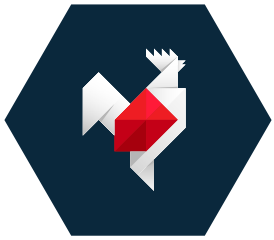


WPROWADZENIE
DO BEZPIECZEŃSTWA IT

REDAKCJA
Michał Sajdak

TOM 1

Łukasz Basa / Gynvael Coldwind / Tomasz Dacka / Marcin Dudek / Bartosz Jerzman
Konrad Jędrzejczyk / Wojciech Lesicki / Paweł Maziarz / Marcin Piosek / Iwona Polak
Piotr Ptaszek / Marek Rzepecki / Michał Sajdak / Wiktor Sędkowski / Grzegorz Trawiński
Tomasz Turba / Krzysztof Wosiński / Marek Zmysłowski

securitum

Projekt okładki: Krzysztof Kopciowski
Projekt typograficzny: Krzysztof Kopciowski
Redaktor: Michał Sajdak
Redakcja merytoryczna: Iwona Polak, Patryk Siaśkiewicz, Tomasz Turba, Mariusz 'maryush' Witkowski, Marek Zmuda

Redaktor prowadzący: Katarzyna Sajdak
Redakcja przypisów: Magdalena Anioł
Adiustacja: Katarzyna Sajdak
Skład i łamanie: Krzysztof Kopciowski
Korekta: Magdalena Anioł, Paulina Lenar

ISBN: 978-83-968874-0-5

Wydanie I
Kraków 2023

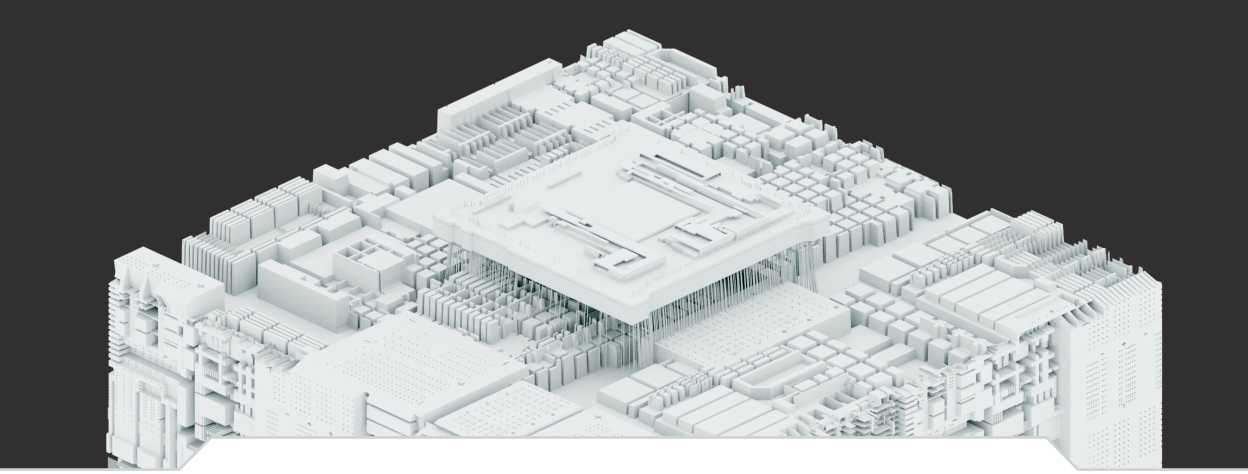
Securitum Wydawnictwo Spółka z ograniczoną odpowiedzialnością
ul. Siostry Zygmunty Zimmer 5
30-441 Kraków
e-mail: wydawnictwo@securitum.pl
www.securitum.pl

Zastrzeżenia prawne

Bezpieczeństwo IT ma coraz większe znaczenie. Nie można profesjonalnie zabezpieczyć aplikacji czy systemu, nie znając technik ich atakowania. Omawiamy je w tej książce, ponieważ to bardzo skuteczny sposób podnoszenia wiedzy i świadomości użytkowników, administratorów i twórców aplikacji. **Wszelkie podawane przez nas informacje powinny jednak być wykorzystywane wyłącznie w granicach prawa**, co z reguły oznacza zakaz wykorzystywania omawianej tu wiedzy bez zgody dysponenta systemu czy sieci.

