Міністерство освіти і науки України

Харківський національний університет імені В.Н. Каразіна

Факультет комп'ютерних наук

Спеціальність 125 «Кібербезпека»

Освітня програма «Безпека інформаційних та комунікаційних систем»

«Допущено до захисту»

Зав.кафедрою БІСТ

Сергій РАССОМАХІН

_____

« » 2022 р.

Пояснювальна записка

до кваліфікаційної роботи магістра

на тему: «Дослідження та аналіз інструментів і методів, що використовує механізм OSINT»

оцінка « »

Голова ЕК

Доценко С.І. _____

Керівник к.т.н. Єсіна М. В.

Рецензент к.т.н. Бобух В. А.

Виконавець : студент групи КБ-61

_____ Азаров М.О.

Харків – 2022

РЕФЕРАТ

Пояснювальна записка до проекту магістра містить 91 сторінки, 8 рисунків, 1 таблицю, 1 додаток, 36 посилань на джерела.

Мета роботи полягає в дослідженні та аналізу методології OSINT, основних інструментів та методів, що використовує OSINT, ресурсів, що використовуються для збору інформації у рамках проведення розвідки на основі відкритих джерел, а також проблематики проведення розвідки та факторів успішного аналізу.

Об’єкт дослідження – механізм розвідки на основі відкритих джерел OSINT.

Предмет дослідження – основні інструменти та методи, що використовує OSINT, ресурси, що використовуються для збору інформації при проведенні розвідки на основі відкритих джерел, проблематика ведення розвідки, а також фактори успішного проведення OSINT

Основними методами досліджень є аналіз та порівняння основних інструментів та методів, а також ресурсів для збору інформації при проведенні розвідки на основі відкритих джерел.

У роботі досліджено: методологію OSINT, основні інструменти та методи, що використовує OSINT, ресурси, що використовуються для збору інформації у рамках проведення розвідки на основі відкритих джерел, а також проблематика проведення розвідки та фактори успішного аналізу.

Результати роботи можуть бути використані у різних наукових виданнях, а також для кращого розуміння розвідки на основі відкритих джерел.

Ключові слова: OSINT, OSINF, РОЗВІДКА НА ОСНОВІ ВІДКРИТИХ ДЖЕРЕЛ, ВІДКРИТА ІНФОРМАЦІЯ, ГЛИБИННИЙ ВЕБ-ПРОСТОР, СОЦІАЛЬНА МЕРЕЖА, АНАЛІТИК, РОЗВІДНИК, ВІДКРИТЕ ДЖЕРЕЛО, РЕСУРС OSINT, ЗБІР ІНФОРМАЦІЇ, ІНСТРУМЕНТ OSINT.

ABSTRACT

The explanatory note to the master's project contains 91 pages, 8 figures, 1 table, 1 appendix, 36 references to sources.

The purpose of the work is to study and analyze the OSINT methodology, the main tools and methods used by OSINT, the resources used to gather information in the framework of conducting intelligence based on open sources, as well as the issues of conducting intelligence and the factors of successful analysis.

The object of research is the intelligence mechanism based on open sources of OSINT.

The subject of the study is the main tools and methods used by OSINT, the resources used to gather information when conducting intelligence based on open sources, the problems of conducting intelligence, as well as the factors of successful conduct of OSINT

The main methods of research are the analysis and comparison of the main tools and methods, as well as resources for gathering information when conducting intelligence based on open sources.

The work examines: the OSINT methodology, the main tools and methods used by OSINT, the resources used to gather information in the framework of conducting intelligence based on open sources, as well as the problems of conducting intelligence and the factors of successful analysis.

The results of the work can be used in various scientific publications, as well as for a better understanding of intelligence based on open sources.

Keywords: OSINT, OSINF, OPEN SOURCE INTELLIGENCE, OPEN INFORMATION, DEEP WEB, SOCIAL NETWORK, ANALYST, INTELLIGENCE, OPEN SOURCE, OSINT RESOURCE, INTELLIGENCE GATHERING, OSINT TOOL.

ЗМІСТ

ПЕРЕЛІК ПОЗНАЧЕНЬ І СКОРОЧЕНЬ

OSINT	– Open Source Intelligence
OSINF	– Open Source Information
FBMS	– Foreign Broadcast Monitoring Service
FBIS	– Foreign Broadcast Intelligence Service
ЦРУ	– Центральне Розвідувальне Управління
OSC	– Open Source Center
BBC	– British Broadcasting Corporation
ЗМІ	– Засоби Масової Інформації
SWB	– Summary of World Broadcasting
FRD	– Federal Research Division
EMM	– European Media Monitor
JRC	– Joint Research Center
WNC	– World News Connection
NTIS	– National Technical Information Service
OSINT-V	– Open Source Intelligence-Verified
OSD	– Open Source Data
IoT	– Internet of Things
ISBN	– International Standard Book Number
ISSN	– International Standard Serial Number
SaaS	– Software as a Service
NLP	– Natural Language Processing

ВСТУП

Значний обсяг інформаційних ресурсів у глобальних мережах містить різні експертні оцінки, частина яких пов'язані з реалізацією кібернетичних інформаційних загроз [1].

У даний час люди використовують Інтернет для передачі цінної інформації або файлів у розважальних цілях, покупки різних продуктів, встановлення зв'язку з іншими людьми, а також використання веб-сайтів соціальних мереж для спілкування з глобальним світом без географічних бар'єрів. Як зазначають підприємства з кібербезпеки, до 2030 року 90% населення світу користуватимуться Інтернетом, малі та літні люди будуть знаходитися в Інтернеті, а це означає, що буде понад 8 мільярдів користувачів Інтернету [2]. Оскільки світ продовжує залишатися цифровим, комп'ютеризовані соціальні порядки надаватимуть колосальну кількість комп'ютеризованої інформації, створеної людьми, та обмінюватимуться інтелектуальними даними в Інтернеті.

Експертні оцінки, що містяться у відкритих документах, можуть бути проаналізовані, синтезовані, а також створені як за основу для подальшого прийняття рішень. Вони відрізняються від традиційних експертних оцінок як обсягами, так і об'єктивністю. Крім того, в мережах може бути інформація, пов'язану з організацією протиправної діяльності, зокрема, кіберзлочинності, кібертероризму.

Виходячи з цього, облік інформації з веб-каналів має велике значення для вирішення завдань у сфері забезпечення кібербезпеки.

Сьогодні так званий Open Source INTelligence (OSINT) є одним з найважливіших інструментів кібербезпеки. OSINT – це одна з областей розвідки, що включає пошук, відбір та збір розвідувальної інформації, доступної із загальнодоступних джерел, а також аналіз цієї інформації.

Під цим визначенням широкий набір джерел буде вважатися частиною OSINT. Наприклад, дані, що вільно розміщуються на веб-сайтах соціальних мереж, повідомлення на зборах та групових чатах, реєстри незахищених веб-сайтів та будь-які дані, які можна знайти в Інтернеті. Безперечно, більшість активів OSINT неможливо знайти за допомогою пошукових систем, таких як Google або Yahoo!, оскільки багато активів приховані в глибокій і темній мережі, і такі активи становлять понад 96% чистого контенту [2].

OSINT зазвичай проводиться шляхом моніторингу, аналізу та дослідження інформації, що надходить з Інтернету. Матеріали, складені на основі інформації з відкритих джерел, підтримують усі методи та діяльність розвідки шляхом накопичення розвідувальних знань, їх аналізу та розповсюдження.

За даними аналітика ЦРУ Шермана Кента (1947 р.), політики одержують до 80% інформації, необхідної для прийняття рішень у мирний час, із відкритих джерел. Пізніше генерал-лейтенант Семюел Вілсон, начальник відділу розвідки Міністерства оборони США у 1976-1977 роках, зазначав, що 90% розвідувальних даних надходить із відкритих джерел, і лише 10% – від агентів [1].

Той факт, що відкриті джерела часто надають більшу частину розвідувальних даних, робить OSINT важливою частиною зусиль збору розвідувальних даних з усіх джерел. Кожен спеціаліст з розвідки повинен бути знайомий з джерелами та методами OSINT, тим більше, що аналіз та збір інформації все більше зливаються один з одним. Проте, інформаційна діяльність та використання відкритого вихідного коду мають підтримуватись спеціалізованими елементами, щоб аналітики не відставали від нових технологій та ринку. Спеціалізовані експерти OSINT найбільш кваліфіковані для виявлення потенційних прогалин у можливостях та оцінки того, де підрядники можуть бути корисними. Одним із добрих способів інтеграції знань та навичок приватного сектору до розвідувальної спільноти є програма сертифікації OSINT, яка в даний час впроваджується, наприклад, у США [3].

Через стрімкий розвиток сучасних інформаційних технологій ця технологія набуває популярності й у відповідних силових структурах України.

Особливу актуальність дана технологія в Україні отримала після повномасштабного вторгнення Росії, оскільки початкова інформація з відкритих джерел після певного аналізу і опрацювання може стати цінним знанням, особливо в умовах війни.

Фактично, OSINT – це та ж розвідка, але на заміну величезному людському ресурсу приходять сучасні технології, штучний інтелект, який допомагає швидко дізнатися про те, що, як і де відбувається.

1 АКТУАЛЬНИЙ СТАН ТА МІЖНАРОДНИЙ ДОСВІД ВИКОРИСТАННЯ МЕХАНІЗМУ OSINT

1.1 Базові аспекти розвідки на основі відкритих джерел

Історія використання відкритої інформації сягає появи розвідувальних даних як інструменту, що підтримує рішення та дії уряду. Тим не менш, це не було систематичним зусиллям, поки Сполучені Штати не стали піонерами в інституціоналізації та професіоналізації автономного потенціалу для моніторингу іноземних ЗМІ, створивши Службу моніторингу іноземного мовлення (FBMS), яка виросла з дослідницької ініціативи у Прінстонському університеті. FBMS швидко набрала обертів після нападу японців на Перл-Харбор. У 1947 році вона була перейменована на Службу зовнішньої радіомовної розвідки (FBIS) і передана у відання нещодавно створеного ЦРУ. У 2005 році, після терактів 11 вересня та прийняття Закону про реформу розвідки та запобігання тероризму, FBIS разом з іншими дослідницькими елементами було перетворено на Центр відкритих джерел (OSC) Національної розвідки. З моменту свого заснування зусилля OSINT відповідали за фільтрацію, розшифрування, переклад (в такий спосіб інтерпретацію) та архівування новин та інформації з усіх типів іноземних джерел ЗМІ [3].

У 1939 році британський уряд звернувся до Британської радіомовної корпорації (BBC) з проханням створити цивільну, а потім і комерційну службу, що вивчає іноземну друковану журналістику та радіомовлення, з її дайджестом іноземних радіопередач, пізніше названим «Зведення світових радіопередач» (SWB), а тепер відомий як BBC Monitoring. Як йдеться у довіднику BBC 1940 року, мета полягала в тому, щоб звести «сучасну Вавилонську вежу, де із зразковою концентрацією вони прислухатимуться до голосів друзів та ворогів». До середини 1943 року BBC щоденно відстежувала 1,25 мільйона слів в ефірі. Офіційне партнерство між BBC та її

американським партнером було встановлено у 1947/48 році з угодою про повний обмін продукцією. Також у 1948 році на базі Відділу авіаційних досліджень було створено дослідницький підрозділ Бібліотеки Конгресу США для надання спеціалізованих дослідницьких та аналітичних послуг із використанням великих фондів бібліотеки. Тепер він відомий як Федеральний дослідницький підрозділ (FRD).

Під час холодної війни країни по обидва боки «залізної завіси» створили можливості для збору даних з відкритих джерел, які часто вбудовані в їхні секретні розвідувальні служби. Відкриті джерела не тільки «становили більшу частину всієї розвідувальної інформації, але й зрештою стали провідним джерелом» інформації про військовий потенціал та політичні наміри противників, включаючи раннє попередження та прогнозування загроз[3]. Наприклад, міністерство державної безпеки Східної Німеччини (відоме як «Штазі») щомісяця аналізувало 1000 західних журналів та 100 книг, а також щодня узагальнювало понад 100 газет та 12 годин західнонімецького радіо та телебачення [3].

Відкриті джерела під час холодної війни вже були джерелом інформації, що добре зарекомендувало себе, часто першим засобом для націлювання на інші зусилля зі збору або «зовнішнім шматочком головоломки». Завдяки інтернет-технологіям загальнодоступна інформація дуже вплинула на всі аспекти сучасного політичного, соціального та економічного життя. Однак потрібно знати, що Інтернет сам по собі не є джерелом (за винятком його метаданих), скоріше це засіб передачі інформації та віртуальне місцезнаходження.

Більшість розвідувальних спільнот повільно оцінили цінність Інтернету з двох причин [3]:

- Розвідувальні агентства шукають інформаційну перевагу, таємно працюючи із секретами. Опора на відкриту інформацію та відповідні обмеження авторського права суперечить цій ідеї.
- У більшості випадків складніше, ризиковано і дорого застосовувати підпільні методи для отримання секретних джерел, що справляє враження, що ці

джерела повинні бути більш цінними, ніж відкриті джерела, плутаючи метод із продуктом або помилково приймаючи секретність для знань.

Після розпаду Радянського Союзу західні спецслужби переорієнтували свої операції на нові географічні та тематичні пріоритети, такі як Африка та Азія, недержавні діячі, конфлікти низької інтенсивності в експедиційних умовах, політичний та релігійний тероризм, поширення зброї масової ураження та вразливості комп'ютерних мереж, що призвело до більшого акценту на відкритих джерелах. Військові США вперше вигадали термін OSINT наприкінці 1980-х років, стверджуючи, що реформа розвідки необхідна, щоб упоратися з динамічним характером інформаційних вимог, особливо на тактичному рівні на полі бою. У 1992 році Закон про реорганізацію розвідки визначив цілі збору інформації як «надання своєчасних, об'єктивних та неупереджених відомостей на основі всіх джерел, доступних розвідувальному співтовариству США, як відкритих, так і закритих». У 1996 році Комісія з функцій і можливостей розвідувальної спільноти США (відоміша як Комісія Аспіна-Брауна) дійшла висновку, що «необхідно докласти більше зусиль для використання великої інформації, доступної в даний час з відкритих джерел». Паралельні зусилля НАТО зі створення основи для використання OSINT призвели до публікації кількох довідників, букварів та практичних посібників різної якості. За допомогою European Media Monitor (EMM) та OSINT Suite, серед інших інструментів та проектів, Об'єднаний дослідний центр (JRC) Комісії Європейського Союзу (ЄС) розробляє свої власні інструменти для вирішення проблем, які стають перед ними.

11 вересня 2001 року стало переломним моментом для OSINT: Національна комісія з терористичних атак на Сполучені Штати (Комісія 11 вересня) у 2004 році рекомендувала створити агентство з відкритим вихідним кодом без подальших коментарів чи подробиць. Ця концепція була підхоплена у 2005 році разом з відповідними рекомендаціями Комісії з розвідувальних можливостей США щодо зброї масового ураження, коли директор Національної розвідки заснував OSC, поглинувши FBIS ЦРУ із World News Connection (WNC) під контролем Національної

служби технічної інформації (NTIS). OSC позиціонує себе як «головний постачальник іноземних розвідувальних даних з відкритих джерел для уряду США і надає інформацію з іноземних політичних, військових, економічних і технічних питань, що виходять за рамки звичайних засобів масової інформації, з безлічі відкритих джерел, що постійно розширюється». У той же час було призначено помічника заступника директора Національної розвідки з відкритих джерел, що підвищило впізнаваність Національного підприємства з відкритих джерел. З розвитком регіональних центрів обробки даних, які зосереджені на питаннях внутрішньої безпеки та правоохоронних органів, OSINT стає основним джерелом для об'єднання та консолідації відповідної інформації у дієві продукти.

1.2 Сутність механізму розвідки на основі відкритих джерел

Інтернет-розвідка стрімко набирає популярності у сфері кібербезпеки і не тільки. Збором таємної інформації користуються детективи, слідчі, поліцейські, розвідувальні групи. А в умовах нинішньої ситуації в Україні навіть звичайнісінька людина може допомогти ЗСУ боротися з ворогом за допомогою програм для розвідки. Пошук потрібних даних серед відкритих джерел можна проводити з більшою ефективністю, ніж просто користуватися Інтернет-пошуковиками.

Розвідка за відкритим джерелами (Open Sourse Intelligence) є несекретною інформацією, яка була навмисно виявлена, виділена, очищена і поширена серед обраної аудиторії для вирішення конкретного питання [4]. Це забезпечує дуже міцну основу інших дисциплін розвідки. При систематичному застосуванні, продукти OSINT можуть знизити потребу в секретних ресурсах для збору розвідданих, обмежуючи запити інформації лише тими питаннями, на які не можна відповісти з відкритих джерел.

Враховуючи деякі вхідні дані, а також застосування передових методів збору та аналізу, OSINT постійно розширює знання про мету. Таким чином, знайдена інформація знову підживлює процес збору, щоб наблизитись до кінцевої мети [5]. У

даний час OSINT широко використовується урядами та спецслужбами для проведення розслідувань та боротьби з кіберзлочинністю [6]. Тим не менш, він використовується не тільки для державних справ, але й застосовується для різних цілей.

Дійсно, поточні дослідження зосереджені (але не обмежуються ними) на трьох основних додатках, які представлені на рис. 1.1 та описані далі:

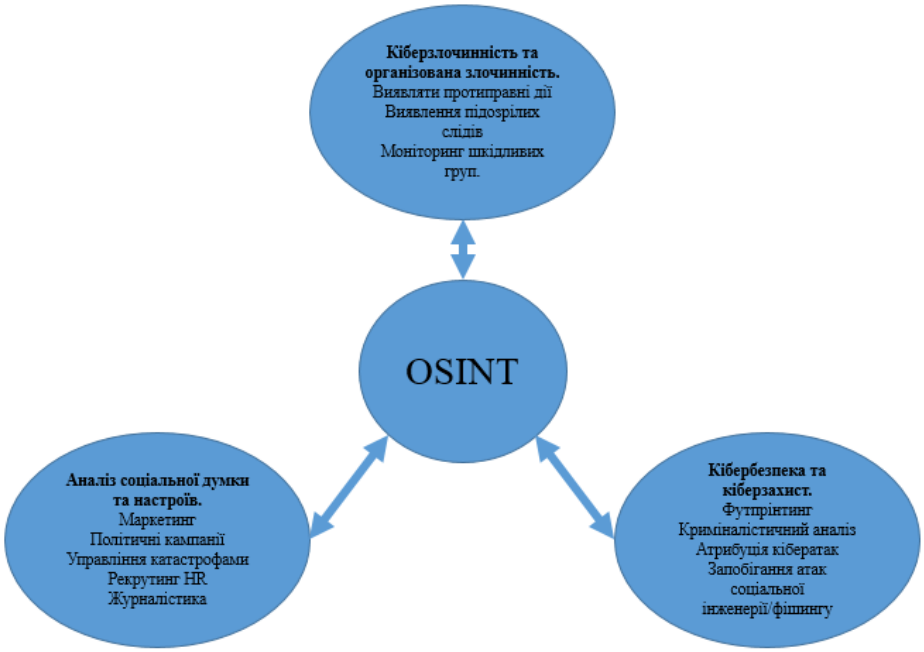

Рисунок 1.1 – Основні варіанти використання OSINT

– Аналіз громадської думки та настроїв. Поряд з бумом онлайнових соціальних мереж можна збирати дані про взаємодії користувачів, повідомлення, інтереси та переваги для отримання неявних знань. Докази, зібрані із соціальних мереж, мають далекосяжні наслідки і широко використовуються [7]. Такий збір та аналіз можуть застосовуватися, наприклад, у маркетингу, політичних кампаніях чи управлінні стихійними лихами [8].

– Кіберзлочинність та організована злочинність. Відкриті дані постійно аналізуються та зіставляються за допомогою процесів OSINT, щоб виявити злочинні наміри на ранній стадії. Зважаючи на патерни зловмисників та взаємозв'язок між злочинами, розвідка по відкритим джерелам може надати силовикам можливість

оперативно виявляти протиправні дії [9]. У цьому напрямі, використовуючи відкриті дані, можна було б відстежувати діяльність терористичних організацій, які дедалі активніше діють у Інтернеті [10], [11].

– Кібербезпека та кіберзахист. Системи ІКТ (інформаційні та комунікаційні технології) постійно піддаються атакам з боку злочинців, які прагнуть порушити доступність послуг [12]. Отже, дослідження стають критично важливими для захисту цих систем від кібератак, зокрема шляхом вирішення проблем, які все ще залишаються відкритими в галузі кібербезпеки [13]. У цьому сенсі наука про дані застосовується як визначення слідів у пентестах, але й превентивного захисту організацій та компаній. Зокрема, методи інтелектуального аналізу даних можуть допомогти, виконуючи аналіз щоденних атак, зіставляючи їх та підтримуючи процеси прийняття рішень для ефективного захисту, а також для швидкого реагування [14]. Так само OSINT можна розглядати у цьому контексті як джерело інформації для відстеження та розслідувань. Судово-цифровий аналіз може містити OSINT для доповнення цифрових доказів, залишених інцидентом[15].

