, національне законодавство про захист персональних даних) висвітлені менш детально.

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ДОСЛІДЖЕНЬ

Проведене комплексне дослідження доводить, що в умовах глобальної цифровізації традиційна парадигма кібербезпеки, орієнтована виключно на захист видимого контенту, є критично недостатньою та створює небезпечну «ілюзію безпеки». Метадані, які фоново і непомітно генеруються сучасним апаратним та програмним забезпеченням, функціонують як «невидиме цифрове ДНК». Це цифрове ДНК здатне розкрити повний контекст створення та транзиту інформації незалежно від того, чи застосовується криптографічне шифрування до самого файлу.

Дослідження продемонструвало, що автоматизація процесів збору даних за допомогою OSINT-інструментів (таких як ExifTool, Metagoofil чи FOCA) у поєднанні з алгоритмами штучного інтелекту перетворила приховані цифрові артефакти на потужну зброю. Аналіз реальних інцидентів та векторів атак засвідчив, що транзитні, описові та структурні метадані активно експлуатуються для деанонімізації осіб, проведення фізичної розвідки (аж до кінетичного впливу) та підготовки високоточних фішингових атак (Spear Phishing). Інтеграція метаданих із генеративним ШІ підвищує результативність таких атак до 54%.

Експериментальне тестування розширеної вибірки зі 100 файлів (9 різних форматів) математично та практично підтвердило масовість проблеми: 77% документів успішно деанонімізували своїх авторів, а 100% PDF-файлів розкрили структурні дані про використане програмне забезпечення. Розроблена та апробована в ході дослідження математична модель оцінки ризиків Metadata Information Risk Score (MIRS) довела свою високу аналітичну здатність. Статистична валідація моделі через експертну оцінку та матрицю помилок підтвердила її надійність із показником F_1-score на рівні 98%, що дозволяє ефективно автоматизувати процес класифікації файлів за рівнем загрози.

Критичним інсайтом дослідження стало практичне доведення неспроможності базових алгоритмів очищення (Data Scrubbing) протистояти складним файловим архітектурам. Експеримент показав, що тоді як для медіафайлів та PDF показник ризику (MIRS) успішно знижувався до нуля, 100% офісних пакетів (DOCX, PPTX, XLSX) та файлів розмітки зберегли чутливі артефакти у своїх глибоких XML-сховищах.

Підсумовуючи, ефективна протидія неконтрольованому витоку метаданих вимагає відмови від фрагментарних методів захисту на користь багатоешелонованих оборонних стратегій. Для гарантування інформаційної та фізичної безпеки необхідно:

1. Впроваджувати комплексні технологічні рішення: експериментально доведено, що замість базового очищення (Data Scrubbing) корпоративний сектор повинен переходити до технологій радикальної деконструкції та реконструкції контенту (CDR).

2. Застосовувати суворі організаційні політики операційної безпеки (OPSEC) та керуватися галузевими стандартами деідентифікації даних, такими як NIST.