Wyjście poza te granice może skutkować zarówno odpowiedzialnością cywilną (np. obowiązkiem naprawienia wyrządzonej szkody), jak i odpowiedzialnością karną. Przykładowo, zgodnie z polskim kodeksem karnym nieuprawnione uzyskanie dostępu do systemu informatycznego lub jego części podlega grzywnie, karze ograniczenia wolności albo pozbawienia wolności do lat 2 (art. 267 §2 kodeksu karnego). Z kolei nieuprawnione zakłócenie w istotnym stopniu pracy systemu informatycznego, systemu teleinformatycznego lub sieci teleinformatycznej przez transmisję, zniszczenie, usunięcie, uszkodzenie, utrudnienie dostępu lub zmianę danych informatycznych podlega karze pozbawienia wolności od 3 miesięcy do lat 5 (art. 269a kodeksu karnego).

Zwracamy na to uwagę, ponieważ **nie jest naszym zamiarem wspieranie jakichkolwiek bezprawnych działań**. Dlatego zastrzegamy, że w najszerszym prawnie dopuszczalnym zakresie wyłączamy naszą odpowiedzialność za skutki takich działań.



OSINT – WPROWADZENIE

Tomasz Turba

Tomasz 'z3' Turba. Specjalista do spraw bezpieczeństwa. Swoją przygodę z komputerami rozpoczął od hackowania Amigi 500. W branży IT działa od 2006 roku, przeszedł przez wszystkie szczeble kariery.

Autor kilku innowacyjnych szkoleń o tematyce cyberbezpieczeństwa oraz wielokrotny laureat nagród za publikacje na temat bezpieczeństwa IT. Współpracował z wieloma instytucjami jako konsultant do spraw zabezpieczeń, pentester i inspektor RODO. Ma duże doświadczenie jako szef zespołu CSIRT.

Prelegent podczas Mega Sekurak Hacking Party. Redaktor w portalu sekurak.pl. Pracuje w Securitum jako trener, pentester i analityk.

Prywatnie łowca demonów z doświadczeniem bojowym i biało-wywiadowczym. <3#ALRJ#

Podziękowania

Dziękuję mojej ukochanej Żonie, bez której ten rozdział i całe moje życie nie byłyby takie same. Za Twoje nieskończone wsparcie i cierpliwość.

Dziękuję moim kochanym Córkom, małym hackerkom, które są ciągłym przypomnieniem, dlaczego stale dążę do bycia lepszym. Za Waszą nieustającą ciekawość i uśmiechy dające mi siłę do działania.

WSTĘP

Zastanawiam się, czy każdy z nas zna jakąś historię o namierzeniu osoby lub obiektu w Internecie tylko na podstawie kilku – pozornie niezależnych od siebie – cech rozpoznawczych? Przytoczę na wstępie jedną z nich[1].

Agencja Bezpieczeństwa Wewnętrznego poszukiwała biznesmena, Janusza M., który był podejrzewany o uszczuplenie należnego podatku VAT w kwocie nie mniejszej niż 50 milionów złotych. Bogactwo poszukiwanego pozwalało mu skutecznie ukrywać się przed wymiarem sprawiedliwości. Nie przebywał zbyt często na lądzie, tylko pływał samotnie od portu do portu swoim bardzo drogim jachtem. Nie afiszował się luksusowym trybem życia, nie pojawiał się na portalach społecznościowych (choć dzisiaj robi to większość społeczeństwa i tzw. influencerzy, monetyzujący dzielenie się swoją prywatnością).

Los jednak chciał, że 60-letni Janusz poznał 25-letnią Paulinę i zakochał się w niej. Ona od początku uważała go za księcia z bajki i postanowiła pochwalić się na Instagramie zdjęciem z ukochanym, tak by cały świat dowiedział się, jaka jest szczęśliwa. Zdjęcie udostępniła bez jego wiedzy. Służby, mające w swym portfolio wizerunek osoby poszukiwanej europejskim nakazem aresztowania (ENA)[2], odnalazły publiczny profil Pauliny stosunkowo szybko i niezwłocznie rozpoczęła się wykorzystująca techniki OSINT analiza miejsc, w których przebywał poszukiwany. Do identyfikacji włączono także sieć ENFAST (European Network of Fugitive Active Search Teams)[3], której głównym założeniem jest błyskawiczna, międzynarodowa wymiana informacji pomiędzy służbami w sprawie tropu. Zaledwie po kilkunastu dniach od publikacji owego zdjęcia, w wigilijny poranek 2019 roku, służby odnalazły poszukiwanego w Wenecji[4]. Jedno zdjęcie z jego wizerunkiem, upublicznione w Internecie, udaremniło skuteczne ukrywanie się przed wymiarem sprawiedliwości[5].