1.2.1 Категорії інформації з відкритих джерел

Існують різні види інформації, з якою можна зіткнутися під час проведення OSINT-аналізу. Згідно з опублікованим у 2001 році Довідником НАТО з розвідки за відкритими джерелами версії, існує чотири категорії відкритої інформації та розвідувальних даних[4]:

– Дані з відкритих джерел (OSD): це загальні дані, що надходять із основного джерела. Приклади включають супутникові зображення, дані телефонних дзвінків та метадані, набори даних, дані опитування, фотографії та аудіо- або відеозаписи, в яких записано подію.

– Інформація з відкритих джерел (OSINF): це загальні дані, які спочатку зазнали деякої фільтрації, щоб відповідати певному критерію чи потребі; ці дані

можна назвати вторинним джерелом. Приклади включають книги з певної теми, статті, дисертації, витвори мистецтва та інтерв'ю.

– Розвідка з відкритих джерел (OSINT): включає всю інформацію, яка була виявлена, відфільтрована та призначена для задоволення конкретних потреб чи цілей. Ця інформація може бути використана безпосередньо у будь-якому розвідувальному контексті. Коротко OSINT можна визначити як результат обробки матеріалів із відкритих джерел.

– Затверджений OSINT (OSINT-V): це OSINT із високим ступенем достовірності. Дані повинні бути підтверджені (перевірені) з використанням джерела, що не відноситься до OSINT, або джерела OSINT з високою репутацією. Це важливо, оскільки деякі зовнішні зловмисники можуть поширювати неточну інформацію OSINT з наміром ввести в оману OSINT. Хорошим прикладом цього є, коли телевізійна станція транслює у прямому ефірі прибуття президента до іншої країни. Така інформація є OSINT, але вона має велику міру достовірності.

Як можна зрозуміти, OSD і OSINF включають в себе основні джерела (первинні і вторинні) інформації, яку OSINT використовує для отримання результатів.

Існує ще одна проблема, яку необхідно зрозуміти в контексті OSINT, – це різниця між даними, інформацією та знаннями. Ці три терміни зазвичай використовуються як взаємозамінні; однак кожен із них має різне значення, хоча всі троє взаємодіють один з одним.

– Дані: це набір фактів, що описують будь-що без подальшого пояснення чи аналізу. Наприклад, "Ціна золота за унцію становить 1212 доларів".

– Інформація: тип даних, які були правильно інтерпретовані для надання корисного значення в конкретному контексті. Наприклад, "Ціна золота за унцію впала з 1212 до 1196 доларів за один тиждень".

– Знання: це поєднання інформації, досвіду та розуміння, які були вивчені чи отримані після деяких експериментів. Знання описують те, що ваш мозок записав у минулому, і ці записи можуть допомогти вам приймати більш правильні рішення про

майбутнє, стикаючись із подібними контекстами. Наприклад: "Коли ціна на золото падає більш, ніж на 5 відсотків, це означає, що ціна на нафту теж впаде".

1.2.2 Типи OSINT

OSINT включає всі загальнодоступні джерела інформації. Цю інформацію можна знайти онлайн або офлайн, зокрема в таких місцях[16]:

– Інтернет, який включає наступне та багато іншого: форуми, блоги, сайти соціальних мереж, сайти для обміну відео, такі як YouTube, wikis, Whois records про зареєстровані доменні імена, метадані та цифрові файли, ресурси темної мережі, дані геолокації, IP адреси, системи пошуку людей і все, що можна знайти в Інтернеті.

– Традиційні засоби масової інформації (наприклад, телебачення, радіо, газети, книги, журнали).

– Спеціалізовані журнали, наукові публікації, дисертації, матеріали конференцій, профілі компаній, річні звіти, новини компаній, профілі співробітників і резюме.

– Фотографії та відео, включаючи метадані.

– Геопросторова інформація (наприклад, карти та комерційні зображення).

Зробивши висновки із вищевказаного, можна зрозуміти, що OSINT охоплює не лише онлайн-джерела. Паперові видання загальнодоступних джерел також повинні бути ретельно досліджені в рамках процесу збору OSINT. Проте онлайн-джерела становлять найбільший сегмент OSINT.

Сьогодні ми живемо у вік інформації, і видавці, а також корпорації, університети та інші постачальники джерел OSINT переводять свої бізнес-процеси у цифрові формати. Кількість користувачів на сайтах соціальних мереж також продовжуватиме зростати, а кількість пристроїв Інтернету речей (IoT) у майбутньому збільшуватиметься, що призведе до величезного збільшення обсягу цифрових даних, що надходять від мільярдів датчиків та машин у всьому світі. Іншими словами, більшість джерел OSINT у майбутньому будуть онлайн-джерелами.

Обсяг цифрових даних стрімко зростає. За даними IDC Research [17], до 2020 року загальний обсяг цифрових даних, створених у всьому світі, досягне 44 зеттабайт, і протягом п'яти років це число зростатиме швидше, досягнувши 180 зеттабайт у 2025 році. За оцінками дослідницької групи Gartner, до 2020 року середньостатистична людина проводитиме більше часу, взаємодіючи з автоматичними ботами, ніж із живими людьми, і, звичайно, всі ці взаємодії будуть цифровими. За іншою оцінкою, у 2021 році 20% всіх дій людини будуть пов'язані з використанням сервісу як мінімум однієї з гігантських IT-компаній (Google, Apple, Facebook, Amazon). Не кажучи вже про те, що більшість людей віддадуть перевагу використанню голосових команд для взаємодії зі своїми комп'ютерними пристроями, ніж друкувати необхідне руками. Ці цифри дають уявлення про те, як виглядатиме найближче майбутнє в епоху цифрових технологій. Обсяг цифрових даних поряд із збільшенням кількості людей, які використовують Інтернет для виконання своєї роботи, зробить онлайн-джерела основним джерелом OSINT як для урядів, так і для комерційних корпорацій у майбутньому.

1.2.3 OSINT-організації

Деякі спеціалізовані організації надають послуги OSINT. Деякі з них є державними, а інші є приватними компаніями, які пропонують свої послуги різним сторонам, таким як державні установи та бізнес-корпорації на основі підписки. У цьому підрозділі будуть визначені основні організації OSINT у всьому світі[16].

1) Урядові організації

Урядові організації, які займаються аналізом OSINT, як і раніше, вважаються кращими через ресурси, доступні від їх урядів для виконання своєї роботи. Двома найвідомішими урядовими агентствами, які проводять OSINT у всьому світі, є «Open Source Center» в США та «BBC Monitoring» у Великій Британії.

– Open Source Center. OSC – найбільша організація OSINT, що має величезні ресурси для виконання своєї роботи. OSC тісно співпрацює з іншими місцевими спецслужбами США та пропонує свої послуги урядовим спецслужбам США.

Перед Open Source Center, створеним у 2005 році, було поставлено завдання збирати та аналізувати інформацію з відкритих джерел, яка представляє розвідувальну цінність, із усіх засобів масової інформації – друкованих, радіомовних та онлайн. OSC був наступником Інформаційної служби іноземного мовлення (FBIS), яка збирала та перекладала світові новини та іншу інформацію з відкритих джерел протягом півстоліття.

У рамках зусиль із модернізації Агентства, оголошених директором Бреннаном, 1 жовтня 2015 року Open Source Center (OSC) змінив свою назву на Open Source Enterprise (OSE), – згідно із заяви представника ЦРУ Райана Трапані. OSE, як і раніше, займається збором, аналізом та розповсюдженням загальнодоступної інформації, що має значення для розвідки. Нова назва організації відображає широку актуальність та масштаб місії відкритого вихідного коду.

OSE зберігає свою роль центру передового досвіду розвідувальної спільноти для збору, аналізу та торгівлі відкритими джерелами.

Як і FBIS протягом кількох десятиліть, Open Source Center випускав загальнодоступну лінійку продуктів, включаючи переклади та аналіз відкритих джерел інформації. Але наприкінці 2013 року, на розчарування багатьох давніх передплатників, ЦРУ раптово закрило цей канал публічної інформації, пославшись на витрати та доступність альтернативних загальнодоступних джерел.

– BBC Monitoring. BBC Monitoring (https://monitoring.bbc.co.uk/login) – це підрозділ Британської радіомовної корпорації (BBC), який відстежує іноземні ЗМІ по всьому світу. Дана організація виконує ту саму роль, як і Open Source Center у США, з тією різницею, що він не належить британській розвідці. BBC Monitoring фінансується заінтересованими сторонами, а також багатьма комерційними та державними організаціями по всьому світу. Вперше вона була створена у 1939 році і

має офіси у різних країнах світу. Підрозділ Британської радіомовної корпорації активно відстежує телебачення, радіопередачі, друковані ЗМІ, Інтернет та нові тенденції зі 150 країн більш, ніж 70 мовами. BBC Monitoring управляється BBC і пропонує свої послуги з передплати заінтересованим сторонам, таким як комерційні організації та офіційні органи Великобританії.

2) Організації приватного сектору

Існує багато приватних корпорацій, які розробили передові програми та методи збору даних із загальнодоступних джерел у комерційних цілях. Дійсно, більшість приватних корпорацій OSINT співпрацюють із державними установами, щоб надати їм таку інформацію.

– Jane’s Information Group. Jane's Information Group (https://www.janes.com) – британська компанія, заснована у 1898 році. Jane's – провідний постачальник, що спеціалізується на військовій сфері, тероризмі, державній стабільності, серйозній та організованій злочинності, розвідувальних даних про розповсюдження та закупівлі, аерокосмічній та транспортній тематиці. Дана компанія видає безліч журналів та книг, пов'язаних з питаннями безпеки, на додаток до своїх джерел OSINT, які відстежують та прогнозують проблеми безпеки у 190 штатах та 30 територіях.

– Economist Intelligence Unit. Economist Intelligence Unit (https://www.eiu.com/home.aspx) – це підрозділ бізнес-аналітики, досліджень та аналізу британської Economist Group. Основною сферою діяльності Economist Intelligence Unit є бізнес- та фінансові прогнози. Також дана компанія пропонує щомісячний звіт на додаток до економічного прогнозу країни на найближчі п’ять років із всебічним уявленням про поточні тенденції з економічних та політичних питань.

– Oxford Analytica. Oxford Analytica (http://www.oxan.com) – відносно невелика OSINT-компанія порівняно з двома попередніми. Oxford Analytica спеціалізується на геополітиці та макроекономіці. Дана компанія має глобальну мережу експертів з макроекономіки, які консультують своїх клієнтів щодо передових

методів стратегії та ефективності при доступі до складних ринків. Експертні мережі Oxford Analytica налічують понад 1400 експертів. Більшість із них є вченими у своїй галузі, старшими викладачами провідних університетів та висококласними фахівцями у своїй галузі.

3) Продавці сірої літератури

Даний тип даних, що використовується при зборі даних OSINT, має велику розвідувальну цінність. Сіра література переважно виробляється світовими видавничими компаніями. Сюди входять книги, журнали, газети та все, що публікується публічно. Однак існує ще один тип «сірої літератури», яка називається «сірою інформацією», до якої висуваються інші вимоги.

Зазвичай терміни «сіра література» та «сіра інформація» взаємозамінні. Однак у галузі розвідки вони дещо відрізняються. Сіра література відноситься до всіх публікацій, які можна отримати традиційними каналами книгарень, а сіра інформація відноситься до інших публікацій, які не можна отримати традиційними способами. Отже, сіра інформація має власні канали, і її може бути важко ідентифікувати і отримати. Сіра інформація включає наступне: академічні статті, препринти, матеріали конференцій і дискусій, дослідні звіти, маркетингові звіти, технічні специфікації та стандарти, дисертації, тези, галузеві публікації, меморандуми, урядові звіти та документи, не опубліковані в комерційних цілях, переклади, інформаційні бюлетені, огляди ринку, звіти про поїздки та програми фестивалів тощо.

Сіру літературу можна поділити на три основні види:

– Біла: включає все, що опубліковано для продажу через традиційні канали книгарень. Видання має мати номер ISBN або ISSN і може бути отримане у продавця. Книги, журнали та газети потрапляють до цієї категорії.

– Тимчасова: цей тип недовговічний. Приклади включають розклади польотів, чернетки, копії рахунків-фактур, рекламні оголошення, плакати, квитки, візитні картки та все, що публікується самостійно.

– Сіра: містить поєднання двох згаданих типів. Як правило, можна отримати сіру літературу, заплативши абонентську плату за такий контент або купивши книги, журнали та інші публікації безпосередньо в книгарнях. Щоб отримати більше схованої сірої інформації, розвідникам доводиться використовувати інші спеціалізовані сервіси. Нижче наведено деякі з них:

– Factiva (http://new.dowjones.com/products/factiva) – глобальна база даних новин із ліцензованим контентом. Дана організація збирає дані з більш ніж 33000 джерел преміум-класу, і багато цих джерел (74%) ліцензовані і не можуть бути знайдені у вільному доступі в Інтернеті. Factiva збирає джерела 28 мовами на додаток до своєї унікальної можливості надання доступу до ресурсів, які ще не були опубліковані їхніми творцями.

– LexisNexis (https://www.lexisnexis.com/en-us/gateway.page) в даний час належить RELX Group (раніше Reed Elsevier). Спочатку дана компанія була зосереджена на наданні високоякісних юридичних та журналістських документів, але розширила сферу свого охоплення, включивши до нього більше послуг, таких як інструменти моніторингу ЗМІ, інструменти управління поставками, рішення для аналізу продажів, інструменти аналізу ринку та рішення для управління ризиками, які аналізують загальнодоступний та галузевий контент для прогнозування ризиків та поліпшення прийняття рішень.

Нижче перераховані інші компанії, які спеціалізуються на зборі онлайн-інформації як з відкритих, так і приватних джерел:

– InsideView (https://www.insideview.com) – це компанія, яка займається програмним забезпеченням як послугою (SaaS), яка збирає статистичні дані та зв’язки з понад 40000 джерел бізнес-інформації, контактних даних, онлайн-новин і соціальних мереж. Заснована у 2005 році, InsideView в основному використовується відділами маркетингу, продажів і операцій для виявлення та збору інформації про клієнтів і потенційних клієнтів. У квітні 2021 року компанію придбала Demandbase.

– Semantic Visions (www.semantic-visions.com) – це заснована на програмному забезпеченні аналітична фірма, що базується в Празі та Лондоні, використовує систему розвідки з відкритим вихідним кодом (OSINT) військового рівня, яка збирає та аналізує 90% світового контенту новин. Компанія Semantic Visions, заснована у 2011 році, використовує багатомовну систему раннього попередження для захисту клієнтів від загроз до того, як вони матеріалізуються. Semantic Visions прагне захищати демократію за рахунок виявлення дезінформації та ворожої пропаганди, а також за рахунок забезпечення спільної ситуаційної поінформованості про події та тенденції, що виникають.

– DigitalGlobe (www.digitalglobe.com) – американська компанія, комерційний оператор кількох цивільних супутників дистанційного зондування Землі, великий постачальник результатів супутникової зйомки та геопросторових даних (зокрема для Google Maps/Earth та Virtual Earth). Компанія була заснована у 1992 році, випустила акції на NYSE у травні 2009 року, отримавши капіталізацію у 279 мільйонів доларів США [18].

Вона є оператором цивільних супутників дистанційного зондування Землі надвисокої роздільної здатності WorldView-1 (роздільна здатність – 50 см), WorldView-2 (46 см), QuickBird (61 см), GeoEye-1 (41 см) та IKONOS (1 м). Загальна добова продуктивність угрупування – понад 3,5 млн. км^2.

1.2.4 Сторони, які зацікавлені в OSINT-інформації

OSINT може бути корисним для різних учасників. У даному підрозділі буде перераховано і роз’яснено, що спонукає кожного шукати ресурси OSINT[16].

– Уряд. Державні органи, особливо військові відомства, є найбільшим споживачем джерел OSINT. Величезні технологічні розробки та широке використання Інтернету в усьому світі зробили уряд величезним споживачем даних OSINT. Урядам потрібні джерела OSINT для різних цілей, таких як національна безпека, боротьба з тероризмом, кібервідстеження терористів, розуміння громадської думки всередині

країни та за кордоном з різних питань, надання політикам необхідної інформації для впливу на їхню внутрішню та зовнішню політику, а також використання іноземних ЗМІ, таких як телебачення для отримання інформації, миттєві трансляції різних подій, що відбуваються ззовні. Спецслужби об'єднують легально доступну інформацію зі своїми таємно отриманими розвідувальними даними (наприклад, з використанням супутникових зображень-шпигунів, електронних станцій прослуховування та шпигунів), щоб відповісти на конкретне запитання або передбачити майбутнє. У цих людей є необхідні ресурси (гроші та обладнання) для збирання та аналізу величезних обсягів даних в Інтернеті. Очікується, що дії урядів з видобутку даних OSINT активізуватимуться в міру того, як ми неухильно рухаємося до того, що зараз є цифровим віком.

– Міжнародні організації. Міжнародні організації, такі як ООН, використовують джерела OSINT для підтримки миротворчих операцій у всьому світі. ООН врівноважує інтереси наддержав і національних держав, що розвиваються, при розробці своєї політики, яка вимагає, щоб вона була максимально прозорою. Для цього ООН виявила, що зручніше використовувати джерела OSINT (включно з комерційними супутниковими зображеннями) для розвідувальних потреб, а не покладатися на звіти держав-членів, політика яких може суперечити один одному. Гуманітарні організації, такі як Міжнародний Червоний Хрест, використовують джерела OSINT, щоб допомогти їм у зусиллях з надання допомоги під час кризи або стихійного лиха. Вони використовують дані OSINT для захисту свого ланцюжка поставок від терористичних груп, аналізуючи сайти соціальних мереж та програми для обміну повідомленнями в Інтернеті, щоб передбачити майбутні терористичні дії. НАТО сильно залежить від джерел OSINT для розвідувальних цілей і для планування операцій з підтримки миру. Крім того, комерційні супутникові знімки використовуються для планування операцій, оскільки не всі держави-члени мають такі об'єкти.

НАТО опублікувало три стандартні довідники щодо використання OSINT для громадськості:

- Довідник НАТО з розвідки за відкритими джерелами (https://archive.org/details/NATOOSINTHandbookV1.2).
- Програма для читання інформації з відкритих джерел НАТО (http://www.au.af.mil/au/awc/awcgate/nato/osint_reader.pdf).
- «Використання розвідувальних даних НАТО в Інтернеті» (http://nsarchive2.gwu.edu//NSAEBB/NSAEBB436/docs/EBB-005.pdf).

– Правоохоронні органи. Поліція використовує джерела OSINT для захисту громадян від жорстокого поводження, сексуального насильства, крадіжки особистих даних та інших злочинів. Це можна зробити, відстежуючи канали соціальних мереж на наявність цікавих ключових слів та зображень, щоб допомогти запобігти злочинам до їх ескалації. Правоохоронні органи використовують OSINT для моніторингу та відстеження мереж злочинців у різних країнах. Наприклад, вони використовують тактику OSINT для збору інформації про людей, що цікавлять, щоб створити повний профіль для кожного з них. Вони також використовують джерела OSINT для онлайн-підробки та порушення авторських прав.

– Бізнес-корпорації. Інформація – це сила, і корпорації використовують джерела OSINT для дослідження нових ринків, моніторингу діяльності конкурентів, планування маркетингової діяльності та прогнозування всього, що може вплинути на їхню поточну діяльність та майбутнє зростання. У минулому використання джерел OSINT було обмежено великими компаніями з добрим бюджетом розвідки. У даний час, з широким використанням Інтернету, невеликі компанії з обмеженим бюджетом можуть ефективно використовувати джерела OSINT та включати отриману інформацію до своїх бізнес-планів.

Підприємства також використовують аналітику OSINT для інших нефінансових цілей, таких як:

- Боротьба з витоком даних, знаючи, що розкриття бізнесу конфіденційної інформації та вразливість безпеки їхніх мереж є причиною майбутніх кіберзагроз.
- Створення своїх стратегій розвідки загроз шляхом аналізу джерел OSINT як зовні, так і всередині організації, а потім об'єднання цієї інформації з іншою інформацією для реалізації ефективної політики управління кіберризиками, яка допомагає їм захистити свої фінансові інтереси, репутацію та клієнтську базу.

OSINT особливо корисний для компаній, що працюють в оборонній промисловості, оскільки такі компанії повинні бути повністю обізнані про навколишні обставини своїх клієнтів, щоб розробляти та орієнтувати їх на відповідне обладнання.