3. Здійснити фундаментальний перехід розробників ПЗ та корпоративних екосистем до концепції «Privacy by Design» (приватність за замовчуванням), де автоматична санація цифрових слідів є базовою архітектурною властивістю систем, а не додатковою опцією.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Joint Task Force. (2020). *Security and privacy controls for information systems and organizations (NIST SP 800-53)*. NIST Technical Series Publications. https://nvlpubs.nist.gov
2. Wikipedia contributors. (n.d.). Exif. In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Exif
3. Eckersley, P. (2010). How unique is whose web browser? The role of demographics in browser fingerprinting among US users. *Privacy Enhancing Technologies Symposium (PETS)*. https://petsymposium.org
4. Wu, Z., et al. (2023). Rendered private: Making GLSL execution uniform to prevent WebGL-based browser fingerprinting. *USENIX Security Symposium*. https://usenix.org
5. Infosec Institute. (2021). *Information gathering using Maltego*. https://infosecinstitute.com
6. Eskandari, S., et al. (2023). The masks we (think we) wear: Privacy threats of browser-extension wallets in the Web3 ecosystem. *Privacy Enhancing Technologies Symposium (PETS)*. https://petsymposium.org
7. Hou, S., et al. (2023). *Breadcrumbs in the digital forest: Tracing criminals through torrent metadata with OSINT*. arXiv. https://arxiv.org
8. Roy, S., et al. (2024). *Context-aware spear phishing: Generative AI-enabled attacks against individuals via public social media data*. arXiv. https://arxiv.org
9. Vectra AI. (2023). *AI phishing: How attackers achieve 54% click rates in 5 minutes*. https://vectra.ai
10. Hern, A. (2018). Fitness tracking app Strava gives away location of secret US army bases. *The Guardian*. https://theguardian.com
11. The Record. (2023). Russian naval officer killed near home may have been tracked on Strava app. *The Record Media*. https://therecord.media
12. Wikipedia contributors. (n.d.). Rafic Hariri. In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Rafic_Hariri
13. Kurz, H., et al. (2021). Shadow attacks: Hiding and replacing content in signed PDFs. *NDSS Symposium*. https://ndss-symposium.org
14. Tenable. (2022). *CVE-2022-25641 vulnerability details*. Tenable Research. https://tenable.com
15. BRG. (2023). *Nervous system: How legal tech helped catch the BTK killer*. Berkeley Research Group. https://thinkbrg.com
16. Make Tech Easier. (2017). *In May 2017, an NSA contractor named Reality Winner mailed a printed classified document*. https://maketecheasier.com
17. Voisin, J. (2023). *MAT2 - Metadata Anonymisation Toolkit* [GitHub repository]. GitHub. https://github.com/jvoisin/mat2
18. MediaCircle. (2022). *Clearswift adaptive redaction*. https://mediacircle.de
19. Red Eagle Tech. (2023). *What is content disarm and reconstruction (CDR)? Complete guide*. https://redeagle.tech
20. Garfinkel, S. (2023). *De-identifying government datasets: Techniques and governance (NIST SP 800-188)*. NIST Technical Series Publications. https://nvlpubs.nist.gov
21. NIST. (2020). *NIST informative references for the Privacy Framework*. https://nist.gov
22. Dataintelo. (2024). *Metadata removal tools market research report 2034*. https://dataintelo.com

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (34), 2026
ISSN 2663 – 4023

Melnychuk Maksym
Student, Department of Information Protection
Lviv Polytechnic National University, Lviv, Ukraine
ORCID: 0009-0009-8606-2029
maksym.melnychuk.kb.2025@lpnu.ua

Opirskyy Ivan
D.Sc., Professor, Head of the Department of Information Protection
Lviv Polytechnic National University, Lviv, Ukraine
ORCID: 0000-0002-8461-8996
ivan.r.opirskyi@lpnu.ua

AUTOMATED METADATA ANALYSIS AS AN INFORMATION LEAKAGE VECTOR

Abstract. This paper investigates a fundamental problem of modern cybersecurity—the automated collection and analysis of metadata, which functions as the "invisible digital DNA" of electronic documents, media files, and network communications. The study challenges the dangerous "illusion of security," whereby users and organizations focus primarily on the cryptographic protection of visible content while overlooking hidden transit, descriptive, and structural artifacts. The paper provides a comprehensive analysis of the specific architectural vulnerabilities of popular formats (Microsoft Office, PDF, and JPEG) and, based on real-world incidents, demonstrates how the integration of OSINT tools with artificial intelligence algorithms transforms these digital traces into instruments for deanonymization, physical reconnaissance, and the preparation of highly targeted spear-phishing attacks. To empirically validate the theoretical framework, the paper presents the results of an extended experimental study involving a sample of 100 digital objects across 9 different formats. The research revealed a critical level of background sensitive-information leakage, with successful author deanonymization achieved in 77% of cases. To algorithmize protection processes, a mathematical model for quantitative threat assessment, the Metadata Information Risk Score (MIRS), was developed and statistically validated, with its accuracy confirmed by an *F1-score* of 98%. The study also demonstrates that conventional basic data-scrubbing techniques are ineffective against deep XML structures in office documents. In conclusion, the paper substantiates the need for multilayered defensive strategies to counter these threats, ranging from comprehensive content decomposition technologies such as Content Disarm and Reconstruction (CDR) and NIST de-identification standards to the fundamental implementation of the "Privacy by Design" concept.