W tym rozdziale skupimy się na białym wywiadzie. Przedstawię skuteczne techniki realizacji poszukiwań, narzędzia oraz dobre rady. By rozpocząć przygodę z OSINT-em, nie jest wymagana specjalistyczna wiedza techniczna, choć przydaje się ona w przypadku używania skomplikowanych narzędzi, które często nie posiadają graficznego interfejsu użytkownika. Nie stanowi to jednak przeszkody w realizacji poszukiwań, gdyż rozległość OSINT-u powoduje, że nawet w czasie, kiedy powstawał ten rozdział*, prawdopodobnie powstało kilka nowych, zautomatyzowanych narzędzi, które wcześniej były dostępne tylko dla „wtajemniczonych".

* Każde odwołanie do czasu powstania rozdziału dotyczy czerwca 2023 roku, o ile nie zaznaczono inaczej [przyp. red.].

OSINT, czyli Open-Source Intelligence, jest formą wywiadu jawnego, często nazywaną właśnie białym wywiadem. Różne praktyki wywiadowcze znane są ludzkości już od czasu powstania papieru, a może nawet wcześniej, niż przedstawiono to historycznie w tabeli 1. Podczas polowań ludzie pierwotni namierzali zwierzynę, badali jej cechy, określając „podatności", a to wszystko w ściśle z góry określonym celu nadrzędnym: żeby przetrwać.

Dzisiaj jest podobnie. Obserwujemy obiekt, który może być osobą, miejscem lub np. infrastrukturą. Zbieramy dane, badamy podatności w konkretnym celu (ataku lub ochrony). Psychologia została ta sama, tylko narzędzia mamy nowsze. Współczesny OSINT, ten z XXI wieku, nie jest jednak tym, czym był jeszcze 30 lat temu, kiedy to służby specjalne wykorzystywały lub pozyskiwały informacje z prasy, telewizji czy książek. **Obecnie OSINT jest formą zdobywania wiedzy z szeroko udostępnianego, cyfrowego, „zaszumionego" świata danych szczątkowych w Internecie.** Zarówno jednak dla jego starej, jak i nowej formy występuje cecha wspólna: w obu przypadkach wywiad jawny wymaga dużego nakładu pracy w pozyskaniu informacji, często potencjalnie bez znaczenia, które razem utworzą ważną całość. Ogromną rolę odgrywa także odsianie wspomnianego szumu informacyjnego, nawarstwiającego się z każdym skanowaniem.

Cele tego działania mogą być różne: indywidualne, dziennikarskie, związane z pracą służb lub biznesowe, a zamiary – dobre lub złe. Z dobrych warto wymienić: próbę odnalezienia osoby, adresu, obiektu lub weryfikację danych źródłowych, np. walkę z fake newsami. Te złe można by wymieniać długo: obserwacja, stalking, nauka nawyków do przeprowadzenia szerszej operacji (chociaż to też może być dobry powód), kradzież danych, phishing wraz z socjotechniką, fałszerstwo przemysłowe… i wiele, wiele innych. Z obu tych klasyfikacji można jednak wyróżnić podstawowe cele biznesowe: uzyskanie przewagi technologicznej nad konkurencją, weryfikację kontrahentów lub pracowników oraz analizę parametrów gospodarczych ważnych dla przedsiębiorstwa i służących jego rozwojowi.

Wiele osób rozpoczynających przygodę z białym wywiadem twierdzi, że „Google wystarczy". Nie wystarczy. Chociażby dlatego, że przeglądarki zupełnie inaczej klasyfikują i prezentują wyniki wyszukiwania. Nieważne, jak dobrze znamy Google Dorks*, nieważne, w jakim stopniu opanowaliśmy konkretne narzędzie analityczne czy kwerendy (zapytania) na portalu społecznościowym. **Biały wywiad to zrozumienie i nauczenie się pewnego schematu myślenia.** Schematu, który nie powinien nieść za sobą ograniczeń. To odszukiwanie informacji, których inni mogą nie dostrzec, np. analiza miejsca, w którym osoba poszukiwana zrobiła zdjęcie, bez dysponowania jego metadanymi, czy obserwacja znajomych w celu określenia ich nawyków, zachowań i reakcji. Wszystko ma znaczenie: numer rejestracyjny, reklama na budynku, napis na butelce z wodą, a nawet pora dnia i roku, które takie przykładowe zdjęcie utrwaliło.