– Тестувальники на проникнення та хакери Black Hat, злочинні організації. OSINT широко використовується хакерами та тестувальниками на проникнення (пентестерами) для збору інформації про конкретну мету в Інтернеті. Розвідка на основі відритих джерел також вважається цінним інструментом для допомоги у проведенні атак соціальної інженерії. Перший етап будь-якої методології тестування на проникнення починається з розвідки (тобто з OSINT). На рис. 1.2 показано основні етапи тестування на проникнення.

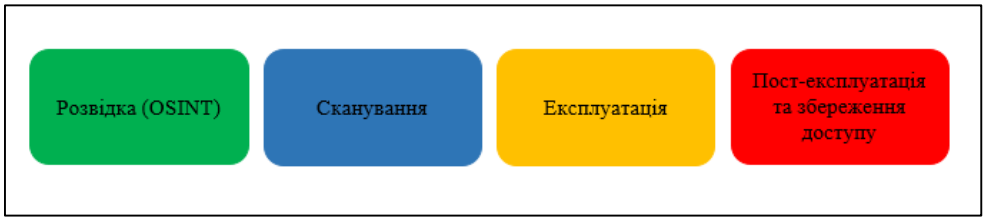

Рисунок 1.2 – Методика тестування на проникнення[16]

Компанії платять пентестерам за те, що вони зламують внутрішні мережі, щоб показати слабкі місця та засоби захисту від атак сторонніми каналами. Пентестери відрізняються від «чорних» хакерів, які використовують ці вразливості для отримання несанкціонованого доступу до конфіденційних даних. Проте обидві групи хакерів використовують одні й самі методи і інструменти розвідки для виконання своєї роботи.

– Люди, які піклуються про конфіденційність. Це звичайні люди, які можуть захотіти перевірити, як сторонні можуть проникнути в їхні обчислювальні пристрої і, що про них знає їхній Інтернет-провайдер. Їм також необхідно знати свій рівень онлайн-уразливості, щоб закрити будь-яку прогалину у безпеці та видалити будь-які особисті дані, які могли бути опубліковані ненавмисно. OSINT – чудовий інструмент для перегляду того, як ваша цифрова особа виглядає для зовнішнього світу, що дозволяє зберігати конфіденційність.

Приватні особи також можуть використовувати OSINT для боротьби з крадіжкою особистих даних, наприклад, якщо одна людина видає себе за іншу.

Дійсно, всі користувачі Інтернету так чи інакше використовують методи OSINT, наприклад, при пошуку чогось в Інтернеті. Будь то компанія, школа, університет чи людина, яку ви шукаєте, ви збираєте деяку форму розвідувальної інформації OSINT.

– Терористичні організації. Терористи використовують джерела OSINT для планування атак, збору інформації про цілі перед їх атакою (наприклад, використання супутникових зображень, таких як Google Maps, для дослідження місця розташування потенційної цілі), отримання військової інформації, випадково розкритої урядами (наприклад, як конструювати бомби) і поширення своєї пропаганди у світі, використовуючи різні канали засобів.

1.2.5 Види збору інформації

Джерела OSINT можна збирати трьома основними методами: пасивним, напівпасивним та активним. Використання одного на користь іншого залежить від сценарію, в якому працює процес збору, а також від типу даних, які цікавлять. Три методи збору зазвичай використовуються для опису способів роботи футпринтінгу, іншими словами, отримання технічної інформації про цільову IT-інфраструктуру (типи ОС, топологія мережі, імена серверів і т. д.)[16].

1.2.5.1 Пасивний збір

Це тип, що найчастіше використовується при зборі даних OSINT. Дійсно, всі методи розвідки OSINT повинні використовувати пасивний збір, тому що основною метою збору OSINT є збирання інформації про ціль лише через загальнодоступні ресурси. У цьому типі ваша ціль нічого не знає про вашу діяльність зі збору розвідданих. Цей вид пошуку є високо анонімним і має проводитися таємно. З технічної точки зору цей тип збору розкриває обмежену інформацію про ціль, оскільки ви не надсилаєте жодного трафіку (пакетів) на цільовий сервер – ні прямо, ні побічно – і основні ресурси, які ви можете зібрати, обмежені архівною інформацією (переважно застаріла інформація), незахищені файли, що залишилися на цільових серверах, і контент, що є на цільовому веб-сайті.

1.2.5.2 Напівпасивний збір

З технічної точки зору цей тип збору відправляє обмежений трафік на цільові сервери для отримання загальної інформації про них. Цей трафік намагається бути схожим на типовий Інтернет-трафік, щоб не привертати уваги до вашої розвідувальної діяльності. Таким чином, розвідник не проводить глибокого дослідження цільових онлайн-ресурсів, а проводите поверхове розслідування, не спричиняючи тривоги на боці цілі. Хоча цей тип збору вважається деяким анонімним, ціль може знати, що відбувається розвідка, якщо вона досліджує проблему (перевіривши журнали сервера або мережевого пристрою). Однак вони не повинні мати можливості віднести його до зловмисника.

1.2.5.3 Активний збір

У цьому типі аналітик безпосередньо взаємодіє із системою, щоб зібрати інформацію про неї. Ціль може дізнатися про процес розвідки, оскільки людина/організація, що збирає інформацію, буде використовувати передові методи

для збору технічних даних про цільову IT-інфраструктуру, таких як доступ до відкритих портів, сканування вразливостей (системи Windows без виправлень), сканування програм веб-сервера і багато іншого. Цей трафік буде виглядати як підозріла або зловмисна поведінка і залишить сліди в системі виявлення вторгнень (IDS) або системі запобігання вторгненням (IPS). Проведення атак соціальної інженерії на ціль також вважається видом активного збору інформації.

Із вищесказаного можна зрозуміти, що активний збір та напівпасивний збір – це типи збору інформації, які зазвичай використовуються рідко під час збору OSINT. Пасивний збір краще, тому що він може таємно збирати інформацію із загальнодоступних джерел, і у цьому суть OSINT.

2 СУТНІСТЬ ОСНОВНИХ ІНСТРУМЕНТІВ ТА МЕТОДІВ, ЩО ВИКОРИСТОВУЄ МЕХАНІЗМ РОЗВІДКИ ЗА ВІДКРИТИМИ ДЖЕРЕЛАМИ

2.1 Операційний цикл OSINT

Лише частина величезного обсягу OSINF, яка щодня поширюється та передається, може вважатися актуальною, своєчасною та корисною для OSINT-аналітика [19]. Визначення того, що є менш чи більш актуальним, потребує величезних зусиль, розподілених на весь спектр розвідувальних даних, від початкового збору до поширення результатів до їх отримання розробником політики. Перетворення інформації з необроблених даних включає кроки, необхідні для надання контексту для оцінки достовірності та надійності звіту.

Однак OSINT все ще потребує чіткої методології [20]. Існує кілька існуючих моделей для опису методології розвідки. Цикл розвідки ЦРУ описує цей процес як планування та напрям, збір, обробку, аналіз та виробництво, а також поширення. У Довіднику з досліджень інтелекту ці етапи описуються як збір, обробка, аналіз та виробництво, класифікація та розповсюдження.

Методологія OSINT фокусується на чотирьох ключових етапах: збір, обробка, використання та виробництво, як показано на рис. 2.1 (обробка та використання можуть відбуватися не повністю послідовно, а скоріше паралельно чи узгоджено). Простіше кажучи, ці етапи можна описати як одержання інформації, перевірку цієї інформації, визначення цінності інформації та надання інформації клієнтам. У підрозділах, що представлені нижче кожна з цих областей розбита на складові.

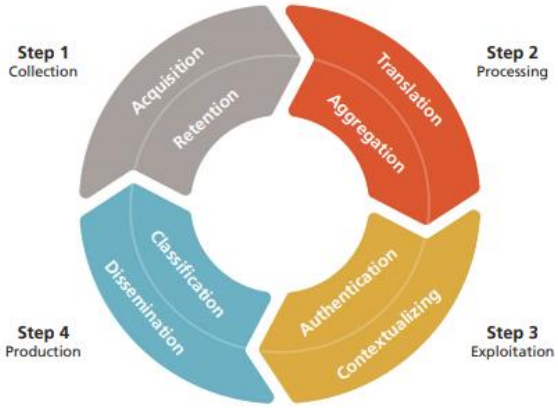

Рисунок 2.1 – Операційний цикл OSINT[5]

2.1.1 Збір інформації

Перший етап, збір інформації, включає виявлення потенційно корисної інформації та збереження цього матеріалу. На цьому етапі потрібен посібник – явний чи загальний – для збирачів інформації з відкритих джерел, щоб визначити види інформації, яку слід збирати, та визначити пріоритети зусиль зі збору, щоб відобразити вимоги розвідувальної спільноти[5]. Придбання – це фізичний чи електронний збір цієї інформації. Утримання – це володіння придбаним OSINF.

З усіх типів OSINF, що розглядаються у першому розділі, найлегше збирати контент новинних медіа. Для OSINT першого покоління фізичний збір даних ЗМІ представляв логістичні проблеми, які вимагали розосередження FBIS по кількох географічних точках для перехоплення передач. Збирання друкованих матеріалів залежало від присутності дипломатичного службовця або таємного колекціонера для фізичного придбання опублікованих матеріалів. Однак сьогодні, коли більша частина інформації ЗМІ доступна в Інтернеті, логістичні завдання перемістилися з обробки на управління інформацією. Збереження інформації засобів масової інформації досить просто. Обсяг такої інформації є керованим і інформація зазвичай надходить у стандартизованому текстовому форматі.

Сіру літературу, як і контент засобів, стає легше збирати з тих самих причин. Творці сірої літератури повільніше, ніж новинні ЗМІ, переходять до онлайн-контенту,

тому все ще бувають випадки, коли аналітику потрібно фізично отримувати інформацію в друкованому вигляді, особливо в країнах, що розвиваються, де використання Інтернету установами може бути не так широко поширене. Як і у випадку з контентом новин, отримати «сіру» літературу не так вже й складно.

Інформація в соціальних мережах, навпаки, представляє безліч унікальних проблем на етапі збору як для короткого, так і для повного контенту. По-перше, отримати повну картину необроблених даних може бути складно. На початковому етапі аналізу контенту соціальних мереж аналітика соціальних мереж була легко доступна, інколи ж навіть безкоштовна для використання. Одна компанія Topsy, наприклад, надала публічний доступ до повного індексу матеріалів Twitter з моменту заснування Twitter у 2006 році. Однак, оскільки аналітика соціальних мереж стала усталеною галуззю, такі платформи, як Topsy, були куплені та закриті більшими компаніями, які прагнуть монетизації цих платформ. Компанії-агрегатори соціальних мереж, які продають дані соціальних мереж, часто надають лише частину даних із платформи соціальних мереж або наборів даних лише за певний період часу. Крім того, ці провайдери також, як правило, зосереджуються на даних соціальних мереж з платформ, що базуються в США, насамперед Twitter і Facebook, хоча власні платформи є більш актуальними для деяких ключових інтересів розвідувальної спільноти. Крім того, навіть, якщо розвідувальна спільнота може отримати повний набір даних соціальних мереж, а не підмножину, дані не є репрезентативною вибіркою для населення. Демографічні групи нерівномірно використовують соціальні мережі, і в багатьох місцях, що становлять інтерес для розвідувальної спільноти, використання може сильно залежати від соціально-економічного класу.

Збір даних соціальних мереж також порушує юридичні питання, пов'язані із захистом громадян США, що є особливо актуальним для утримання. Таких проблем менше у «сірій» літературі і, як правило, їх немає у засобах масової інформації. Оскільки дані соціальних мереж можуть легко включати дані, що стосуються громадян США, розвідувальна спільнота повинна дотримуватися суворих процедур,

пов'язаних зі збором та зберіганням інформації. Ці процедури докладно описані у різних нормативних актах, у тому числі у Виконавчому указі 12333 та Директиві Міністерства оборони США 5240.01. Крім того, як повний, так і короткий контент у соціальних мережах динамічніший, ніж контент у новинах або маловідомій літературі. Стаття новин (за винятком виправлень), як правило, не є живим документом – якщо історія змінилася, буде згенеровано окрему нову статтю. Навпаки, тенденція обговорення може викликати інтерес та оновлення протягом кількох днів чи тижнів, а може тривати роками. Придбання та збереження контенту соціальних мереж, зокрема, має відбуватися в режимі реального часу та постійно, оскільки впливовий контент може бути розміщений та видалений протягом короткого періоду часу, якщо він провокує суперечки або розкриває конфіденційну інформацію – випадки, які можуть представляти особливий інтерес для розвідувальної спільноти. Нарешті, як повний, так і короткий контент у соціальних мережах дедалі частіше представлений у форматах, відмінних від тексту. Відеоролики YouTube є прикладом повного контенту соціальних мереж в іншому форматі, а короткі дані соціальних мереж у нетекстовому форматі включають зображення на таких платформах, як Flickr, та «живі» відео на таких платформах, як Facebook та Twitter.

2.1.2 Обробка інформації

Другий етап, обробка інформації, включає перевірку інформації та її придатність для використання[5]. Обробка може приймати різні форми, включаючи переклад вихідних матеріалів з мови оригіналу на англійську та перетворення відеоматеріалів або фотографій на корисну інформацію. Обробка в OSINT другого покоління є кардинальними змінами порівняно з обробкою в OSINT першого покоління як щодо змін існуючих методів, так і вимог нових методів. Багато завдань, які виконуються під час обробки, тепер можуть бути виконані легше і з меншими витратами за допомогою програмного забезпечення, включаючи професійні версії Google Translate. У той же час OSINT тепер має багато доступної інформації в менш

структурованому форматі, що значно ускладнює обробку. В методології обробки інформації при проведенні OSINT існує два компоненти обробки: трансляція та агрегація. Ці компоненти не обов'язково повинні з'являтися у заданій послідовності, хоча в деяких випадках один може допомагати іншому.

Обробка даних засобів масової інформації насамперед включає переклад джерел. Колись основним завданням FBIS був переклад, який радикально вплинув на швидкий розвиток машинного перекладу, принаймні, для мов, для яких задокументовано загальний синтаксис. Хоча лінгвісти OSC, як і раніше, відіграють важливу роль – надаючи нюанси культурного контексту іншомовному матеріалу, – тепер вони можуть зосередити свої зусилля на забезпеченні аналітичної цінності на етапі експлуатації. Машинний переклад найбільш ефективний для контенту засобів масової інформації, який використовує стандартну лексику і часто дотримується шаблонної структури. Сіра література зазвичай також дотримується стандартів професійного листа, які підходять для машинного перекладу, хоча складні та специфічні теми, охоплені сірою літературою, іноді потребують втручання людини. Інформація у соціальних мережах має переваги та недоліки для машинного перекладу. З одного боку, повідомлення в соціальних мережах зазвичай містять обмежену кількість символів – наприклад, Twitter обмежує користувачів до 140 символів. З іншого боку, повідомлення в соціальних мережах, швидше за все, будуть містити сленг, стенографію, смайлики або значки. Вони також можуть використовувати кілька мов і, ймовірно, найчастіше містять друкарські помилки. У той час як повний контент у соціальних мережах може містити достатньо інформації, щоб забезпечити якийсь несуперечливий запис, за яким можна зробити висновок про стиль або позицію автора, короткий контент у соціальних мережах з меншою ймовірністю надасть такий запис, якщо тільки основна частина матеріалу чи діяльність компілюється[4].

Агрегування, яке, як правило, не потрібне для даних засобів масової інформації та «сірої літератури», є важливим кроком для аналізу багатьох типів контенту соціальних мереж, особливо короткого контенту соціальних мереж. Агрегація може

також включати скорочення або інтеграцію при переведенні масиву даних у придатну для використання форму. Багато комерційних компаній надають послуги з агрегування даних, що позбавляє розвідувальну спільноту необхідності займатися безпосереднім збором даних. Хоча ці агрегатори даних можуть мінімізувати збір і обробку інформації, вони можуть не надавати дані з декількох платформ і не надавати повні зразки даних. Для розвідувальної спільноти також може бути складно точно знати, які дані були включені до набору даних, що ускладнює його здатність автентифікувати дані та поміщати їх у відповідний контекст.

2.1.3 Використання (експлуатація) інформації

Експлуатація прагне визначити, чи є інформація тим, на що вона претендує, і яка її цінність для розвідувальної спільноти[5]. Експлуатацію також іноді називають аналізом. Як зазначає колишній офіцер ЦРУ та вчений-розвідник Артур Хулник, однією з найсерйозніших проблем, пов'язаних з використанням продуктів OSINT, є величезний обсяг інформації, яка знаходиться у відкритому доступі, та ступінь надійності, притаманна цій інформації. Таким чином, багато часу при аналізі OSINT має бути витрачено на відділення надійної, "хорошої" розвідувальної інформації від "поганої" [21]. Аналітики повинні вміти «збирати, оцінювати та сортувати інформацію, знати обмеження та справлятися з ними, а також розуміти різних користувачів, потреби, завдання, структуру інформації, організацію, інститути та закони» [22]. Готовий продукт повинен містити аналітичні висновки на основі доступних джерел.

Експлуатація містить в собі три етапи: автентифікація, оцінка достовірності та контекстуалізація. Автентифікація прагне перевірити, чи є інформація тим, чим вона є. Для інформації, що надходить із інституційних джерел, це досить просто. Статті про «Нью-Йорк Таймс» з великою ймовірністю можуть бути свідомо та цілеспрямовано опубліковані «Нью-Йорк Таймс». Так само можна з високим ступенем впевненості припустити, що «сіра» література, опублікована на державних веб-сайтах, була

підготовлена і поширена урядом. Автентифікація контенту соціальних мереж набагато складніша. Користувачі можуть свідомо приховувати свою справжню особистість або надавати неправдиву інформацію про свою особистість. Це виходить за межі простого істинного імені користувача. Наприклад, людина може бути нечесною щодо свого місцезнаходження чи особистих якостей. Якщо розвідувальна спільнота намагається встановити атмосферу всередині країни, дуже важливо, щоб користувачі були корінними мешканцями цієї країни, а не членами діаспори. Автентифікація може бути потрібна одночасно з функціями агрегування даних, щоб гарантувати, що вибірка даних або композит не будуть помилково спотворені.

Оцінка правдоподібності, як і автентифікація, досить проста для традиційного медіа-контенту та сірої літератури, але надзвичайно складна для контенту соціальних мереж. Міра достовірності спрямована на визначення того, чи заслуговує на довіру інформація, тобто чи була вона надана без наміру спростувати або ввести в оману і чи має її джерело достовірний доступ до неї. «Нью-Йорк Таймс», наприклад, майже завжди публікує матеріали з певною метою – її зміст має бути точним, а джерела прозорими. Це може бути менш вірним для іноземних ЗМІ, особливо державних ЗМІ, які мають намір впливати на своє населення або інформувати його. Тим не менш, ймовірно, існує історія матеріалів з таких джерел, які можуть дати деяке уявлення про достовірність їхньої інформації.

На відміну від засобів масової інформації, соціальні мережі, як правило, не є інформацією з інших рук. Контент зазвичай надходить безпосередньо із джерела. Проте це завжди так, і оригінальність джерела може викликати підозри. Ретвіти, репости та боти – це приклади даних із соціальних мереж, які виявилися важливими для приховування намірів першоджерел. Навіть, якщо джерело надає інформацію про подію, свідком якої воно було, чи можна довіряти цьому звіту, враховуючи, що ми можемо мало знати про викриття, упередженість та досвід джерела? Це не означає, що люди завжди свідомо спотворюють факти. Одним із яскравих прикладів є поліцейські натільні камери, які, на думку деяких, проливають світло на зіткнення з

поліцією, але які важко інтерпретувати, а інтерпретація може бути схильна до когнітивних спотворень. не обов'язково розуміти самоцензуровану інформацію джерела. Деякі люди, наприклад, вважають за краще ніколи не згадувати та не показувати фотографії своїх дітей в Інтернеті; інші ніколи не обговорюють свою роботу в Інтернеті.

Контекстуалізація дозволяє аналітику з відкритих джерел передавати експертні знання у предметній галузі кінцевому споживачеві. Це може включати коментарі про джерело, що надають додаткову інформацію, наприклад, інформацію, що стосується достовірності. Контекстуалізація також може включати компіляцію декількох елементів OSINF з будь-якого результату в продукт, який дає більш повне уявлення про проблему[7].

2.1.4 Виробництво інформації (Production)

На заключному етапі виробництва інформація надається споживачеві у зручній для використання формі[5]. Цей споживач найчастіше буде аналітиком розвідувальних даних із усіх джерел, який може включити їх у багатоцільове виробництво розвідувальних даних. Проте продукт з відкритих джерел також може мати високий пріоритет або бути повним, щоб його можна було надати безпосередньо розробнику політики або іншому замовнику розвідувальних даних. Це схоже на інші дисципліни розвідки, де людська, сигнальна або геопросторова розвідки зазвичай включаються до аналітичного продукту з усіх джерел, але іноді надаються у необробленому вигляді безпосередньо клієнту розвідки.