Keywords: metadata, cybersecurity, information leakage, deanonymization, OSINT tools, digital profiling, social engineering, spear phishing, Privacy by Design, Content Disarm and Reconstruction (CDR).

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Joint Task Force. (2020). *Security and privacy controls for information systems and organizations (NIST SP 800-53).* NIST Technical Series Publications. https://nvlpubs.nist.gov
2. Wikipedia contributors. (n.d.). Exif. In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Exif
3. Eckersley, P. (2010). How unique is whose web browser? The role of demographics in browser fingerprinting among US users. *Privacy Enhancing Technologies Symposium (PETS).* https://petsymposium.org
4. Wu, Z., et al. (2023). Rendered private: Making GLSL execution uniform to prevent WebGL-based browser fingerprinting. *USENIX Security Symposium*. https://usenix.org
5. Infosec Institute. (2021). *Information gathering using Maltego*. https://infosecinstitute.com
6. Eskandari, S., et al. (2023). The masks we (think we) wear: Privacy threats of browser-extension wallets in the Web3 ecosystem. *Privacy Enhancing Technologies Symposium (PETS).* https://petsymposium.org
7. Hou, S., et al. (2023). *Breadcrumbs in the digital forest: Tracing criminals through torrent metadata with OSINT*. arXiv. https://arxiv.org
8. Roy, S., et al. (2024). *Context-aware spear phishing: Generative AI-enabled attacks against individuals via public social media data*. arXiv. https://arxiv.org
9. Vectra AI. (2023). *AI phishing: How attackers achieve 54% click rates in 5 minutes*. https://vectra.ai

10. Hern, A. (2018). Fitness tracking app Strava gives away location of secret US army bases. *The Guardian*. https://theguardian.com
11. The Record. (2023). Russian naval officer killed near home may have been tracked on Strava app. *The Record Media*. https://therecord.media
12. Wikipedia contributors. (n.d.). Rafic Hariri. In *Wikipedia, The Free Encyclopedia*. https://en.wikipedia.org/wiki/Rafic_Hariri
13. Kurz, H., et al. (2021). Shadow attacks: Hiding and replacing content in signed PDFs. *NDSS Symposium*. https://ndss-symposium.org
14. Tenable. (2022). *CVE-2022-25641 vulnerability details*. Tenable Research. https://tenable.com
15. BRG. (2023). *Nervous system: How legal tech helped catch the BTK killer*. Berkeley Research Group. https://thinkbrg.com
16. Make Tech Easier. (2017). *In May 2017, an NSA contractor named Reality Winner mailed a printed classified document*. https://maketecheasier.com
17. Voisin, J. (2023). *MAT2 - Metadata Anonymisation Toolkit* [GitHub repository]. GitHub. https://github.com/jvoisin/mat2
18. MediaCircle. (2022). *Clearswift adaptive redaction*. https://mediacircle.de
19. Red Eagle Tech. (2023). *What is content disarm and reconstruction (CDR)? Complete guide*. https://redeagle.tech
20. Garfinkel, S. (2023). *De-identifying government datasets: Techniques and governance (NIST SP 800-188).* NIST Technical Series Publications. https://nvlpubs.nist.gov
21. NIST. (2020). *NIST informative references for the Privacy Framework*. https://nist.gov
22. Dataintelo. (2024). *Metadata removal tools market research report 2034*. https://dataintelo.com

Отримано редакцією журналу / Received: 02.02.26
Прорецензовано / Revised: 16.02.26
Схвалено до друку / Accepted: 24.09.26