Biały wywiad zmienia się każdego dnia, tak jak zmienia się cyfrowy świat oraz dostęp do danych. Współcześnie, już poza OSINT-em, wyróżnia się także jego odłam: **SOCMINT (Social Media Intelligence)**, czyli biały wywiad przeprowadzany za pomocą portali spo-

* Google Dorks są dostępnymi w sieci napisanymi przez kogoś zapytaniami umożliwiającymi wyszukiwanie konkretnego celu.

łecznościowych. Nie jest to jednak wiedza tajemna. Nie trzeba być informatykiem ani nawet osobą bardzo zaawansowaną technicznie, by zajmować się białym wywiadem i zrozumieć ten rozdział. Te cechy bardzo pomagają w poruszaniu się w świecie OSINT-u, ale nie są niezbędne. **Ważne jest ułożenie schematu myślenia jako procesu dojścia do skutecznego wnioskowania.**

Tabela 1. Kalendarium sposobów pozyskiwania informacji z otwartych źródeł (oprac. własne)

ROK	OPIS
...–1455	zapisy z rękopisów i ksiąg, ograniczony dostęp do bibliotek i samego piśmiennictwa
1455–1898	lata zmian społecznych, dostęp do publikacji, ogłoszeń, prasy, listów i książek
1898–1969	lata kryzysów i wojen spowodowały uformowanie i jednoczesne przekształcenie się jednostek wywiadowczych do formy specjalizacji (HUMINT, SIGINT, GEOINT, MASINT, TECHINT, FININT i w końcu OSINT)
1969–1994	powstanie sieci Internet, masowa komunikacja odbiorcza: kino, teatr, prasa, radio, telewizja, pierwsze bazy danych, utwierdzenie się w przekonaniu o ważności specjalizacji OSINT
1994–2004	powstanie wyszukiwarek, forów internetowych, czatów, internetowych baz danych, blogów, rozkwit mobilnego Internetu, powstanie anonimowej sieci Tor, utworzenie MySpace
2004–2008	rozkwit sfery społecznościowej Internetu, pierwsze regulacje prawne, masowa dostępność smartfonów
2008–2016	powstanie wielu portali społecznościowych, zmiana sposobu przedstawiania informacji – z tekstowej na multimedialną, rozkwit memów, pojawienie się specjalizacji SOCMINT jako części OSINT
2016–2019	masowe dostępy *on-demand*, rozkwit wirali (krótkich materiałów wideo o dużej popularności), regulacje w zakresie publikacji danych w sieci Internet, dyrektywa GDPR/RODO
2019–2023	rozwój technologii wspierających społeczności: VR, AR, masowe wirale, memy, rolki (krótkie tymczasowe filmy), dezinformacja na dużą skalę (np. sprzeczne informacje dotyczące pandemii COVID-19 czy wojny w Ukrainie)
2023+	rozwój narzędzi sztucznej inteligencji (ang. *Artifical Inteligence*, AI) w kontekście przeglądania, zdobywania i referowania informacji (wykorzystanie AI do analiz OSINT, ale także do zautomatyzowanej dezinformacji)

Czy jesteśmy gotowi na nadchodzące zmiany w związku z rozwojem AI? Jak szybko będziemy w stanie dostosować się do świata, który już za chwilę prawdopodobnie będzie tak inny od tego, który znamy? Odpowiedzi na te pytania przyniesie najbliższa przyszłość.

Świat IT przyspiesza coraz bardziej, dlatego nauka o podstawach jest tak ważna, bo to one pozwalają zrozumieć zasady i przystosować się do tego zmieniającego się świata, nawet jeśli nie mamy wpływu na kierunek, w którym idzie. Podstawy, które wypełniają ten rozdział, pozwolą zrozumieć, jak OSINT wpływa na życie, zarówno na poziomie osobistym, jak i zawodowym. Bo nawet w szybko zmieniającym się świecie, podstawy pozostają stałe. Są jak latarnie morskie w burzliwym oceanie technologii, pomagając odnaleźć kierunki poszukiwań i wyznaczając bezpieczną drogę, w którą wyruszyć trzeba ze sprawdzonym ekwipunkiem.