Етап виробництва також включає присвоєння OSINT-продукту рівня класифікації. Хоча продукт можна отримати з OSINF, деталі збору, обробки та використання цієї інформації можуть вимагати підвищеного рівня класифікації. OSINF може задовольняти вимоги щодо секретної розвідувальної інформації, особливо у поєднанні з іншою інформацією [4]. Наприклад, інформація може бути отримана офіційними або секретними способами, коли викриття володіння

інформацією може поставити під загрозу її безперервний доступ. Виробники відкритих джерел можуть використовувати технологію секретної обробки та використання, яка виправдовує класифікацію інформації.

Поширення є також компонентом етапу виробництва. Аналіз із відкритих джерел найчастіше поширюється у вигляді письмового звіту. Однак продукти можуть також набувати форми усних брифінгів або графічних візуалізацій.

Середовище, що використовується для поширення, на жаль, часто є найпростішим, а не найефективнішим механізмом розповсюдження. Відео, аудіо або інтерактивна графіка часто можуть бути ефективнішими, ніж письмові звіти для передання конкретної інформації. Аналітики, які працюють з усіма джерелами інформації, зазвичай отримують свої аналітичні звіти з текстових баз даних, таких як Trident, WISE або Pathfinder.

2.2 Сутність основних інструментів OSINT

2.2.1 Інструменти для моніторингу та аналізу соціальних мереж та блогів

Моніторинг соціальних медіа – найважливіший етап для успішного розвитку бізнесу, просування Інтернету, конкурентної розвідки. За допомогою соціальних медіа можна дізнатися найповнішу інформацію про аудиторію товару чи послуги, її думку про роботу компанії[23].

Наведемо приклад кількох сервісів для ефективного моніторингу соціальних медіа, зосередивши увагу на найбільш популярних, корисних та доступних:

- Socialmention (www.socialmention.com) – платформа пошуку та аналізу інформації в соціальних мережах. Слоган – «Пошук та аналіз у соціальних мережах у реальному часі». Система шукає згадки у вибраних мережах або у всіх мережах одразу. Надає аналіз тональності згадок, пов'язані ключові слова, популярні джерела та багато іншого. Охоплення системи – понад 100 соціальних медіа, включаючи соціальні мережі, соціальні закладки, блоги, форуми та багато іншого.

- Hootsuite (www.hootsuite.com, www.seesmic.com) – Сервіс моніторингу соціальних медіа. Слоган сервісу – «Соціальні мережі – це ваша суперсила». Підтримує моніторинг таких ресурсів, як Twitter, Facebook, LinkedIn, Chatter, Ping.fm. Є програми як для веб, так і для персонального комп'ютера, iPhone, Android, Windows Mobile. Сервіс HootSuite є сертифікованим партнером Twitter. Забезпечує постінг (posting) за розкладом, можливість відстежувати повідомлення за ключовими словами та згадками. Система HootSuite також надає повноцінну інтеграцію із Facebook.
- YouScan (www.youscan.io) – система моніторингу російськомовних соціальних медіа. Слоган – «Використовуйте силу соціальних медіа для ухвалення вірних рішень». Система YouScan відстежує згадки в блогах, форумах, соціальних мережах (Facebook, ВКонтакте), Twitter, YouTube, та надає результати моніторингу в аналітичному інтерфейсі з функціями одночасної роботи кількох співробітників. Надає звіти щодо кількості повідомлень зі згадуваннями ключових слів, авторів, джерел, тональності.
- Search4faces (https://search4faces.com/) – сервіс пошуку людей в інтернеті по фотографії. Завдяки технології нейронних мереж та машинному навчанню дана платформа допомагає знайти потрібну людину або дуже на неї схожу протягом кількох секунд. Результатом є посилання на профіль знайденої людини у соціальній мережі «Вконтакте», «Одноклассники», «Tiktok» та «ClubHouse», а в найближчому майбутньому також інших.

На даний момент дана платформа має частково або повністю зібрані 4 бази даних: аватари соціальної мережі «Вконтакте» та головні фотографії соціальної мережі «Одноклассники», фотографії профілю соціальної мережі Вконтакте, аватари користувачів «Tiktok», аватари користувачів «ClubHouse».

- Skopenow (https://www.skopenow.com/) – провідний постачальник комплексних рішень для аналізу загроз та OSINT. Продукти компанії для розслідувань використовують понад 1000 клієнтів, у тому числі 20% компаній зі списку Fortune 500 та численні державні установи. Skopenow підвищує ефективність розслідувань,

аналізуючи найширший і найглибший загальнодоступний контент про людей, компанії та події, виявляючи найбільш актуальні та точні результати. Skopenow, заснована в 2016 році і має штаб-квартиру в Нью-Йорку, є технологічною компанією, що швидко росте, підтримується венчурним капіталом. Skopenow працює зі страховими компаніями, юридичними фірмами, глобальними службами безпеки, державними установами, медіа-компаніями, ліцензованими приватними детективами, дослідниками, правоохоронними органами тощо.

- Google Account Finder (EPIEOS) (https://epieos.com/) – дана платформа шукає зображення профілю та доступні огляди карт Google і фотографії, пов'язані з адресою Gmail. Також перевіряє номери телефонів та адреси електронної пошти у соціальних мережах.

2.2.2 Інструменти для аналізу глибинного веб-простору

Традиційні пошукові системи прагнуть звузити простір глибинного Інтернету, поступово захоплюючи такі ніші, як блоги, наукові сайти, інформаційні агентства. Так, як допоміжні сервіси для пошуку по глибинному веб від Google можна рекомендувати: Google Book Search (books.google.com) – пошук книг, Google Scholar (scholar.google.com) – пошук наукових публікацій, Google Code Search (code. google.com) – пошук програмного коду[24].

Система Goldfire Research від компанії Invention Machine Corp. (inventionmachine.com) дозволяє обробляти контент глибинного веб, розміщений більш, ніж на 2000 сайтів урядових, академічних, дослідницьких та комерційних організацій США. Система Goldfire Research має інформацію про механізми доступу до баз даних глибинного веб і автоматично генерує запити до них.

Існуючі засоби аналізу та просування веб-ресурсів дозволяють по-новому підійти до оцінки співвідношення обсягів видимого та глибинного Інтернету. Так на веб-сайті www.cy-pr.com наводиться інформація про реальну кількість документів на досліджуваному веб-сайті та про кількість документів, заіндексованих різними

пошуковими системами, у тому числі Google. Отримавши репрезентабельну вибірку сайтів, наприклад, за рейтингом top100, можна отримати оцінку співвідношення видимої та глибинної частини веб-простору.

Як показують розрахунки, обсяг інформації, що опинилася в глибинній частини веб-простору, перевищує обсяг інформації з видимої частини приблизно в 3-5 разів. Виявляється, за рідкісним винятком, що чим більший ресурс, тим більша його частина відноситься до глибинного Інтернету. У цьому сенсі невеликі веб-ресурси виграють у доступності. Так як більша частка новинних документів опиняється в глибинному Інтернеті, то для завдань бізнес-аналітики потрібні спеціальні послуги доступу до такої інформації. Саме такий сервіс надають служби інтеграції контенту новин – архіви мережевих ЗМІ. Бізнес-аналітики активно використовують найбільші архіви інформації з відкритих джерел "Інтегрум" (integrum.ru) та InfoStream (www.infostream.ua). Саме використання відкритих джерел дозволяє конкурентній розвідці діяти в рамках правового поля, але при цьому мати високу ефективність.

Можна констатувати, що чим швидше зростає веб-простір, тим гірше він охоплюється традиційними каталогами та пошуковими машинами. Через зростання кількості веб-сайтів і порталів, що використовують бази даних, динамічних систем керування контентом, появи нових версій форматів представлення інформації, глибинний веб зростає дуже інтенсивно. З одного боку, Інтернет як величезне сховище збільшує обсяг інформації, доступної «в принципі», але з іншого боку – зростає інформаційний хаос, ентропія мережевого інформаційного простору збільшується. Дедалі менша частина інформаційних ресурсів стає доступною користувачам реально.

Провідні пошукові системи, як і раніше, намагаються знайти технічні можливості для індексації вмісту баз даних та доступу до закритих веб-сайтів, проте їх завдання об'єктивно розходяться із завданнями бізнес-аналітиків – орієнтація традиційних пошукових служб на масовий сервіс у цьому випадку виправдана. Таким чином, ніша для систем пошуку в глибинному Інтернеті стає все ширшою.

2.2.3 Інструменти для аналізу онлайн-карт, а також геолокації

Технологічні досягнення та революційні інновації сповістили початок нової ери у спостереженні за Землею, зробивши супутники більш актуальним інструментом для бізнесу та суспільних інтересів. Це не тільки змінює спосіб вивчення Землі, а й виводить журналістику на нові орбіти.

- GeoCreepy (https://www.geocreepy.com/) – OSINT-інструмент геолокації. Збирає інформацію, пов’язану з геолокацією, з онлайн-джерел та різних соціальних мереж і дозволяє відображати на карті, фільтрувати пошук на основі точного місцезнаходження та/або дати, експортувати у форматі csv або kml для подальшого аналізу в Картах Google. Користувач може ввести ім'я користувача Twitter або Flickr, і інструмент аналізує повідомлення, щоб визначити місцезнаходження та час.
- Digital Globe (www.digitalglobe.com, https://www.maxar.com/) – американська компанія, комерційний оператор кількох цивільних супутників дистанційного зондування Землі, великий постачальник результатів супутникової зйомки та геопросторових даних (зокрема для Google Maps/Earth та Virtual Earth). У 2015 році дана компанія надала Associated Press зображення (рис. 2.2), що показує як два траулери завантажують на комерційний вантажний корабель морепродукти, видобуті рабською силою.
- RoundShot (https://www.roundshot.com) – платформа, що надає послуги по перегляду камер, які обертаються на 360°, з вибраних регіонів по всьому світу в режимі онлайн. Послуги даного сервісу є безкоштовними.
- NGA GEOINT (https://github.com/ngageoint) – офіційний репозиторій інструментів, пов'язаних з картами Національного агентства геопросторової розвідки на GitHub. Національне агентство геопросторової розвідки (NGA) надає геопросторову розвідку світового класу, яка надає вирішальну перевагу політикам, військовослужбовцям, спеціалістам розвідки та службам швидкого реагування.

- NGA є унікальним поєднанням розвідувального агентства та агентства бойової підтримки. Це світовий лідер у своєчасній, актуальній, точній та дієвій GEOINT. NGA дозволяє розвідувальній спільноті США та Міністерству оборони (DOD) виконувати пріоритети національної безпеки президента для захисту нації. NGA також передбачає майбутні потреби своїх партнерів та вдосконалює дисципліну GEOINT для їхнього задоволення.

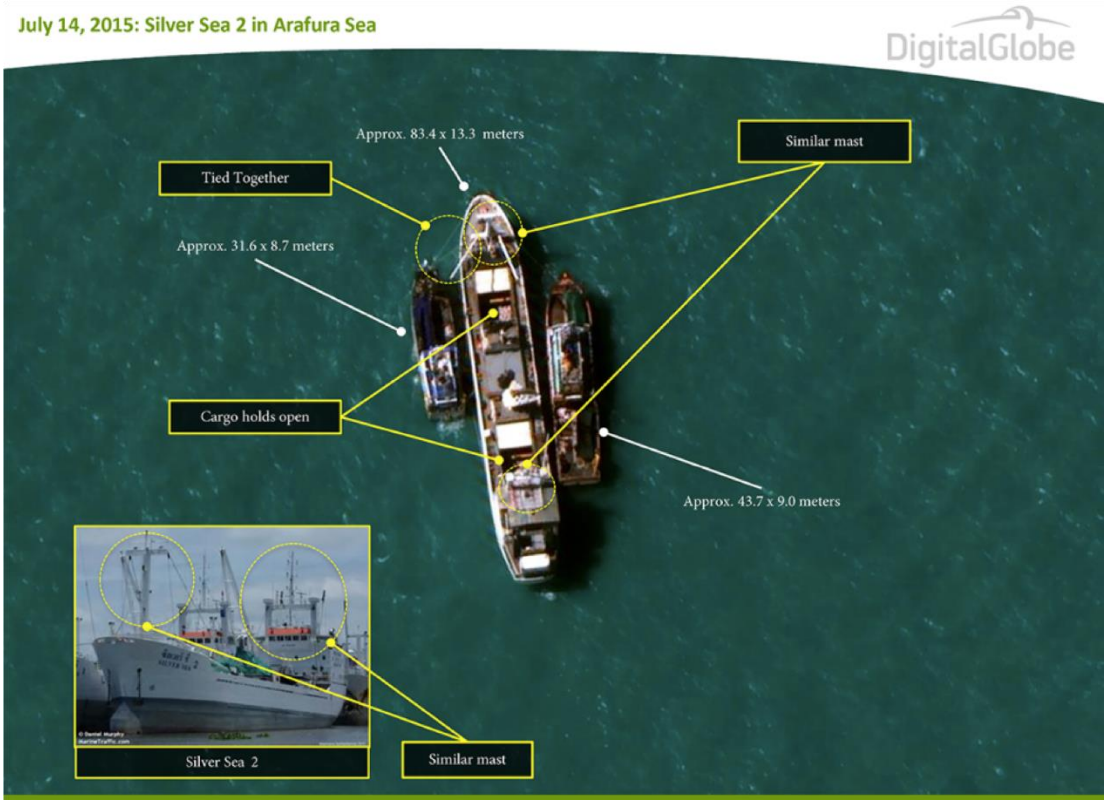

Рисунок 2.2 – Зображення Digital Globe, що було надане до Associated Press у 2015 році

- Resource Watch (https://resourcewatch.org/) – Некомерційна платформа, що все ще перебуває в стадії бета-тестування, що надає сотні датасетів про стан ресурсів та жителів планети. Спонсорується Інститутом світових ресурсів та іншими організаціями. Resource Watch містить сотні наборів даних у одному місці про стан ресурсів планети та громадян. Користувачі можуть візуалізувати виклики, з якими стикаються люди та планета, від зміни клімату до бідності, від ризику води до нестабільності держави, від забруднення повітря до міграції людей тощо.

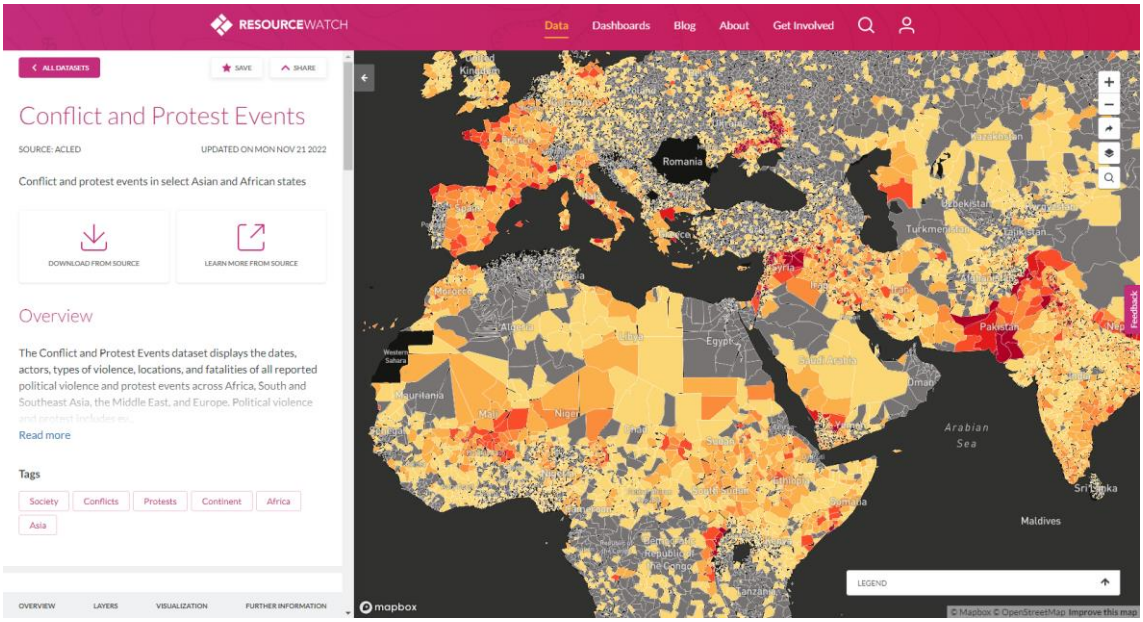

Рисунок 2.3 – Мапа конфліктів і протестних дій

2.2.4 Інші найвідоміші та найсучасніші інструменти OSINT

Використання відповідних інструментів OSINT може допомогти різним компаніям підвищити рівень кібербезпеки. Інструменти розвідки допомагають збільшити обсяг цінної інформації, яку отримує аналітик.

Хоча OSINT-інструменти часто використовуються в різних цілях, вони нерідко орієнтуються на кілька конкретних областей[25]:

– Знаходження нових загальнодоступних активів. Загальнодоступні активи скрізь, але найнебезпечніші активи – це ті, що OSINT-розвідник не бачить. Їх можна назвати невивченими інфраструктурними активами. Ось чому створення повної карти всієї онлайн-інфраструктури компанії та інших типів використовуваних компанією сервісів – це перше, чим займатимуться під час процесу збирання інформації як люди з добрими намірами, так і зловмисники.

Це може забезпечити якісний і надійний захист, або призвести до того, що організація стане жертвою різноманітних кібератак.

– Виявлення критично важливих зовнішніх даних компанії, в якій працює OSINT аналітик. Іноді дані не зберігаються на загальнодоступних веб-ресурсах

компанії, іноді вони розташовуються будь-де, і часто це відбувається, коли компанія працює з великою кількістю сторонніх SaaS-сервісів, а також SaaS-сервісів, які вже сторонні для них. До інших поширених проблем належать злиття та поглинання компаній. Якщо їх згаяти, то це може призвести до того, що вектор зовнішніх атак буде спрямований на головну компанію. І OSINT може виявитися дуже корисним засобом під час проведення аудитів кібербезпеки, спрямованих на виявлення злиття та поглинання компаній.

– Угрупування критично важливих даних у корисні плани. Після виявлення найбільш корисних даних з усіх джерел за допомогою відповідних OSINT-інструментів саме час зіставити та згрупувати всі ці дані, а потім перетворити їх на функціональні плани. Що робити із цими відкритими портами? Хто відповідає за оновлення цього застарілого програмного забезпечення? Коли видаляти ці неактуальні записи DNS? Після того, як аналітики компанії згрупують в єдиний фундамент усі свої важливі дані, всі ці та багато інших питань допоможуть вам при створенні найбільш коректних та практичних планів.

Для виконання всіх цих аспектів існує безліч OSINT інструментів, короткий перелік яких буде представлено нижче:

1) OSINT Framework. OSINT Framework – це не конкретне програмне забезпечення, а набір інструментів, які спрощують завдання OSINT. OSINT Framework надає інформацію у вигляді інтерактивної інтелект-карти на базі Інтернету, яка візуально впорядковує інформацію, що показана на рис. 2.4.

– Даний фреймворк популярний серед пентестерів та фахівців з інформаційної безпеки. За допомогою цієї платформи можна переглядати різні інструменти OSINT, які фільтруються за категоріями.

– Наприклад, деякі категорії – це ім'я користувача, адреса електронної пошти, геолокація/карти, темна мережа, пошукові системи, транспорт, загальнодоступні записи та багато іншого.

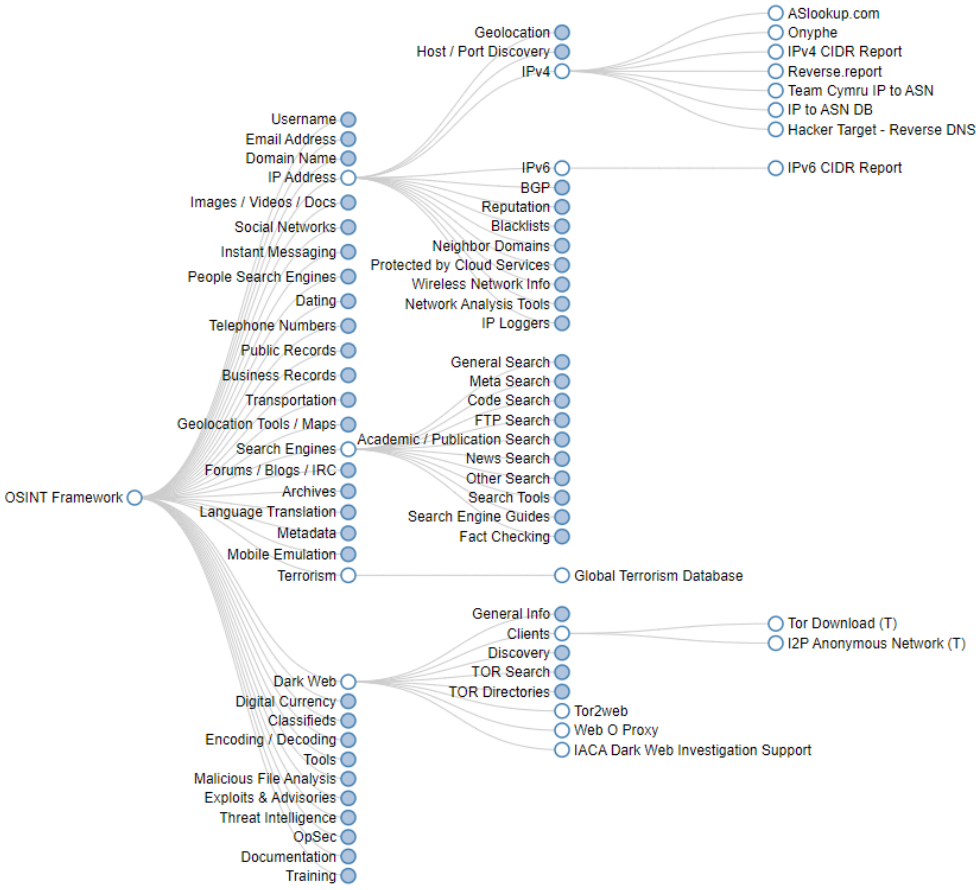

Рисунок 2.4 – Практичне використання OSINT Framework

2) HaveIbeenPwned. Дозволяє перевірити, чи піддавався обліковий запис злому. Якщо користувач підозрює, що його обліковий запис був зламаний, або хоче перевірити, чи не зламали зовнішні облікові записи, цей інструмент підійде. За допомогою даного інструменту можна виявляти кіберзагрози з великої кількості джерел, таких як Gmail, Hotmail, облікових записів Yahoo, а також LastFM, Kickstarter, WordPress.com, LinkedIn та багато інших популярних сайтів.

3) Censys. Пошукова система, яка використовується для отримання найбільш актуальної та точної інформації про будь-який підключений до Інтернету пристрій, наприклад, сервери або доменні імена.

Користувач даного інструменту зможе знаходити всі докладні відомості про географічне розташування та технічні відомості про 80-й та 443-й порти, що

використовуються на будь-якому сервері, а також вміст тіла HTTP- або HTTPS-відповіді на GET-запит до цільового сайту, TLS-рукостискання браузера Chrome.

4) Google Dorks. Інструментом Google Dorks важко назвати, скоріше це метод запиту даних, що включає розширені та розумні аргументи пошуку в Google. Сайти автоматично індексуються під час сканування ботами Google, якщо сайти з конфіденційною інформацією спеціально не блокують ботів, їх вміст відображається як результати пошуку для конкретних запитів Google. Так, користувачі можуть детально перевірити сервер і знайти потрібну інформацію. Цей метод використовує простий синтаксис Google для фільтрації даних пошукової системи. Служить корисним засобом OSINT для початківців. Найчастіше Google Dorks використовують у своїй роботі журналісти, слідчі, інженери з безпеки.

5) Maltego. Дане програмне забезпечення розроблене компанією Paterva і входить до складу дистрибутива Kali Linux.

Використовуючи Maltego, ви можете запускати розвідувальні перевірки, спрямовані на певні цілі.

Одна з найбільш ефективних і важливих можливостей цього програмного забезпечення – так звані «перетворення» (transforms). Перетворення в деяких випадках доступні безкоштовно, але в інших пропонують лише платні версії. Вони допоможуть вам проводити різноманітні тести та інтегрувати дані зі сторонніми додатками.

Щоб використовувати Maltego, потрібно створити безкоштовний обліковий запис на сайті розробників. Потім користувач зможе запустити нову машину або застосувати перетворення до цілі вже існуючої. Після того, як аналітик обере свої перетворення, програма Maltego почне виконувати всі перетворення на своїх серверах.

6) Recon-ng. Ще один відмінний інструмент, який вбудований у дистрибутив Kali Linux і використовується для виконання швидкої та ретельної інформаційної розвідки з віддалених цілей.

Цей програмний каркас для інформаційної розвідки в Інтернеті написаний на Python і включає численні модулі, зручні функції та інтерактивну довідку.

Простий інтерфейс командного рядка дозволить користувачеві виконувати звичні операції, наприклад, взаємодія з базою даних, виконання веб-запитів, управління ключами API або приведення отриманих даних до стандартного вигляду. Інструмент включає корисні для розвідки модулі, такі як google_site_web і bing_domain_web, які можна використовувати для пошуку цінної інформації про домени, що вас цікавлять.

7) theHarvester. Альтернативний інструмент для вилучення цінної інформації про будь-які імена піддоменів, віртуальні хости, відкриті порти та адреси електронної пошти будь-якої компанії або сайту.

Це особливо корисно на початкових етапах тестування на проникнення у вашу власну локальну мережу або сторонні мережі, до яких у вас є доступ. Як і згадані раніше інструменти, TheHarvester включений до складу дистрибутива Kali Linux.

TheHarvester використовує численні ресурси для отримання даних, такі як сервери криптографічних PGP-ключів, Bing, Baidu, Yahoo, пошукова система Google, а також соціальні мережі LinkedIn, Twitter і Google Plus.

Також цей інструмент можна використовувати для проведення активних тестів на проникнення, таких як атака на основі прямого перебору доменних імен за словником, пошук у rDNS та розширення доменів верхнього рівня за допомогою прямого перебору за словником.

8) Shodan. Засіб для спостереження за безпекою мережі та пошукова система, орієнтована на прихований («глибинний») веб та Інтернет речей. Цей інструмент був створений Джоном Матерлі у 2009 році для відстеження загальнодоступних комп'ютерів у будь-якій мережі.

Часто його називають "пошуковою системою для хакерів", оскільки він дозволяє виявляти та досліджувати пристрої будь-якого типу, підключені до мережі, такі як сервери, маршрутизатори, веб-камери та багато інших.

Shodan схожий на Google, але замість показу красивих зображень та сайтів з багатим інформаційним наповненням він покаже вам відомості, які будуть цікаві дослідникам інформаційної безпеки, такі як банери за цільовими серверами, що працюють за протоколами SSH, FTP, SNMP, Telnet, RTSP, IMAP та HTTP, а також загальнодоступна інформація про сервери. Під банерами в Shodan розуміється текстова інформація, що описує служби та пристрої. Результати, що виводяться, будуть сортуватися по країні, операційній системі, мережі та портах.

Користувачі Shodan можуть отримувати доступ не лише до серверів, веб-камер та маршрутизаторів. Цей інструмент можна використовувати для сканування майже будь-якого пристрою, підключеного до Інтернету, включаючи системи керування світлофорами, системи опалення будинку, панелі керування аквапарком, системи водопостачання, атомні електростанції та багато іншого.

9) Nmap. Один з найпопулярніших інструментів для проведення аудиту безпеки. Це безкоштовна утиліта з відкритим вихідним кодом, призначена для проведення аудиту безпеки та дослідження мереж на локальних та віддалених хостах.

Деякі з основних можливостей:

– Виявлення хостів: Nmap надає можливість пошуку хостів усередині будь-якої мережі, які мають певні відкриті порти або які можуть надсилати ICMP- або TCP-пакети.

– Виявлення інформації про IP і DNS: включає тип пристрою, Mac-адреси і навіть зворотні доменні імена.

– Виявлення портів: Nmap може виявляти будь-який відкритий порт у цільовій мережі та дозволяє вам дізнатися про всілякі служби, що працюють на ньому.

– Виявлення операційних систем: інструмент надає виявлення повної версії операційної системи та апаратних характеристик будь-якого підключеного хоста.

– Виявлення версії: Nmap також може отримувати ім'я програми та номер версії.

10) OpenVAS. Open Vulnerability Assessment System – програмний каркас, який включає спеціальні служби та інструменти для фахівців у галузі інформаційної безпеки.

Це сканер вразливостей та інструмент управління безпекою з відкритим вихідним кодом, який був розроблений після того, як компанія-розробник програми Nessus закрила її вихідний код та зробила його пропрієтарним. Після цього початкові розробники сканера вразливостей Nessus вирішили зробити форк оригінального проекту та створити OpenVAS.

Основний інструмент у OpenVAS – це OpenVAS Scanner, ефективний засіб, який проводить усі тести на наявність мережевих вразливостей у цільовому комп'ютері.

Ще один основний компонент – OpenVAS Manager. Це, практично, є рішенням для управління вразливістю, яке дозволяє поміщати просканові дані в базу даних SQLite, тому будь-який користувач зможе легко і нетривіально фільтрувати та впорядковувати результати сканування, а також виконувати пошук по них.

11) Metagoofil. Інструмент інформаційної розвідки, що дозволяє дослідникам безпеки, менеджерам з інформаційних технологій вилучати метадані з різних типів файлів, таких як: doc, docx, PDF, xls, xlsx, ppt, pptx.

За допомогою даного інструменту виконується ретельний, глибокий пошук у пошукових системах, наприклад у Google, зосереджуючись на файлах цих типів. Після виявлення такого файлу програма завантажить його у ваше локальне сховище і перейде до вилучення всіх цінних даних.

Після завершення вилучення даних ви побачите повний звіт з іменами користувачів, банерами ПЗ, версіями програм, іменами хостів та багато іншого. Цей звіт є цінним ресурсом для етапу інформаційної розвідки.

Крім безлічі інших корисних можливостей, Metagoofil також включає кілька параметрів запуску, що дозволяють фільтрувати типи шуканих файлів, деталізувати результати і налаштовувати їх висновок.

12) Aircrack-ng (http://www.aircrackng.org) – обов'язковий інструментарій у проектах тестування на проникнення мереж Wi-Fi, хоча не завжди є потреба використовувати всі інструменти, включені до нього. Інструментарій в першу чергу розроблений для перехоплення трафіку, що передається через бездротові мережі, аудиту WEP і WPA/WPA2-PSK ключів шифрування (перевірка стійкості), а також проведення пентестингу бездротових мереж. Інструмент дозволяє оцінювати безпеку мережі Wi-Fi. За допомогою утиліти можна перевірити ступінь надійності пароля мережі, отримати доступ до мережі сусідів.

13) Exiftool. У той час як багато інструментів OSINT фокусуються на даних, знайдених у загальнодоступних файлах, таких як файли форматів PDF, DOC, HTML, SQL і так далі, існують інструменти, які розроблені спеціально для вилучення критично важливих даних із загальнодоступних зображень, відеозаписів та аудіозаписів. Exiftool читає, записує та витягує метадані з файлів наступних типів: EXIF, IPTC, GPS, XMP, JFIF і багато інших.

Також цей інструмент підтримує безліч камер, таких як Canon, Casio, FujiFilm, Kodak, Sony і багато інших. Крім того, він доступний на різних платформах, включаючи Linux, Windows та MacOS.

3 СУТНІСТЬ ГОЛОВНИХ РЕСУРСІВ, ЯКІ ВИКОРИСТОВУЮТЬСЯ ДЛЯ ЗБОРУ ІНФОРМАЦІЇ У РАМКАХ OSINT

3.1 Сутність ресурсів OSINT

Значимість розвідки за відкритими джерелами була відзначена ще у 1966 році президентом США Ліндоном Джонсоном, коли була виголошена його промова на церемонії складання присяги директором ЦРУ Річардом М. Хелмсом.

На усталену помилкову думку, вся корисна розвідувальна інформація видобувається із секретних джерел агентурним чи оперативним шляхом – насправді це не так. Відоме визнання адмірала Захаріаса – заступника начальника розвідки Військово-морських сил США у роки Другої світової війни спростовує це. Так, за його оцінкою, 95% інформації розвідка військово-морських сил черпала з відкритих джерел, 4% – з офіційних, і лише 1% – з конфіденційних джерел. Часто саме цей один відсоток є тією золотою недостатньою ланкою, яка дозволяє скласти цілісну картину розрізненої мозаїки всіх розвідданих. І, якщо таке співвідношення справедливе для військової розвідки, то тим більше воно буде правильним для розвідки відкритими джерелами для будь-яких інших сфер.

У той самий час, аналіз розсекреченого звіту ЦРУ за 1987 рік «Enterprise-Level Computing in Soviet Economy» (SOV C87-10043) дає уявлення у тому, який колосальний обсяг даних потрібно обробляти аналітикам. Для складання звіту постійно протягом року сканувалося 347 відкритих джерел; для створення зведення обсягом одну сторінку щодня оброблявся інформаційний масив обсягом приблизно 7 млн. слів. Загальновідомо, що основна відмінність розвідки з відкритих джерел від промислового шпигунства – це легітимність та дотримання етичних норм. Тут це положення доведено до абсолюту – виключно всі джерела інформації в цьому випадку доступні та легальні.

Розвідувальна інформація може бути отримана з офіційних джерел, інших відкритих джерел, ЗМІ, оголошень, реклами, внутрішньофірмових, банківських, урядових звітів, баз даних, від експертів шляхом добування (збору), аналізу або спеціальної обробки даних, текстів. Щоправда, при цьому кількість різнорідних відомостей, які необхідно переробити, щоб отримати крихти знань величезна, а тому нині розвідка з відкритих джерел немислима без використання спеціалізованих інформаційних технологій.

Сьогоднішній розвиток інформаційних технологій зробив комп'ютерну розвідку з відкритих джерел доступною навіть для відносно невеликих компаній, сьогодні вона поширена на всіх рівнях економіки.

В інформаційно-аналітичній роботі важливе значення має можливість доступу до джерел даних, інформації та знань. При цьому головною проблемою є знаходження змістовних та надійних джерел із усіх загальнодоступних. Коли такі джерела знайдені, включаються механізми перетворення даних на знання, для чого застосовуються відповідні технології. Під даними зазвичай розуміють «сирі», необроблені відомості, що ґрунтуються на фактах. Це можуть бути статистичні дані, факти з біографій ключових осіб або, наприклад, відомості про звітність окремих компаній. Інформацією є вже певним чином оброблені і проаналізовані дані. Кінцевим інформаційним продуктом будь-якої аналітичної роботи є знання – синтезовані висновки, рекомендації до прийняття рішень.

Інформація, як було зазначено вище, може бути отримана з офіційних, відкритих джерел, ЗМІ, оголошень, реклами, фірмових, банківських, урядових звітів, баз даних від експертів шляхом аналізу або спеціальної обробки даних, текстів.

Нижче у таблиці 3.1, а також далі наведено докладний перелік видів інформаційних джерел, які найчастіше використовуються при розвідці з відкритих джерел.

Таблиця 3.1 – Перелік видів інформаційних джерел, що використовуються під час проведення OSINT

1	Прес-релізи компаній, офіційні заяви від імені компаній про нові технології, нові напрямки, угоди, перспективи.
2	Інтерв'ю співробітників компаній, відповідні матеріали у ЗМІ.
3	Висловлювання співробітників компаній на форумах, блогах, приватних бесідах.
4	Тендери, закупівлі.
5	Патенти, авторські свідоцтва компанії та її співробітників.
6	Розробки компанії: провідні розробки, що фінансуються, якими компанія цікавиться.
7	Активність компанії на ринку злиття та поглинання (M&A).
8	Вакансії компанії (відкриваються, що закриваються), повідомлення про активний пошук співробітників, вимоги до вакансій, умови.
9	Курси підвищення кваліфікації, навчання персоналу – вказівка на пріоритети у розвитку компанії.
10	Подяки та нагороди компанії та її співробітників.
11	Участь у заходах (виставки, конференції, круглі столи, презентації).
12	Участь у організаціях (союзи, асоціації, конфедерації тощо).

1) Прес-релізи компаній, офіційні заяви від імені компаній про нові технології, нові напрямки, угоди, перспективи. Такі прес-релізи створюються компаніями для власної популяризації, привернення уваги потенційних клієнтів, інвесторів, які шукають вигідні варіанти вкладення власних коштів. Часто в таких заявах є інформація про наміри, заплановані події. Прес-релізи доступні на веб-сайтах компаній, PR-службах, на загальних та профільних спеціалізованих майданчиках для розміщення прес-релізів.

2) Інтерв'ю співробітників компаній, що відповідають матеріалам у ЗМІ. В інтерв'ю інтерес становлять плани компаній. При цьому з боку служби розвідки з

відкритих джерел допускається ініціювання інтерв'ю когось із співробітників об'єкта інтересу.

3) Висловлювання співробітників компаній на форумах, у блогах, у приватних бесідах. При цьому можуть виявлятися плани компаній, кадрова політика, атмосфера в колективі тощо. Джерела інформації: 1) інтернет-ресурси (спеціалізовані форуми, блоги співробітників), блоги експертів, групи у соціальних мережах; 2) виставки, конференції, курси підвищення кваліфікації, професійні заходи.

4) Тендери, закупівлі. Предмети закупівель, обладнання, виконавці. Джерела інформації: 1) інтернет-ресурси (сайти компаній, торгові майданчики, профільні форуми); 2) партнери досліджуваної компанії, ті, хто брав участь у їх тендерах, у клієнтів та постачальників.

5) Патенти, авторські свідоцтва компанії та її співробітників. Корисними даними для розвідки з відкритих джерел є їх зміст, спрямованість, списки співавторів. Інформація розміщується на відповідних сайтах. Для України: https://ukrpatent.org/; Google Patents: https://patents.google.com/; Євразійське патентне відомство: www.eapo.org. Патентування можливе у будь-якій країні, кращі варіанти – країна реєстрації організації, країна ведення бізнесу, крім того США, Євросоюз, Японія та Китай.

6) Розробки компанії: провідні, фінансовані, розробки, якими компанія цікавиться. Спостереженню підлягають спроби компанії проводити дослідження: закупівля специфічного устаткування, способи фахівців працювати, переговори, відвідування відповідних організацій тощо.

7) Активність компанії на ринку злиття та поглинання (M&A). Інформація, які організації поглинаються, планують поглинути чи ведуть переговори про поглинання. Інформацію можна отримати в Антимонопольному комітеті (АМК) України, аналогічних відомствах інших країн, за повідомленнями новин на веб-ресурсах, присвячених M&A.

8) Вакансії компанії (які відкриваються, які закриваються), повідомлення про активний пошук співробітників, вимоги до вакансій, умови. Джерело інформації: веб-сайт компанії, сайти з пошуку роботи та сайти агентств, з якими компанія співпрацює.

9) Курси підвищення кваліфікації, навчання персоналу – вказівка на пріоритети у розвитку компанії. Інтерес представляє те, чому навчають, яких фахівців запрошують на навчання, які вимоги висувають при залученні учнів, які навчання, скільки персоналу навчається.

10) Подяки та нагороди компанії та її співробітників.

11) Участь у заходах (виставки, конференції, круглі столи, презентації). З'ясування, у яких заходах беруть участь компанії, їх спрямованість, коло учасників.

12) Участь в організаціях (союзи, асоціації, конфедерації тощо) – інформація про те, в яких об'єднаннях бере участь компанія, як бере активну участь, що отримує від участі, на що розраховує, як використовує.

Інформація характеризується якісними, кількісними та ціннісними показниками. До якісних характеристик зазвичай відносять: достовірність, об'єктивність та однозначність інформації. До кількісних характеристик – її повноту (відсутність нез'ясованих прогалин) та релевантність (ступінь відповідності суті поставлених питань та завдань). Ціннісними характеристиками є вартість та актуальність інформації.

Діяльність OSINT заснована на використанні лише легітимних джерел інформації, яких цілком достатньо для ухвалення управлінських рішень у різних сферах, необхідно лише провести деяку інформаційно-аналітичну обробку наявних відкритих даних. Серед таких джерел інформації можна назвати дані статистики, матеріали з веб-сайтів, соціальних мереж, ЗМІ, галузевих звітів тощо.

Багато служб розвідки з відкритих джерел не завжди можуть відокремити нелегітимну частину інформації від легальної, а замовник, як правило, цікавиться кінцевими результатами, джерела для нього виступають лише як підтвердження,

проміжні дані. Разом з тим солідні замовники самі зацікавлені в тому, щоб інформація видобувалася законними засобами, щоб аналітичний звіт був легальним [24].