CELE I SPOSOBY REALIZACJI OSINT-U

Rodzaje białego wywiadu można sklasyfikować według podziału na cel i sposób pozyskiwania danych (przedstawionego w tabeli 2). Poprzez **cel** rozumie się tu zamiar uzyskania danych (atak lub obronę). Informacje mogą być zbierane, aby przeprowadzić ofensywę na obiekt bądź organizację, np. poprzez skrupulatne utworzenie wektora ataku, czyli sposobu jego przeprowadzenia od początku do końca w rozbiciu na różne fazy. Cele defensywne OSINT-u często określa sama organizacja lub obiekt, aby rozpoznać swoją wiedzę, testując ją prewencyjnie (samotestowanie).

W obu przypadkach sposób zbierania danych z otwartych źródeł może być różny: aktywny lub pasywny. W **sposób aktywny** działamy wtedy, gdy osoba poszukująca informacji wchodzi w interakcję z obiektem. Interakcja może odbywać się **bezpośrednio**, np. poprzez nawiązanie kontaktu (osobiście, telefonicznie, e-mailowo lub np. przez zaproszenie do zawarcia znajomości na portalu społecznościowym) lub **pośrednio**, np. przez skanowanie obiektu obserwacji za pomocą publicznej adresacji IP w Internecie (z której zostaje ślad w postaci logów możliwych do prześledzenia).

Metoda pasywna zakłada, że nie wchodzi się w interakcję z celem. Należy spełnić tu dwa główne kryteria realizacji: nie dać się namierzyć oraz nie wzbudzić cienia podejrzenia podczas ewentualnej fałszywej interakcji z obiektem. Sposób pasywny jest bezpieczniejszą metodą; we współczesnym Internecie istnieje wiele możliwości jej realizacji. W przypadku gdy obiekt posiada dobrze zabezpieczone konta na portalach społecznościowych, tj. nie udostępnia osobom nieuprawnionym swoich danych, wtedy siłą rzeczy należy wejść w interakcję bezpośrednią. Aby to jednak zrobić w dalszym ciągu w sposób pasywny, trzeba stworzyć fałszywą tożsamość lub – co skuteczniejsze – podszyć się pod osobę z grona znajomych obiektu. Z reguły w takim przypadku obiekt sądzi, że znajomy stracił dostęp do konta i stworzył drugie, więc akceptuje znajomość. Należy jednak pamiętać o aspektach prawnych i legalności takiego działania. Nie sugeruję tego rozwiązania jako właściwego, jednak wskazuję na istnienie takiej możliwości, co może ułatwić zrozumienie sposobów działania przestępców wskazanych w art. 190a § 2 *Kodeksu karnego*[*]. W przypadku skanowania organizacji, np. pod kątem enumeracji subdomen, można połączyć się przez szyfrowany tunel VPN (Virtual Private Network) lub wykorzystać darmowe narzędzia online do skanowania. Ostatecznie najlepiej jest zastosować obie metody, tzn. łączyć się do publicznego Internetu za pośrednictwem tunelu VPN (lub anonimowego serwera proxy) bądź z sieci Tor dla nieco bardziej zaawansowanych, a następnie korzystać ze skanerów online. Musimy jednak pamiętać, że tunel VPN niekoniecznie może zapewnić nam prywatność, a jedynie zmianę adresu IP widzianego po stronie skanowanego obiektu.

* § 1. Kto przez uporczywe nękanie innej osoby lub osoby jej najbliższej wzbudza u niej uzasadnione okolicznościami poczucie zagrożenia lub istotnie narusza jej prywatność, podlega karze pozbawienia wolności do lat 3.

§ 2. Tej samej karze podlega, kto, podszywając się pod inną osobę, wykorzystuje jej wizerunek lub inne jej dane osobowe w celu wyrządzenia jej szkody majątkowej lub osobistej.

Tabela 2. Podział analizy białego wywiadu ze względu na obiekt, cel oraz sposób działania

OBIEKT ANALIZY	OSOBA		INSTYTUCJA/FIRMA	
	przykłady: Jan Kowalski – potencjalny przestępca, Will Smith – znany aktor, Pani Janina – księgowa firmy LRATT, Wujek Stanisław – zaginiona osoba		przykłady: CDE Auto – legalna firma, OPUS LUPUS – fałszywa fundacja, FGHJKL.co – podejrzana domena, AAABBB.com.pl/news/10 – artykuł w portalu	
CEL ANALIZY	**ofensywny**	**defensywny**	**ofensywny**	**defensywny**
	pozyskanie danych o osobie w celu dokonania większego ataku (np. na infrastrukturę) typu: uzyskanie listy znajomych, kont e-mail, zdjęcia w celu preparowania ataku phishingowego (fałszywy profil z prawdziwym zdjęciem)	sprawdzenie, ile danych o osobie można uzyskać, np. w celach audytowych (wewnętrznych) lub poszukiwawczych osoby zaginionej; stworzenie własnych zabezpieczeń, np. *canary tokens**	nielegalne przejęcie, wykradzenie danych lub uszkodzenie infrastruktury, np. przez skanowanie infrastruktury w celu znalezienia podatności do dalszego ataku; często kradzież gospodarcza	namierzenie fałszywej instytucji; walka z fałszywymi treściami; działania związane z audytem IT w celu uszczelnienia infrastruktury; stworzenie własnych zabezpieczeń, np. *honeypot*
SPOSÓB DZIAŁANIA	**aktywny**	**pasywny**	**aktywny**	**pasywny**
	interakcja z osobą, np. zaproszenie do znajomych, kontakt telefoniczny, kontakt e-mailowy, wiadomość w komunikatorze, także kontakt ze znajomymi lub rodziną osoby na różnych portalach społecznościowych czy przez e-maila bądź inną formę komunikacji	obserwacja osoby bez wchodzenia w interakcję lub kontakt z celem z wykorzystaniem anonimowej tożsamości, np. użycie narzędzia FBStalker do sprawdzania, kiedy użytkownik aktywnie korzysta z profilu na portalu Facebook	aktywne skanowanie, śledzenie wątków na profilach w portalach społecznościowych, wchodzenie w interakcję z pracownikami; wykorzystanie specjalistycznych portali, aby sprawdzić istnienie konkretnej podatności/otwartego portu w publicznie dostępnej adresacji IP	przeszukiwanie portali merytorycznych i technicznych w celu zebrania jak największej liczby informacji, np. wykorzystanie narzędzia hunter[6] w celu sprawdzenia znalezionych kont e-mailowych

Przygotowanie do działania

W tabeli 2 zaprezentowano bardzo prosty podział współczesnej analizy otwartoźródłowej, jednak podstawą działania jest opracowanie i wykorzystanie planu wyszukiwania informacji, zwanego strategią wyszukiwawczą. Nieważne, jaki jest obiekt oraz cel analizy. Ostateczny sposób działania (pasywny, aktywny lub hybrydowy) jest bezpośrednio skorelowany z wybraną strategią bądź kilkoma strategiami realizowanymi wspólnie lub rozłącznie. W zależności od tego, jaką informację chce się uzyskać, strategia za każdym razem wprost zależy od rodzaju tejże informacji, posiadanej wstępnej wiedzy na temat poszukiwanego obiektu oraz własnego, nabytego wcześniej doświadczenia. Strategia wyszukiwawcza to nic innego, jak doprowadzenie do skutecznego wyszukiwania informacji poprzez ułożenie odpowiednich zapytań, a także ustalenie kolejności poszukiwań

* *Canary tokens* – to metoda wykrycia ataku lub zainteresowania w bardzo wczesnej fazie rekonesansu. Metoda występuje w postaci wygenerowania tokena w różnych słabo zauważalnych formach (np. w postaci pliku *.docx) – więcej w dalszej części rozdziału.

w celu maksymalizacji trafności spodziewanych wyników. W tabeli 3 zebrane zostały podstawowe strategie wyszukiwawcze, które są realizowane przez najbardziej popularne i powszechne wyszukiwarki treści w Internecie.