У технології OSINT в останні десятиліття з'явилося і розвинулося до небачених раніше масштабів нове інформаційне джерело – веб-простір мережі Інтернет. Сьогодні за оцінками експертів, Інтернет за кількістю інформації знаходиться на першому місці, випереджаючи ЗМІ, галузеві видання та новини, котрі одержують від колег, спеціальні огляди, закриті бази даних. При цьому у відкритих джерелах та спеціалізованих базах даних, доступних в Інтернет, міститься більша частина інформації, необхідної для проведення розвідки з відкритих джерел, проте залишається відкритим питання її знаходження та ефективного використання. Останні дослідження інформаційного веб-простору показали, що доступний через традиційні інформаційно-пошукові системи трильйон веб-сторінок – це лише «поверхова видима частина айсберга». Близько 40% всієї інформації в Інтернеті є безкоштовною. Навігацію по даному інформаційному простору забезпечують понад мільйон пошукових систем та каталогів, але й вони охоплюють лише малу частину інформаційних ресурсів. Прихованих і невидимих (deep, invisible) ресурсів мережі Інтернет значно більше – це, насамперед сторінки, що динамічно генеруються, файли різноманітних форматів, інформація з численних баз даних. До "прихованого" Інтернету можна віднести і такі мережі, як BitTorrent, DirectConnect, EMule, Napster та ін.

Сьогодні для OSINT основними джерелами інформації є Інтернет, преса, а також відкриті бази даних. Дуже популярні серед фахівців з розвідки відкритих джерел бази даних державних та статистичних органів, торгово-промислових палат, органів приватизації тощо. Велику користь приносять і доступні окремі бази даних інших органів влади. Останнім часом дедалі популярніші бази даних з урахуванням архівів ЗМІ, зокрема і мережевих.

Традиційно розвідка по відкритих джерелах спирається на такі джерела інформації, як опубліковані документи відкритого доступу, які містять огляди

товарного ринку, інформацію про нові технології, створення партнерств, злиття та придбання, оголошення про робочі вакансії, про виставки та конференції тощо. Широко використовуються відомості, що знаходяться в документах, що вже є в компаніях, що ведуть розвідку OSINT, результати маркетингових досліджень, інформація, отримана на конференціях, при спілкуванні з клієнтами та колегами. Більшість цих даних потрапляє у мережеву пресу, прес-релізи чи публікуються на корпоративних веб-сайтах. Тому останнім часом велику популярність набувають бази даних на основі архівів мас-медіа, у тому числі (і переважно) мережевих.

3.2 Сутність цінності веб-сайті при проведенні OSINT

Веб-простір, що базується на фізичній інфраструктурі мережі Інтернет та протоколі передачі даних HTTP, об'єднує сотні мільйонів веб-серверів, підключених до мережі Інтернет (рис. 2.1). На початку існування веб-простору на невеликій кількості веб-сайтів публікувалася інформація окремих авторів щодо відносно великої кількості відвідувачів. Сьогодні ситуація різко змінилася, відбувся перехід до Інтернету другого покоління. Самі відвідувачі веб-сайтів беруть активну участь у створенні контенту, що призвело до різкого зростання обсягів інформації та динаміки Інтернет.

Сьогодні в Інтернеті вже існує вільно доступна для користувачів інформаційна база такого обсягу, який раніше важко було уявити. Більше того, обсяги цієї бази перевищують на порядок все те, що було доступне десятиліття тому. У серпні 2005 року компанія Yahoo! оголосила про те, що проіндексувала близько 20 млрд документів. Досягнення компанії Google у 2004 році становило менше, ніж 10 млрд. документів. Сьогодні Google заіндексувала понад трильйони веб-документів. За даними служби Netcraft Web Server Survey (news.netcraft.com, рис. 2.1), в даний час кількість адрес веб-сайтів перевищує 1198 млн. (з них близько 200 млн. активних) [24].

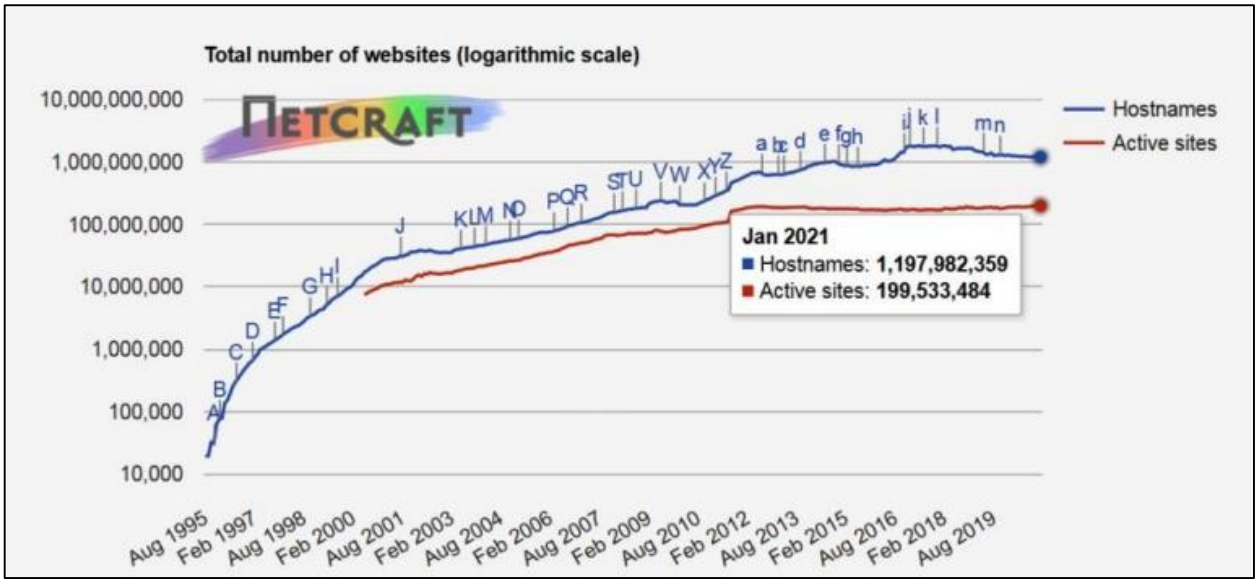

Рисунок 2.1 – Динаміка росту кількості веб-серверів за логарифмічною шкалою (Netcraft, 2021 року)

У відкритих джерелах та спеціалізованих базах даних, доступних у веб-просторі, міститься велика частина інформації, необхідної для проведення аналітичних досліджень, проте залишаються відкритими питання її знаходження та ефективного використання. При використанні веб-простору як найпотужнішого джерела інформації, як уже було зазначено раніше, найважливішими є проблеми обсягу, навігації, наявності інформаційного шуму та динамічного характеру інформації в Інтернеті.

Можливості доступу до інтернет-ресурсів, які приваблюють своєю відкритістю, обсягами та змістовною багатогранністю, на перший погляд здаються безмежними. Проте важливі події у різних галузях свідчать про протилежне. Саме у кризових ситуаціях Інтернет досить часто підводить. Існує безліч проблем – від перевантаженості мережної інфраструктури – до вірусних атак, вразливостей та відмов в обслуговуванні окремих веб-серверів. Ціла низка проблем породжена також обсягами, різноманітністю представлення та динамікою контентного сегменту інформаційного простору.

Незважаючи на такі якості, як відкритість та доступність, існуючу інфраструктуру веб-простору не можна визнати надійною та достовірною. Назвемо ще кілька проблем, властивих веб-простору:

- не вирішено завдання доступу користувачів до різноманітних веб-ресурсів з «одного вікна» для отримання узагальненого представлення потоків інформації з необхідної тематики;
- не забезпечена можливість своєчасного «нагадування» та «проштовхування» профільної для користувача інформації, що публікується на великій кількості веб-сайтів;
- досить велика ймовірність відмови в обслуговуванні критично важливих веб-ресурсів у найнеобхідніший час.

Відомо, що сьогодні існують технології інтеграції контенту, які дозволяють частково вирішувати ці проблеми, забезпечуючи ефективний пошук та навігацію у веб-просторі, моніторинг та агрегацію відкритих веб-ресурсів. Для професійного пошуку та агрегації інформації з веб-простору використовуються спеціалізоване програмне забезпечення, інформаційно-пошукові системи та сервіси.

3.3 Використання соціальних мереж та блогів в якості ресурсу для розвідки

Термін «соціальна мережа» означає зосередження соціальних об'єктів, які можна представити як мережу (чи граф), вузли якої – об'єкти, а зв'язки – соціальні відносини. Цей термін було запроваджено 1954 року соціологом з «Манчестерської школи» Дж. Барнсом (J. Barnes) у роботі «Класи та збори в норвезькому острівному приході». У другій половині XX століття поняття «соціальна мережа» стало популярним у західних дослідників, причому як вузли соціальних мереж стали розглядати не лише представників соціуму, а й інші об'єкти, яким притаманні соціальні зв'язки. Сьогодні термін «соціальна мережа» означає поняття, що виявилося ширшим за свій соціальний аспект, воно включає, наприклад, багато інформаційних

мереж, у тому числі і WWW. Розглядають як статистичні, а й динамічні мережі, для розуміння структури яких необхідний облік принципів їх еволюції [25].

Сьогодні під терміном «соціальні мережі» (Social Networks) розуміють насамперед онлайн-сервіси в мережі Інтернет, призначені для формування, відображення та впорядкування соціальних взаємин. Особливості соціальних мереж:

– надання користувачам широкого спектра можливостей для обміну інформацією;

– створення профілів користувачів, у яких потрібно вказати деяку кількість персональної інформації;

– друзями у соціальній мережі стають переважно не віртуальні, а реальні друзі.

Веб-ресурс соціальної мережі надає можливості:

– активного спілкування;

– створення публічного або закритого профілю (Profile) користувача, що містить персональні дані;

– організації та ведення користувачем списку інших користувачів, з якими має деякі соціальні відносини;

– перегляду зв’язків між користувачами всередині соціальної мережі;

– утворення груп користувачів за інтересами;

– управління вмістом у межах свого профілю;

– синдикації контенту;

– підключення різних програм.

Соціальні медіа є сукупністю онлайн-сервісів та інтернет-додатків, які дозволяють користувачам спілкуватися один з одним у тому числі, і в режимі реального часу. При цьому користувачі можуть обмінюватися між собою думками, новинами, інформацією, в тому числі мультимедійною.

Соціальні медіа базуються на ідеологічній та технологічній базі веб 2.0, що дозволяє створення та обмін контентом, що створений самими користувачами (UserGenerated Content), на відміну від попередньої концепції вебу, що передбачає, як і у випадку традиційних ЗМІ, централізоване створення контенту, що постачається користувачам-читачам.

Очевидно, соціальні медіа є великою цінністю, як джерело інформації для OSINT, надаючи абсолютно на легальних умовах різнобічну інформацію про людей, події, компанії, бренди, продукти.

Вирізняють сім різновидів соціальних медіа: соціальні мережі, блоги, форуми, сайти відгуків, сервери фото- та відеохостингу, віртуальні служби знайомств та геосоціальні мережі. Слід зазначити, що чіткі межі між цими різновидами розмиті.

Під соціальною мережею в мережі Інтернет (social networking service) розуміється онлайн-сервіс, призначений для побудови, відображення та організації соціальних взаємин, що забезпечує надання широкого спектра можливостей для обміну інформацією, можливість користувача надати інформацію про себе (створити свій профіль), побудувати зв'язки, знайти друзів за інтересами, підключити родичів, колег, однокласників тощо.

Під блогом (blog, від web-blog) розуміють веб-сайт, основний зміст якого – це записи, що періодично додаються користувачами (текст, зображення або мультимедіа). Для блогів характерні короткі записи (особливо, у випадках так званих «мікроблогів») тимчасової значущості, блоги зазвичай публічні і припускають сторонніх читачів, які можуть вступити в публічну полеміку з автором (у коментарі до блогозапису або в своїх блогах). Сукупність всіх блогів у мережі Інтернет називають блогосферою.

Веб-форуми є веб-додатками, призначеними для організації спілкування відвідувачів деяких інтернет-ресурсів (веб-сайтів або порталів). На ресурсах веб-форуму користувачі задають цікаві для них теми, які потім обговорюються й іншими користувачами шляхом розміщення повідомлень (постингу) всередині цих тем.

Веб-сайти відгуків створюються з метою підвищення ефективності та якості послуг та товарів, що надаються (не обов'язково в інтернет-середовищі). Користувачі, відвідуючи веб-сайти відгуків, залишають там свої повідомлення, беруть участь в анкетуваннях, формують думки про ту чи іншу послугу чи товар.

Фотохостинг (photo hosting) – це веб-сайт, що дозволяє публікувати будь-які зображення (найчастіше цифрові фотографії) в мережі Інтернет. Основна перевага фотохостингу – зручність демонстрації фотографій. Відповідно, відеохостинг – це веб-сайт, що дозволяє завантажувати та переглядати відеоінформацію у веб-браузері. Відеохостинг набирає популярності у зв'язку з розвитком широкосмугового доступу до Інтернету.

Віртуальна служба знайомств є інтернет-сервісом, що надає послуги з віртуального знайомства користувачів з цілями спілкування, створення сім'ї, серйозних стосунків та ін. (стаття, вік, мета знайомства, інтереси, фотографії). Після реєстрації користувач може спілкуватися з іншими користувачами, отримувати повідомлення та відповідати на них.

Геосоціальні мережі (GeoSocial Network) – це різновид соціальних мереж, в яких користувачі залишають дані про своє місцезнаходження, що дозволяє об'єднувати та координувати їхні дії на підставі інформації про те, які люди присутні в тих чи інших місцях, які події відбуваються у цих місцях.

3.3.1 Сутність основних соціальних мереж

До списку найбільших соціальних мереж, які можуть бути корисними для OSINT, можна включити мережі, що представлені рис. 2.2 [24]:

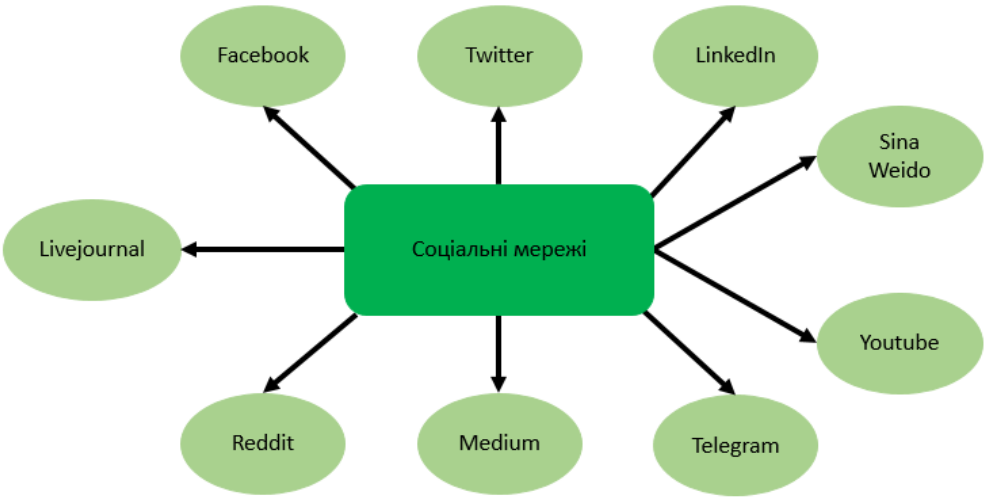

Рисунок 2.2 – Основні соціальні мережі, які користуються популярністю

Facebook (www.facebook.com) – найбільша соціальна мережа, заснована 2004 року М. Цукербергом та його компаньйонами. Починаючи з вересня 2006 року, соціальна мережа доступна для користувачів мережі Інтернет. На червень 2017 року аудиторія Facebook становила 2 мільярди користувачів. Добова активна аудиторія у березні становила 720 мільйонів осіб. Близько 500 млн. осіб на місяць використовують мобільні програми Facebook. Щодня в соціальній мережі користувачі залишають 6 мільярдів «лайків» та коментарів та публікують 300 мільйонів фотографій. На сайті зафіксовано 200 мільярдів «дружніх зв'язків». Щомісячна кількість переглядів сторінок Facebook перевищує 1 трлн.

Twitter (twitter.com) – сервіс, що дозволяє користувачам надсилати короткі текстові нотатки (до 140 символів), використовуючи веб-інтерфейс, SMS, засоби миттєвого обміну повідомленнями або сторонні програми-клієнти. Власником системи Twitter є компанія Twitter Inc., головний офіс якої знаходиться у Сан-Франциско. Станом на 1 січня 2011 року сервіс налічує понад 200 млн. користувачів. 100 млн користувачів виявляють активність хоча б раз на місяць, із них 50 мільйонів користуються Twitter щодня. 55% користуються Twitter на мобільних гаджетах, близько 400 мільйонів унікальних відвідувань отримує за місяць безпосередньо сайт

twitter.com. Особливістю Twitter є публічна доступність розміщених повідомлень; це називається мікроблоггінгом.

LinkedIn (www.linkedin.com) – соціальна мережа для пошуку та встановлення ділових контактів. Соціальну мережу LinkedIn було засновано Рідом Хоффманом у грудні 2002 року, запущено у травні 2003 року. 13 червня 2016 року Microsoft оголосила про придбання LinkedIn за ціною 196 доларів за акцію (загальна ціна угоди – 26,2 мільярда доларів), що є найбільшим придбанням Microsoft на сьогоднішній день. Наприкінці липня 2020 року LinkedIn оголосила про звільнення 960 співробітників; скорочення були спричинені наслідками глобальної пандемії Covid-19. Ця соціальна мережа надає можливість зареєстрованим користувачам створювати та підтримувати список ділових контактів. Контакти можуть бути запрошені як із сайту, так і ззовні, однак LinkedIn вимагає попереднього знайомства з контактами. Список контактів LinkedIn може використовуватися для розширення зв’язків, пошуку компаній, людей та груп за інтересами, публікації резюме та пошуку роботи, рекомендувати користувачів, публікувати вакансії, створювати групи за інтересами. Соціальна мережа LinkedIn також дозволяє публікувати інформацію про ділові поїздки та конференції. На 2020 рік загальна кількість користувачів LinkedIn досягла 675 мільйонів, з них 310 мільйонів активних.

Sina Weibo (кит. 新浪微博, http://weibo.com) – сервіс мікроблогів, запущений компанією Sina Corp. 14 серпня 2009 року, один із найпопулярніших інтернет-сервісів (платформ соціальних мереж) у Китаї та у світі. На початку 2018 року він перевищив ринкову оцінку у 30 мільярдів доларів США. Станом на лютий 2013 року кількість користувачів сервісу складає понад 500 мільйонів. У червні 2020 року Sina Weibo досягла 523 мільйонів активних користувачів на місяць.

YouTube (youtube.com) – відеохостинг, що надає користувачам послуги зберігання, доставки та показу відео. Сервіс, створений у лютому 2005 року трьома колишніми співробітниками PayPal – Чадом Херлі, Стівом Ченом та Джаведом Карімом, – купили Google у листопаді 2006 року за 1,65 мільярда доларів США. Згідно

з рейтингом Alexa в Інтернеті, YouTube є другим за відвідуваністю сайтом після Google Search. Користувачі можуть завантажувати, переглядати, оцінювати, коментувати, додавати до вибраного та ділитися тими чи іншими відеозаписами. Станом на 2019 рік на YouTube завантажують близько 300 годин відео щохвилини, а кількість щоденних переглядів відео досягла 5 млрд.

Telegram (telegram.org) – кросплатформний месенджер, що дозволяє обмінюватися повідомленнями та медіафайлами багатьох форматів. Користувачі можуть надсилати повідомлення та обмінюватися фотографіями, стікерами, голосовими та відео повідомленнями, файлами будь-якого типу, а також робити аудіо- та відеодзвінки. Кількість щомісячних активних користувачів сервісу станом на січень 2021 року становить близько 500 млн. осіб. Станом на березень 2020 року офіційні клієнти для Telegram включають:

- Мобільні програми для Android та iOS/iPadOS;
- Настільні програми для Windows, Linux та macOS;
- Веб-додаток, веб-програми для Chrome app, веб-додаток для React.

З 28 січня 2021 року в Telegram з'явилася можливість імпортувати чати та історію листування з інших месенджерів, зокрема й WhatsApp.

Medium (medium.com) – платформа соціальної журналістики. Сервіс запущений у серпні 2012 року співзасновниками Twitter – Еваном Вільямсом та Бізом Стоуном. Вільямс, раніше співзасновник Blogger та Twitter, спочатку розробив Medium як засіб публікації листів та документів, довжина яких перевищує максимум 140 символів Twitter (тепер 280 символів). Серед 75 співробітників компанії 15 авторів та редакторів. Платформа випускає видання Matter, Cuepoint, Backchannel, Re:form, Vantage та The Nib. Станом на травень 2017 року Medium мав 60 мільйонів унікальних читачів на місяць.

Reddit (reddit.com) – соціальний новинний сайт, на якому зареєстровані користувачі можуть розміщувати посилання на будь-яку інформацію в інтернеті. Як і багато інших подібних сайтів, Reddit – один з найпопулярніших сайтів у світі, займає

19-е місце за відвідуваністю за даними Alexa Internet. Reddit був заснований 23 червня 2005 року випускниками Віргінського університету Стівом Хаффманом та Алексісом Оганяном. Оцінна вартість Reddit складає $6 млрд. У 2019 році щомісяця налічувалося близько 430 мільйонів користувачів Reddit, відомих як «редитори».

LiveJournal, LJ (www.livejour-nal.com) – платформа для ведення онлайн-щоденників (блогів), створена 1999 р. американським програмістом Бредом Фіцпатріком. LiveJournal надає користувачам можливість публікувати свої та коментувати чужі записи, вести колективні блоги («спільноти»), додавати в друзі інших користувачів та стежити за їх записами у «стрічці друзів». До кінця грудня 2016 року сервери LiveJournal перебували в США та система належала американській компанії LiveJournal, Inc., але з грудня 2016 року LiveJournal розміщено на серверах російської компанії Rambler&Co. Серед опцій «Живого Журналу» слід виділити: різні типи записів та можливості їх коментування, вказівку розширених відомостей про користувача, друзі та стрічка друзів, картинки користувачів, функції безпеки облікового запису. На кінець 2012 року, в LiveJournal було зареєстровано понад 40 млн. користувачів, з них 368 805 активних.

3.4 Сутність глибинного веб-простору

Останні дослідження веб-простору показали, що доступні через традиційні інформаційно-пошукові системи більш, ніж трильйон веб-сторінок – це лише «поверхнева видима частина айсберга».

Важливою проблемою є пошук інформації в «прихованому» або «глибинному» веб-просторі, де, як було зазначено вище, міститься незрівнянно більша кількість даних, потенційно важливих для OSINT, ніж у відкритій частині Інтернету.

Це, перш за все, динамічні веб-сторінки, інформація з численних баз даних, які можуть становити великий інтерес для аналітичної роботи. До розряду «прихованого» Інтернету відносяться повнотекстові інформаційні системи типу LexisNexis або Factiva.

До «прихованих» ресурсів Інтернету можна віднести також пірингові мережі, такі як BitTorrent, EDonkey, EMule, Gnutella, Kazaa.

Як було зазначено раніше, необхідної інформації у мережі Інтернет значно більше, ніж її охоплюють універсальні пошукові машини.

Передбачається, що на відміну від частини мережі Інтернет, що «пізнається», «прихована» частина виявилася в сотні разів більш об'ємною.

Аналітик часто стикається з ситуацією, коли йому відомо про існування у веб-просторі якогось документа, але не може знайти його за допомогою традиційних пошукових систем, якими сьогодні можна вважати такі системи, як Google, Yahoo!, Bing, Baidu, Рамблер або Мета. Однак, згадавши або знайшовши в закладках адресу (URL) цього документа, він легко виходить на нього. Тобто у веб-просторі цей документ є, а знайти його звичним способом не можна. Користувач стикається із невидимим для пошукових систем ресурсом.

Сукупність джерел у веб-просторі, недоступних користувачам традиційних пошукових систем, утворює так званий «глибинний веб» – поняття, запроваджене Джиллом Іллсвортом (Jill Ellsworth) 1994 р. тобто під глибинним веб (invisible web, deep web, hidden web) прийнято розуміти ту частину веб-простору, яка індексується роботами (web crawlers) пошукових систем. Використовуючи аналогію, інформація, недоступна для пошуку, знаходиться «в глибині» (англ. – deep). При цьому не варто плутати глибинний веб із ресурсами, зовсім недоступними з мережі Інтернет – це темний веб (dark web), і про нього тут не йтиметься. Деякі ресурси, доступ до яких відкритий лише для зареєстрованих користувачів, також належать до глибинного Інтернету.

У 2000 році американська компанія BrightPlanet (www.brigh-tplanet.com) опублікувала сенсаційну доповідь, в якій стверджується, що у веб-просторі в сотні разів більше сторінок, ніж їх вдалося проіндексувати найпопулярнішими на той час пошуковими системами. Компанія розробила програму LexiBot, яка дозволяє сканувати деякі динамічні веб-сторінки, що формуються з баз даних, та, запустивши

її, отримала несподівані дані. З'ясувалося, що в глибинному Інтернеті знаходиться в 500 разів більше документів, ніж доступно через пошукові системи. Звісно, ці цифри неточні. Крім того, стало відомо, що середня сторінка глибинного веб-простору на 27% компактніша за середню сторінку з видимої частини веб-простору [24].

Сьогодні ситуація змінилася, наприклад, провідні пошукові системи можуть індексувати документи, які представлені у форматах, що містять текст. Звичайно, це насамперед pdf, rtf та doc. У 2006 році Google запатентувала спосіб пошуку в глибинному Інтернеті.

У глибинному Інтернеті знаходяться веб-ресурси, не пов'язані з іншими ресурсами гіперпосиланнями – наприклад, сторінки, що динамічно створюються за запитами до баз даних, документи з баз даних, доступні користувачам через пошукові веб-форми (але не за гіперпосиланнями). Такі документи залишаються недоступними для робота, нездатного в режимі реального часу, правильно заповнити поля форми значеннями (формувати запити до баз даних).

Ось що йдеться про глибинний Інтернет у книзі [26]: «Більшість сторінок невидимого Інтернету можуть бути проіндексовані технічно, але не індексуються, тому що пошукові системи вирішили їх не індексувати… Більшість «невидимих» сайтів мають високоякісний контент. Просто ці ресурси не можуть бути знайдені за допомогою пошукових машин загального призначення.

…Деякі сайти використовують технологію баз даних, що справді складно для пошукової машини. Інші сайти, однак, використовують поєднання файлів, які містять текст та мультимедіа, а тому частина з них може бути проіндексована, а частина – ні.

… Деякі сайти можуть бути проіндексовані пошуковими машинами, але це не робиться тому, що пошукові машини вважають це непрактичним – наприклад, через вартість або тому, що дані настільки короткоживучі, що індексувати їх просто безглуздо – наприклад, прогноз погоди, точний час прибуття конкретного літака, який здійснив посадку в аеропорту тощо».

Основні обмеження, пов'язані з роботами пошукових машин можна пояснити такими основними причинами: для публічних пошукових служб важливіше забезпечити точність пошуку, ніж повноту, іноді важливіше забезпечити отримання відповіді на запит у прийнятний час, ніж точність. Звідси – обмеження на глибину сканування веб-ресурсів, спроби «фільтрації» контенту за змістом, відсіювання сторінок, що містять зайві вихідні гіперпосилання тощо.

Загальновизнано, що цінність ресурсів глибинного Інтернету іноді вища за цінність ресурсів видимої частини веб-простору.

Можна згадати ще одну причину поповнення глибинного Інтернету – власники свідомо не хочуть, щоб їх веб-ресурси знаходили за допомогою пошукових систем. Найчастіше такі веб-ресурси репрезентують щось не зовсім законне, хакерські форуми, архіви неавторизованого контенту тощо. Зрозуміло, що багато таких ресурсів дуже цікаві для вивчення аналітиками.

Багато компаній спочатку підключаються до спільної Мережі і лише потім витрачають великі кошти на захист. Власники сайтів можуть спробувати заборонити індексацію тих чи інших сторінок своїх ресурсів, прописавши команду, що забороняє, у файлі robots.txt, але пошукові системи можуть її проігнорувати. Тому такі ресурси або видаляють, або видаляють гіперпосилання, переводячи дані ресурси до глибинного веб-простору. Наприклад, деякі бізнес-каталоги відмовляються віддавати свої оголошення роботам пошукових систем, тобто, захищаючи свої інформаційні активи компанії переводять свої ресурси в глибинний веб.

Існує кілька типів ресурсів глибинного веб-сайту, наприклад, як було зазначено вище, це можуть бути швидко застарілі веб-сторінки. Крім того, до глибинного веб-простору відносяться веб-ресурси, що є мультимедійною інформацією. Як відомо, зараз ще не існує задовільних алгоритмів пошуку не текстової інформації. Сторінки, що динамічно генеруються за запитом, також часто потрапляють у глибинний веб. Часто без запиту таких сторінок немає, вони генеруються при запиті до баз даних. Виходить, що інформація начебто присутня у веб-просторі, але виникає вона лише в

момент обробки запиту, а універсального алгоритму заповнення їх роботами пошукових форм не існує. І, нарешті, якщо на веб-ресурс не ведуть жодні посилання, то роботи пошукових систем ніяк не можуть дізнатися про його існування.

Засновник компанії BrightPlanet Майкл Бергман (Michael K. Bergman) зміг виділити 12 різновидів глибинних веб-ресурсів, що належать до класу баз даних. У списку опинилися як традиційні бази даних (патенти, медицина та фінанси), так і публічні ресурси – оголошення пошуку роботи, чати, бібліотеки, довідники. Бергман зарахував до глибинних ресурсів спеціалізовані пошукові системи, які обслуговують певні галузі чи ринки, бази даних яких не включаються до глобальних каталогів традиційних пошукових служб.

До глибинного Інтернету також належать численні системи інтерактивної взаємодії з користувачами – допомоги, консультування, навчання, що вимагають участі людей для формування динамічних відповідей від серверів. До них також можна віднести і закриту (повністю або частково) інформацію, доступну користувачам Мережі тільки з певних адрес, груп адрес, іноді міст або країн. До «прихованої» частини Мережі багато хто зараховує і веб-сторінки, зареєстровані на безкоштовних серверах, які індексуються, у кращому разі, лише частково – пошукові системи, щоб уникнути рекламного спаму, не прагнуть обходити їх у повному обсязі [25].

До глибинного Інтернету також відноситься категорія так званих «сірих» сайтів, що функціонують на основі динамічних систем керування контентом (Dynamic Content Management Systems). У пошукових системах зазвичай обмежується глибина індексування таких сайтів, щоб уникнути можливого циклічного перегляду тих самих сторінок.

4 СУТНІСТЬ ГОЛОВНИХ ПЕРЕВАГ ТА НЕДОЛІКІВ ВИКОРИСТАННЯ МЕХАНІЗМУ OSINT. ФАКТОРИ УСПІШНОГО АНАЛІЗУ

4.1 Переваги використання OSINT

У сьогоднішню інформаційну епоху не можна недооцінювати життєво важливу роль, яку OSINT грає у різних галузях розвідки. Переваги OSINT охоплюють багато сфер сучасного світу. Нижче наведено основні з них[16]:

– Менш ризиковано: використання загальнодоступної інформації для збору розвідувальних даних не пов'язане з ризиком порівняно з іншими формами розвідувальних даних, такими як використання супутників-шпигунів або використання джерел землі для збору інформації, особливо у ворожих країнах.

– Економічність: збір даних OSINT зазвичай обходиться дешевше в порівнянні з іншими джерелами розвідувальних даних. Наприклад, використання людських джерел або супутників-шпигунів для збору даних є дорогим. Малі підприємства з обмеженим бюджетом розвідки можуть використовувати джерела OSINT із мінімальними витратами.

– Простота доступу: джерела OSINT завжди доступні, де б ви не були, і завжди актуальні. Джерела OSINT можуть використовуватися різними сторонами у будь-якому розвідувальному контексті. Все, що вам потрібно, це необхідні навички/інструменти для правильного збору та аналізу OSINT. Наприклад, військові відомства можуть прогнозувати майбутні атаки, аналізуючи активність у соціальних мережах, а корпорації можуть використати це для побудови своїх нових стратегій розширення ринку.

– Юридичні питання. Ресурси OSINT можуть бути розділені між різними сторонами, не переймаючись порушенням будь-якої ліцензії на авторське право,

оскільки ці ресурси вже опубліковані у відкритому доступі. Звичайно, при обміні сірою літературою використовуються деякі обмеження.

– Допомога фінансовим слідчим: наприклад, OSINT дозволяє спеціалізованим урядовим установам виявляти осіб, які ухиляються від сплати податків. Багато відомих знаменитостей і деякі гігантські компанії замішані в ухиленні від сплати податків, і моніторинг їх облікових записів у соціальних мережах, відпусток та способу життя має велике значення для державного інспектора, який може переслідувати їх через незадекларований дохід.

– Боротьба з контрафактною продукцією в Інтернеті: методи OSINT можуть використовуватися для виявлення підроблених продуктів/послуг та вказівок правоохоронним органам закривати такі сайти або надсилати користувачам попередження про припинення роботи з ними. Це велика перевага OSINT, особливо у боротьбі з контрафактною фармацевтичною продукцією та натуральними продуктами для здоров'я.

– Підтримка національної безпеки та політичної стабільності: це може бути найважливішою роллю OSINT. Також це допомагає урядам зрозуміти ставлення свого народу та діяти швидко, щоб уникнути будь-яких зіткнень у майбутньому. Розумні уряди використовують OSINT у своїх майбутніх стратегіях, особливо у внутрішній політиці.

4.2 Виклики OSINT

Усі методології збору розвідувальних даних мають деякі обмеження, і OSINT не є винятком із цього правила. У даному підрозділі будуть наведені деякі проблеми, з якими стикається збір даних OSINT.

– Величезний обсяг даних. Збір даних OSINT призведе до отримання величезної кількості даних, які необхідно проаналізувати, щоб вважати їх цінними. Звичайно, для цієї мети існує безліч автоматизованих інструментів, і багато урядів та гігантських компаній розробили власний набір інструментів та методів штучного

інтелекту для фільтрації отриманих даних. Однак величезний обсяг даних залишиться проблемою для збирача OSINT.

– Неструктурована інформація. Публічна інформація, доступна в Інтернеті, за своєю суттю сильно дезорганізована. Це означає, що дані, зібрані за допомогою OSINT, настільки різнорідні, що їх складно класифікувати, зв'язати та вивчити, щоб отримати відповідні взаємозв'язки та знання [8]. У цьому сенсі OSINT потребує таких механізмів, як інтелектуальний аналіз даних, обробка природної мови (NLP) або текстова аналітика, щоб гомогенізувати неструктуровану інформацію, щоб мати можливість її використовувати[27].

– Дезінформація. З цієї нагоди при реалізації функцій OSINT необхідно враховувати нехарчову інформацію, яка не допускає дотримання вимог щодо поширення харчових продуктів. Дії OSINT завжди повинні мати справу з надійною інформацією та слідувати перевіреним напрямам дослідження, щоб отримати позитивні та переконливі результати [28].

– Надійність джерел даних. Достовірність та авторитетність інформації дійсно є ключем до успіху OSINT-розслідувань [29]. Джерела OSINT, особливо коли вони використовуються у розвідувальному контексті, повинні бути ретельно перевірені секретними джерелами, перш ніж їм можна буде довіряти. Багато урядів передають неправильну інформацію, щоб ввести в оману процес збору даних OSINT. В ідеалі зібрані дані мають розпочатися з авторитетних, перевірених та достовірних джерел (офіційні документи, наукові звіти, надійні засоби масової інформації) [30]. На практиці OSINT також співіснуватиме з суб'єктивними або неавторитетними джерелами, що створюють як соціальні мережі або керовані ЗМІ [31]. Незважаючи на те, що цей тип джерел часто дезінформує, насправді з них можна отримати більше інформації для розслідування людей, груп або компаній[32].

– Людські зусилля. З вищесказаного можна зрозуміти, що сам обсяг даних вважається найбільшою проблемою збору OSINT. Людям необхідно переглядати вихідні дані автоматизованих інструментів, щоб знати, чи надійні зібрані дані і чи

заслуговують вони на довіру. Також необхідно порівняти їх з деякими секретними даними (це застосовно до деякої військової та комерційної інформації), щоб переконатися в їх надійності та актуальності. Це ефективно забирає час та коштовні людські ресурси.

– Сильні етичні/правові зміни. Численні побоювання з приводу конфіденційності, поваги та особистої недоторканності виникають з розвитком OSINT [33]. У цьому напрямі слід зазначити, що питання про те, чи є OSINT етичною проблемою, зазвичай знаходиться в галузі етики збору розвідданих [34]. З одного боку, незважаючи на те, що OSINT є загальнодоступним, він може розкривати інформацію, яка явно не розміщена в Інтернеті. Непокриті результати повинні поважати конфіденційність користувачів та не розкривати інтимні та особисті проблеми [35], беручи до уваги поточні відповідні правила (наприклад, GPDR [14]). У цьому відношенні такі аспекти, як сексуальна орієнтація, релігійні переконання, політичні погляди або компрометуюча поведінка можуть бути виведені з Інтернету, і цей процес розкриття інформації сьогодні може бути проблематичним у багатьох країнах. З іншого боку, область пошуку на основі OSINT має бути за визначенням обмежена відкритими джерелами даних. За жодних обставин засоби контролю доступу або методи автентифікації не можуть бути обійдені для отримання інформації[27].

4.3 фактори успішного аналізу(источник «конкурентная разведка»)

Зокрема роль OSINT при проведенні розвідки визначається низкою аспектів, серед яких оперативність надходження, обсяг, якість, ясність, легкість подальшого використання, вартість отримання і т. д. Наступні фактори впливають на процес планування та підготовки ведення OSINT[25]:

– Ефективне інформаційне забезпечення. Більшість потрібних довідкових матеріалів про об'єкти інформаційних операцій видобувається з відкритих джерел. Це

переважно досягається шляхом збору інформації зі ЗМІ. Накопичення даних із відкритих джерел є основною функцією OSINT.

– Релевантність. Доступність, глибина та масштаби публічно доступної інформації дозволяють знаходити необхідну інформацію без залучення спеціалізованих людських та технічних засобів розвідки.

– Спрощення процесів добування даних. OSINT надає необхідну інформацію, за винятком потреби у залученні зайвих технічних та людських методів ведення розвідки.

– Глибина аналізу даних. Як офіційна частина розвідувального процесу, OSINT дозволяє керівництву здійснювати глибокий аналіз загальнодоступної інформації з метою прийняття відповідних рішень.

– Оперативність. Різке скорочення часу доступу до інформації в Інтернеті. Скорочення людино-годин, пов'язаних з пошуком інформації, людей та їх взаємовідносин на основі відкритих джерел. Швидке отримання цінної оперативної інформації. Обстановка, що стрімко змінюється, під час криз повніше відображається в поточних репортажах CNN з місця подій.

– Об’єм. Можливість масового моніторингу певних джерел інформації, з метою пошуку контенту, що цікавить, людей і подій. Як показує досвід, грамотно зібрані фрагменти інформації з відкритих джерел у сукупності можуть бути еквівалентними або навіть значнішими, ніж професійні розвідувальні звіти.

– Якість. Порівняно зі звітами спеціальних агентів, інформація з відкритих джерел виявляється кращою вже тому, що позбавлена суб'єктивізму, не розбавлена брехнею.

– Зрозумілість. Так що, якщо у разі використання OSINT надійність відкритих джерел буває як ясною, так і неясною, то у разі таємно здобутих даних ступінь їх надійності завжди викликає сумніви.

– Легкість використання. Будь-які таємниці прийнято оточувати бар'єрами із грифів таємності, особливих режимів доступу. Що ж до даних OSINT, їх можна легко

передавати у будь-які зацікавлені інстанції. Можливе проведення комплексного розслідування на підставі даних з Інтернету.

– Вартість. Вартість видобутку даних у OSINT мінімальна, визначається лише вартістю використовуваного сервісу.

Зокрема, запропоновані сьогодні для OSINT програмно-технологічні рішення забезпечують:

– збір даних із соціальних мереж, таких як Facebook, Twitter або Youtube, аналіз зібраних даних;

– екстрагування із зібраного контенту суті подій;

– агрегування інформації, отриманої з Інтернету;

– інформаційний вплив у мережі Інтернет;

– оцінку достовірності інформації;

– моніторинг та розпізнавання ідентичності в мережі Інтернет, у тому числі за допомогою геолокації;

– роботу з інформацією, одержаною з невидимих, за допомогою традиційних мережевих пошукових систем, сегментів веб-простору (dark web, hidden web, deep web).

ВИСНОВКИ

Розвідка за відкритим джерелами (Open Sourse Intelligence) є несекретною інформацією, яка була навмисно виявлена, виділена, очищена і поширена серед обраної аудиторії для вирішення конкретного питання. Це забезпечує дуже міцну основу інших дисциплін розвідки. При систематичному застосуванні, продукти OSINT можуть знизити потребу в секретних ресурсах для збору розвідданих.

Згідно з опублікованим у 2001 році Довідником НАТО з розвідки за відкритими джерелами версії, існує чотири категорії відкритої інформації та розвідувальних даних:

- Дані з відкритих джерел (OSD).
- Інформація з відкритих джерел (OSINF).
- Розвідка з відкритих джерел (OSINT).
- Затверджений OSINT (OSINT-V).