Tabela 3. Strategie wyszukiwania

NAZWA STRATEGII	OPIS
wyszukiwanie proste	wyszukiwanie poprzez wykorzystanie jednej lub kilku informacji połączonych za pomocą logiki boolowskiej[7]; najczęściej stosowane do znalezienia konkretnego rekordu na podstawie już istniejących danych cząstkowych, np.: „sekurak", „sekurak AND książka", „sekurak szkolenia"
formowanie klas	zidentyfikowanie wyszukanych pojęć i tworzenia relacji między nimi na zasadzie synonimów, np.: „policja", „sekurak", „bezpieczeństwo", „szkolenia"
kolejne klasy	zawężanie pojęć i redukcja wyników poprzez poszerzenie obszaru wyszukiwania; najczęściej realizowane pełnotekstowo, np.: „wojna w Ukrainie" + „kurs rubla" + data (od–do)
mnożenie odwołań	powtarzanie trzech poprzednich strategii w celu głębokiego precyzowania i wnioskowania na temat wyników, także przez używanie innych narzędzi i wyszukiwarek do tych samych wcześniejszych zapytań (w większości przypadków wyniki będą inne)
strategia cytowań	śledzenie cytowań i zmian treści informacji, linków, multimediów

Strategie wyszukiwań mogą być od siebie wzajemnie zależne. Przygotowując scenariusz poszukiwań, można informacje z tabeli 3 traktować jako swoisty drogowskaz w iteracji kolejnych wyszukiwań. Zaczynamy od prostych zapytań i w toku nabywania wiedzy cząstkowej kwerendy wyszukiwań są bardziej precyzyjne i dają węższą grupę wyników. Niezależnie jednak od tego, jaka strategia wyszukiwawcza zostanie wybrana, należy pamiętać, że działania definiują cele. Istnieje kilka wskazówek dla początkujących, którzy w miarę nabierania doświadczenia w przyszłości zrealizują swoje zadania szybciej, pewniej oraz przede wszystkim bardziej skutecznie. Fundamentem prawidłowego rozwoju nie są narzędzia czy sposoby weryfikacji informacji, a **mapa myśli**. Narzędzia i wnioskowanie są ważne, jednak bez dobrej mapy myśli nie jesteśmy w stanie sensownie rozwiązać danego problemu. Przepełni się nam w głowie bufor zebranych informacji w postaci szumu, który spowoduje, że analizę będziemy musieli rozpocząć prawie od nowa.

Poniżej przedstawiam pięć prostych kroków, które mogą być pomocne (mnie pomagają) w rozwijaniu kompetencji analityka.

DOBRE PRAKTYKI: DOSKONALENIE KOMPETENCJI ANALITYCZNYCH

KROK 1: **Odpoczynek**

Warto odpocząć. To nie żart. Jeżeli nasz umysł jest zmęczony, to przede wszystkim nie będzie funkcjonował na wysokich obrotach. W realnym scenariuszu, np. wpatrując się szesnastą godzinę w monitor, pijąc szóstą kawę, łatwo przeoczyć jedną drobną informację, która dla całości analizy mogła okazać się kluczowa.

KROK 2: **Porządek**

Nie ma nic bardziej przeszkadzającego w analizie niż chaos panujący na stanowisku pracy. Analiza OSINT czasami wymaga błyskawicznego działania i wnioskowania, gdyż od decyzji analityka może zależeć ludzkie życie lub przetrwanie infrastruktury.

Tak było, gdy zostałem poproszony o namierzenie uprowadzonej nastolatki[8]. Liczyła się każda minuta, bo choć podejrzewaliśmy uprowadzenie, to jednym z rozważanych scenariuszy była próba samobójcza[9]. Jeżeli komputer uruchamia się długo – warto zadbać o jego czyszczenie lub konserwację (np. wymiana sprzętu, aktualizacja oprogramowania lub odświeżenie systemu operacyjnego). Jeżeli biurko jest usłane papierami, być może zgubisz tę jedną istotną karteczkę – ważną dla całości analizy. Warto również pozbyć się wszelkich „przeszkadzajek" – najważniejsze jest skupienie.

KROK 3: **Narzędzia**

Znajomość narzędzi, sposobów wyszukiwania i miejsc może być kluczowa w osiągnięciu dobrych rezultatów. Nasze doświadczenie musi być jak największe. Wykorzystujmy narzędzia, które lubimy, znamy, w których się specjalizujemy, pamiętając o innych, często egzotycznych. Choć większość narzędzi zwraca podobną grupę wyników, to jednak zostały stworzone w różny sposób i mogą dać odpowiedź na to, co było nieosiągalne za pomocą tych „ulubionych".

Tak było w przypadku zdjęcia przetrzymywanej i molestowanej dziewczyny w USA, które opublikowano z jej wizerunkiem. W żadnym znanym rejestrze nie udało się porównać jej podobizny, dopiero narzędzie PimEyes[10] wskazało bardzo stare konto na Instagramie, gdzie funkcjonariusze potwierdzili jej imię i nazwisko[11]. Dla mnie kluczowe narzędzia – poza technologią – to: kartka papieru, ołówek, flamastry i linijka.