OSINT включає всі загальнодоступні джерела інформації. Цю інформацію можна знайти онлайн або офлайн, зокрема в таких місцях:

- Інтернет, який включає наступне та багато іншого: форуми, блоги, сайти соціальних мереж, сайти для обміну відео, такі як YouTube, wikis, Whois records про зареєстровані доменні імена, метадані та цифрові файли, ресурси темної мережі, дані геолокації, IP адреси, системи пошуку людей і все, що можна знайти в Інтернеті.
- Традиційні засоби масової інформації (наприклад, телебачення, радіо, газети, книги, журнали).
- Спеціалізовані журнали, наукові публікації, дисертації, матеріали конференцій, профілі компаній, річні звіти, новини компаній, профілі співробітників і резюме.
- Фотографії та відео, включаючи метадані.
- Геопросторова інформація (наприклад, карти та комерційні зображення).

Зробивши висновки із вищевказаного, можна зрозуміти, що OSINT охоплює не лише онлайн-джерела. Паперові видання загальнодоступних джерел також повинні бути ретельно досліджені в рамках процесу збору OSINT. Проте онлайн-джерела становлять найбільший сегмент OSINT.

Аналіз методології OSINT показав, існує три методи збору інформації, а саме:

– Пасивний збір
– Полупасивний збір
– Активний збір

Активний збір та напівпасивний збір – це типи збору інформації, які зазвичай використовуються рідко під час збору OSINT. Пасивний збір краще, тому що він може таємно збирати інформацію із загальнодоступних джерел, і у цьому суть OSINT.

Дослідження операційного циклу OSINT показав, що даний механізм має 4 ключові етапи:

– Збір інформації
– Обробка інформації
– Використання (експлуатація) інформації
– Виробництво інформації

Аналіз основних інструментів показав, що існує безліч інструментів, механізмів, та сервісів для проведення розвідки на основі відкритих джерел. Проте не всі можуть бути корисними для конкретно поставленої мети. Деякі з них можуть працювати, інші – ні, але це невід'ємна частина стратегії OSINT – фахівцю потрібно з'ясувати, які джерела є корисними, а які зовсім не стосуються досліджень. Залежно від наявних даних та кінцевої мети правильний вибір найбільш відповідного інструмента матиме значення. Проте, різноманітне поєднання їх насправді є ключем до досягнення правдоподібних результатів.

Аналіз, щодо ведення розвідки на основі відкритих джерел, показав такі переваги та недоліки:

- Переваги:

- Менший ризик;
- Економічність;
- Простота доступу;
- Юридичні питання;
- Допомога фінансовим слідчим;
- Боротьба з контрафактною продукцією в Інтернеті;
- Підтримка національної безпеки та політичної стабільності

- Недоліки:
  - Величезний обсяг даних;
  - Неструктурована інформація;
  - Дезінформація;
  - Надійність джерел даних;
  - Людські зусилля;
  - Сильні етичні/правові зміни;

Наступні фактори впливають на процес планування та підготовки ведення OSINT:

- Ефективне інформаційне забезпечення;
- Релевантність;
- Спрощення процесів добування даних;
- Глибина аналізу даних;
- Оперативність;
- Об’єм;
- Якість;
- Зрозумілість;
- Легкість використання;
- Вартість.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Lande D.V., Shurko-Tabakova E. V. OSINT as a part of cyber defense system. Kyiv. 2019.
2. Sondrava S., Sharma P., Dholariya D. Prevention to sensitive information disclosure via OSINT. International Journal of Scientific Research in Science, Engineering and Technology, 2021. P. 109-114.
3. Schaurer F., Störger J. The evolution of open source intelligence (OSINT). Comput Hum Behav, 2013. P. 53-56.
4. ORGANISATION, North Atlantic Treaty. NATO Open Source Intelligence Handbook. Brussells: North Atlantic Treaty Organisation, 2001.
5. Williams H. J., Blum I. Defining second generation open source intelligence (OSINT) for the defense enterprise. Rand Corporation. USA. 2018.
6. Cybercrime investigators are users too! Understanding the socio-technical challenges faced by law enforcement / Nouh M., Nurse J. R., Webb H., Goldsmith M. arXiv preprint arXiv:1902.06961, 2019.
7. Bello-Orgaz G., Jung J. J., Camacho D. Social big data: Recent achievements and new challenges. Amsterdam: Information Fusion. 2016. P. 45-59.
8. Bakshy, E., Rosenn, I., Marlow, C., Adamic, L. The role of social networks in information diffusion. Proceedings of the 21st international conference on World Wide Web. 2012. 519-528.
9. Using Open Data to Detect Organized Crime Threats: Factors Driving Future Crime / Larsen H. L., Blanco J. M., Pastor R. P., Yager R. R. Cham: Springer, 2017.
10. Dawson M., Lieble M., Adeboje A. Open source intelligence: Performing data mining and link analysis to track terrorist activities. Information Technology-New Generations. Cham: Springer. 2018. P.159-163.

11. Counter terrorism on online social networks using web mining techniques / Ali F., Khan F. H., Bashir S., Ahmad U. Singapore: Springer, 2018.P. 240-250.
12. Jang-Jaccard, J., Nepal, S. Surya. A survey of emerging threats in cybersecurity. Amsterdam: Journal of Computer and System Sciences. 2014. 973-993.
13. I don't trust ICT: Research challenges in cyber security / Gómez Mármol F., Gil Pérez M., Martínez Pérez G. Cham: Springer, 2016. P. 129-136.
14. Optimal countermeasures selection against cyber attacks: A comprehensive survey on reaction frameworks / Nespoli P., Papamartzivanos D., Mármol F. G., Kambourakis G. IEEE Communications Surveys & Tutorials, 2017. P. 1361-1396.
15. Quick D., Choo K. K. R. Digital forensic intelligence: Data subsets and Open Source Intelligence (DFINT+ OSINT): A timely and cohesive mix. Future Generation Computer Systems. Elsevier, Amsterdam. 2018. P. 558-567.
16. Hassan N. A., Hijazi R. Open Source Intelligence Methods and Tools. Apress, New York. 2018.
17. Fas. Final Report of the Commission on the Intelligence Capabilities of the United States Regarding Weapons of Mass Destruction. 2017. https://fas.org/irp/offdocs/wmdcomm.html (дата звернення: 05.10.2022)
18. DigitalGlobe trading higher after raising $279 million in IPO. USA: Wayback Machine, 2009. URL: https://web.archive.org/web/20140810160139/http://www.dailyfinance.com/2009/05/14/digitalglobe-trading-higher-after-raising-279-million-in-ipo/ (дата звернення: 05.10.2022)
19. Darla C, and Scola N, Mapping the World's 4.3 Billion Internet Addresses, США, The Washington Post, 2015. URL: https://www.washingtonpost.com/graphics/business/world-ip-addresses/ (дата звернення: 06.10.2022)
20. Central Intelligence Agency, Establishment of the DNI Open Source Center, США News and Information, 2005, URL:

https://www.cia.gov/news-information/press-releases-statements/press-release-archive-2005/pr11082005.html (дата звернення: 10.10.2022)

21. Detecting linguistic markers for radical violence in social media / Cohen K., Johansson F., Kaati L., Mork J. C. London: Routledge, 2014. P. 246-256.
22. Colquhoun, Cameron, A Brief History of Open Source Intelligence, The Netherlands, Bellingcat. 2016. URL: https://www.bellingcat.com/resources/articles/2016/07/14/a-brief-history-of-open-source-intelligence/ (дата звернення: 15.10.2022)
23. Распознавание информационных операций / Додонов А. Г., и др.; Киев, 2017.
24. Компьютерная конкурентная разведка / Додонов А. Г., Ландэ Д. В., Прищепа В. В., Путятин В. Г. Киев, 2021.
25. 25 лучших OSINT-инструментов для тестирования на проникновение в систему. XMLDATAFEED. 2021. URL: https://xmldatafeed.com/25-luchshih-osint-instrumentov-dlya-testirovaniya-na-proniknovenie-v-sistemu/ (дата звернення: 20.10.2022)
26. Sherman C. B., Sherman C., Price G. The invisible Web: Uncovering information sources search engines can't see. Melford: Information Today. 2001.
27. The not yet exploited goldmine of OSINT: Opportunities, open challenges and future trends / Pastor-Galindo J., Nespoli P., Mármol F. G., Pérez G. M. IEEE Access, 8, 2020. P. 10282-1030.
28. Mármol F. G., Pérez M. G., Pérez G. M. Reporting offensive content in social networks: toward a reputation-based assessment approach. IEEE internet computing. 2014. P. 32-40.
29. Gong S., Cho J., Lee C. A reliability comparison method for OSINT validity analysis. IEEE Transactions on Industrial Informatics. 2018. P. 5428-5435.
30. Fleisher C. S. Using open source data in developing competitive and marketing intelligence. European journal of marketing. 2008.

31. Screening out social bots interference: Are there any silver bullets? / Zago M., et. al.; IEEE Communications Magazine, 2019. P. 98-104.
32. Weir G. R. S. The limitations of automating OSINT: understanding the question, not the answer. Syngress: Automating Open Source Intelligence. 2016. P. 159-169.
33. Casanovas P. Cyber warfare and organised crime. A regulatory model and meta-model for open source intelligence (OSINT). Ethics and Policies for Cyber Operations. Springer, Cham. 2017. P. 139-167.
34. Bean H. Is open source intelligence an ethical issue? Government Secrecy. Emerald Group Publishing Limited, 2011.
35. Which side are you on? A new Panopticon vs. privacy / Kandias M., Mitrou L., Stavrou V., Gritzalis D. International Conference on Security and Cryptography (SECRYPT). – IEEE, 2013. P. 1-13.
36. Simola J. Privacy issues and critical infrastructure protection. Emerging Cyber Threats and Cognitive Vulnerabilities. Academic Press. 2020. P. 197-226.

ДОДАТОК А

ДОСЛІДЖЕННЯ ТА АНАЛІЗ ІНСТРУМЕНТІВ І МЕТОДІВ, ЩО ВИКОРИСТОВУЄ МЕХАНІЗМ OSINT

Єсіна Марина Віталіївна
доцент, к.т.н.
Азаров Михайло Олегович
студент
Національний університет імені В. Н. Каразіна
м. Харків, Україна

Вступ./Introduction. Цінність інформації з відкритих джерел на додаток до секретної розвідки давно визнана, однак з поширенням Інтернету та розвитком соціальних мереж, аналітики великих даних за останні два десятиліття зробили революцію у розвідці з відкритих джерел (OSINT). Нові зусилля зі збору та використання даних приносять цінні, оригінальні джерела даних у розвідувальному співтоваристві та оборонних підприємствах. Відкриті джерела також можуть замінити або доповнити деякі види доступу, які колись можна було отримати лише за допомогою більш небезпечних та дорогих традиційних платформ збору розвідданих. Хоча розвідувальна спільнота займається OSINT більше 50 років, визначення OSINT і те, як його характеризують як розвідувальну дисципліну, все ще є предметом суперечок. У документі 2011 року, випущеному Управлінням директора національної розвідки, OSINT визначається як «розвідувальні дані, отримані із загальнодоступної інформації, які збираються, використовуються та своєчасно поширюються серед відповідної аудиторії з метою задоволення конкретних вимог розвідки». Розвиток Інтернету та соціальних мереж ще більше ускладнили цю проблему. OSINT стає дедалі складнішим з погляду, як джерел, так і методів. Люди роблять інформацію доступною способами, яких раніше не існувало, включаючи онлайн-вираз особистих почуттів, фотографії місцевості та подій, а також загальнодоступні соціальні та професійні мережі. Поєднання обчислювальної потужності та методів обробки даних дозволяє зберігати та

обробляти велику кількість загальнодоступних даних. Машинне навчання, комп'ютерні алгоритми та автоматичні міркування ще більше розширюють можливості обробки цієї інформації та пошуку наслідків, які становлять цінність для розвідки.

Ці нові методи та системи також потребують високого рівня технічних знань для фахівців зі збору даних та аналітиків, які опрацьовують загальнодоступну інформацію. Історичним завданням OSINT був насамперед переклад – забезпечення доступу до зарубіжних статей новин для аналітиків розвідки з усіх джерел. Як тільки в значній мірі оригінальний звіт з відкритих джерел було переведено, аналітики, які використовують усі джерела, зазвичай могли включити його до готового аналітичного продукту. Крім того, інформація часто надходить від приватних осіб, що створює нові складності для незалежної розвідки у захисті конфіденційності громадян. Всі ці зміни вимагають більш надійного визначення OSINT, оскільки відкриті дані можуть надходити в багатьох формах.

Мета роботи./ The purpose of the work. Метою даної роботи є аналіз поточного стану методу для розвідки на основі відкритих джерел OSINT та проведення всебічного огляду парадигми з акцентом на послуги та методи, що розширюють область кібербезпеки. OSINT включає в себе пошук, вибір і збір інформації, отриманої із загальнодоступних джерел і її аналіз. Також до мети роботи можна віднести визначення ключових факторів для успішного проведення розвідки за відкритими джерелами і аналіз основних інструментів та методів, що використовуються для збору інформації.

Матеріали та методи./Materials and methods. Розвідка за відкритими джерелами, або OSINT (Open Source INTelligence), є несекретною інформацією, яка була навмисно виявлена, виділена, очищена і поширена серед обраної аудиторії для вирішення конкретного питання. Це забезпечує дуже міцну основу інших дисциплін розвідки.

Відкриті джерела інформації є винятковою прерогативою розвідувальних служб. Розвідка ніколи не повинна прагнути обмежити доступ до відкритих

джерел. Навпаки, оперативні дані повинні сприяти використанню відкритих джерел усіма співробітниками, які потребують доступу до відповідної надійної інформації. Штаби розвідки повинні зосередитися на застосуванні перевірених розвідувальних процесів для використання відкритих джерел для покращення своєї розвідувальної продукції, що базується на всіх джерелах. Знайомство з доступними відкритими джерелами дозволить співробітникам розвідки направляти та давати поради іншим співробітникам у їхньому власному використанні відкритих джерел.

Існують два найпоширеніші випадки використання розвідки за відкритими джерелами, такі як:

1. Етичний злом та тестування на проникнення. Фахівці з безпеки використовують дані з відкритих джерел для виявлення потенційних слабких місць у дружніх мережах, щоб їх можна було усунути до того, як ними скористаються зловмисники. До недоліків, що часто зустрічаються, відносяться:

– Випадкові витоки конфіденційної інформації, наприклад, через соціальні мережі.

– Відкриті порти або незахищені пристрої, підключені до Інтернету.

– Невиправлене програмне забезпечення, наприклад, веб-сайти, на яких запущено старі версії найпоширеніших продуктів CMS.

– Витоки або відкриті активи, такі як власний код у pastebins.

2. Виявлення зовнішніх загроз

Інтернет – чудове джерело інформації про головні загрози для організацій. Від визначення того, які нові вразливості активно використовуються, до перехоплення «листувань» зловмисників про майбутню атаку, аналітика відкритих джерел дозволяє фахівцям безпеки розставляти пріоритети у своєму часі та ресурсах для усунення найбільш серйозних поточних загроз.

У більшості випадків цей тип роботи вимагає, щоб аналітик ідентифікував і зіставив кілька точок даних, щоб підтвердити загрозу, перш ніж

робити будь-які дії. Наприклад, якщо хоча один твіт із погрозами може не викликати занепокоєння, то цей же твіт буде розглядатися в іншому світлі, якщо він буде пов'язаний із групою погроз, яка, як відомо, активна у певній галузі.

Одна з найважливіших речей, які потрібно зрозуміти про розвідку з відкритих джерел, полягає в тому, що даний механізм часто використовується в поєднанні з іншими підтипами розвідки. Інформація із закритих джерел, таких як внутрішня телеметрія, закриті даркнет-спільноти та зовнішні спільноти з обміну інформацією, регулярно використовується для фільтрації та перевірки інформації з відкритих джерел. Існує безліч інструментів, які допомагають аналітикам виконувати ці функції.

Існує безліч механізмів і методів Open Source Intelligence, проте не всі можуть бути корисними для конкретно поставленої мети. Деякі з них можуть працювати, інші – ні, але це невід'ємна частина стратегії OSINT – фахівцю потрібно з'ясувати, які джерела є корисними, а які зовсім не стосуються досліджень.

Перш, ніж приступити до збору інформації з відкритих джерел, необхідно визначитися із завданням:

– Яку інформацію потрібно знайти та дослідити?

– Яка мета пошуку?

– Хто чи що мета пошуку та подальших досліджень?

– Яким чином збір та аналіз даних буде виконуватись?

Відповіді на ці питання – перший крок у OSINT.

Найбільш популярні методи OSINT, які знаходяться у відкритому доступі та використовуються в кібербезпеці проти людини та організацій:

– визначення співробітників (ПІБ, посади), а також програмне забезпечення, з яким працюють співробітники;

– збирання інформації через пошукові системи Google (особливо за допомогою Google Dorks), Yandex, Yahoo, Bing та інші;

– аналіз інформації у соціальних мережах (Instagram, Facebook,

Twitter, тощо), на форумах, блогах та інших віртуальних ресурсах;

– пошук за фотографіями, наприклад, через Google, Yandex, TinEYE та інші подібні ресурси;

– визначення контактних номерів телефону та подальший пошук за номером у соціальних мережах, месенджерах, інших сайтах;

– перегляд збережених копій сайтів через Google;

– вивчення веб-сайтів в архіві Інтернету, наприклад, через сервіс Wayback Machine;

– використання Google Maps та інших джерел супутникових зображень для отримання географічного розташування користувачів.

Збір інформації з кола джерел – це трудомістка робота, але є безліч інструментів, які дозволяють спростити збір розвідданих. Найбільш популярні спеціалізовані OSINT технології, які знаходяться у відкритому доступі, проте потребують більш просунутого рівня володіння та знання інформаційних технологій:

– пошук через OSINT Framework, який містить посилання на велику колекцію ресурсів для вирішення найрізноманітніших завдань – від збору адресу електронної пошти до пошуку в соціальних мережах та dark Web;

– використання автоматизованих інструментів OSINT для отримання інформації, наприклад, через Spiderfoot, Maltego, Recon-ng, FOCA та інші;

– застосування інструментів для пошуку підключених до Інтернету пристроїв, наприклад, через пошукову систему Shodan або Censys;

– використання інструментів збору даних про людей, таких як Pipl, які допоможуть вам розкрити багато інформації про людей в одному місці;

– збір інформації через метапошукову систему Searx, що дозволяє збирати дані анонімно з більш, ніж 70 пошукових сервісів;

– відстеження розташування людини за його фотографіями, наприклад, через сервіс GeoCreepy;

– ще один чудовий інструмент, який можна використовувати для збору загальнодоступної інформації, є Metagoofil – використовує пошукову

систему Google для отримання загальнодоступних PDF-файлів, документів Word, Powerpoint і Excel із заданого домену;

– застосування спеціальних OSINT-розширень, таких як Open Source Intelligence Browser Extension;

– вивчення служб DNS, доменів, піддоменів та IP-адрес;

– застосування інших інструментів Kali Tools дозволяє виконувати не тільки розвідку, а й безліч інших завдань;

– знаходження працюючих служб через сканування портів в інфраструктурі цільової компанії;

Результати та обговорення./Results and discussion. Збираючи інформацію та аналізуючи дані з відкритих джерел, зловмисник чи пентестер має можливість сформувати повноцінний профіль жертви, визначити існуючі та потенційні вразливості. Цілеспрямовані кібератаки, як і військові атаки, починаються з попередньої розвідки, і перший етап цифрової розвідки – це OSINT, пасивне отримання розвідданих без запобігання меті. Без активного залучення своєї жертви, зловмисник або пентестер може використовувати отриману інформацію для побудови моделі загрози, розробки плану атаки або захисту.

Можна також виконати зворотний доксинг та OSINT по відношенню до себе чи свого бізнесу, адже це чудовий спосіб проаналізувати, яка інформація може стати відома потенційним зловмисникам. Які уразливості виявляє ваша публічна інформація? Які дані можуть бути отримані зловмисником? Що він може використовувати для застосування фішингових атак чи соціальної інженерії? Маючи дану інформацію, у Вас та Вашої команди буде можливість розробити ефективні методи протидії та захисту.

Висновки./Conclusions. Широке використання форумів, соціальних мереж або засобів масової інформації, а також велика кількість наявних даних перетворюють розвідку відкритих джерел (OSINT) на наступну золоту жилу Інтернету. Вилучення знань із загальнодоступних джерел є способом вирішення існуючих проблем з іншого та інноваційної точки зору. Зокрема, результати, які

може запропонувати цей тип розвідки, можуть значно покращити кібербезпеку та кіберзахист. Отже, мають бути реалізовані автоматизовані процеси OSINT, здатні проводити розслідування в усіх частинах Інтернету та розширювати нашу свідомість через Інтернет.

У роботі також представлені деякі методи OSINT для пошуку та описані найсучасніші інструменти OSINT для розширених розслідувань. Залежно від наявних даних та кінцевої мети правильний вибір найбільш відповідного інструмента матиме значення. Проте, різноманітне поєднання їх насправді є ключем до досягнення правдоподібних результатів.