KROK 4: **Spadochron**

Analiza OSINT może wydawać się banalna, jednak w rzeczywistości taka nie jest. W dobie rozwoju współczesnego Internetu istnieje ogromny szum informacyjny, który przeszkadza we wnioskowaniu. W większości przypadków, z jakimi możemy mieć do czynienia, warto zastosować strategię *out of the box*, by podejść do problemu z każdej możliwej strony. Umysł musi być jak spadochron – będzie działał lepiej, jeśli będzie otwarty. Czasami do celu może doprowadzić rozwiązanie, które wydaje się pozbawione sensu.

Tak było podczas poszukiwań uprowadzonego dziecka, kiedy do celu doprowadziło mnie zdjęcie dziewczynki trzymającej w ręce lizaka w kształcie jednorożca (to unikalny w skali kraju smakołyk, tworzony tylko przez jednego producenta)[12]. Trzeba pamiętać o bardzo ważnej kwestii, jaką jest błąd poznawczy. To ogólne pojęcie irracjonalnego postrzegania rzeczywistości przez błędy w zachowaniu, przekonaniach, podejmowaniu decyzji oraz tendencji do wnioskowania stereotypowego społeczeństwa. Spora grupa tych błędów wpływa jednoznacznie na polaryzację osądów. To jedna z najważniejszych trudności w pracy analityka, o której warto wiedzieć, by móc wydać właściwą opinię.

KROK 5: **Reset, „... czyli tam i z powrotem"**

Nawet bardzo doświadczonym analitykom, znającym wszystkie narzędzia, wykorzystującym każdą znaną technikę, może się zdarzyć, że nie będą w stanie uzyskać żadnej informacji, która pozwoli skutecznie wnioskować i osiągnąć cel. Nie załamujmy się wtedy. Odpocznijmy. Zresetujmy się, przeczytajmy krok czwarty i zacznijmy od nowa. Jeżeli zadanie nie wymaga natychmiastowego rozwiązania, to warto się z tematem „przespać". Jeżeli jesteśmy członkami zespołu – warto porozmawiać o swoich wnioskach z innymi, którzy mają inne spojrzenie na dany temat. Wspólnie, drużynowo zawsze jesteśmy w stanie więcej osiągnąć.

Niesamowitym przykładem jest tutaj historia Marcusa Hutchinsa, hackera, który „uratował Internet", odnajdując przypadkowo wyłącznik ransomware'u WannaCry[13] w 2017 roku, chroniąc sieć przed jeszcze większym potencjalnym rozprzestrzenieniem się tego złośliwego oprogramowania. Został wstępnie źle zidentyfikowany przez FBI jako napastnik, a dopiero powtórna, niezależna analiza wykazała jego niewinność i bohaterstwo[14].

Mapa myśli – płaszczyzna ataku

Zaprezentowane powyżej kroki nie są fragmentem żadnego poradnika (poza tą książką). Zebrałem je jako własne doświadczenie i dla mnie stanowią podstawę dobrej pracy. Każdy jednak może mieć własne przemyślenia na ten temat – i nie ma w tym nic złego. Wręcz przeciwnie. Te kroki to nie „pentalog"; są to raczej wskazówki i pomoc we własnych poszukiwaniach. Najważniejsze jest to, co myślimy i w jaki sposób zrobimy z tego pożytek. Tu z pomocą przychodzi narzędzie, które nazywam **mapą myśli** – czyli **fizyczną i utrwaloną czasowo reprezentacją tego, „co nam w głowie siedzi**".

Mapa myśli w języku angielskim często nazywana jest płaszczyzną ataku (ang. *attack surface*), jednak biorąc pod uwagę fakt, że rekonesans białowywiadowczy często jest zaledwie wstępną fazą ataku, będąc fazą rozpoznania, i wcale atakiem być nie musi, wolę stosować pojęcie mapy. W Internecie można znaleźć mapy dla konkretnych otwartych źródeł. Jest to ogromna pomoc w porządkowaniu własnych myśli i realizacji OSINT-u w sposób usystematyzowany. Jedno z najlepszych takich źródeł stanowi projekt **OSINT Collections** autorstwa użytkownika sinwindie (rysunek 1[15]), dostępny na portalu GitHub[16], a także na stronie OSINT Dojo[17].

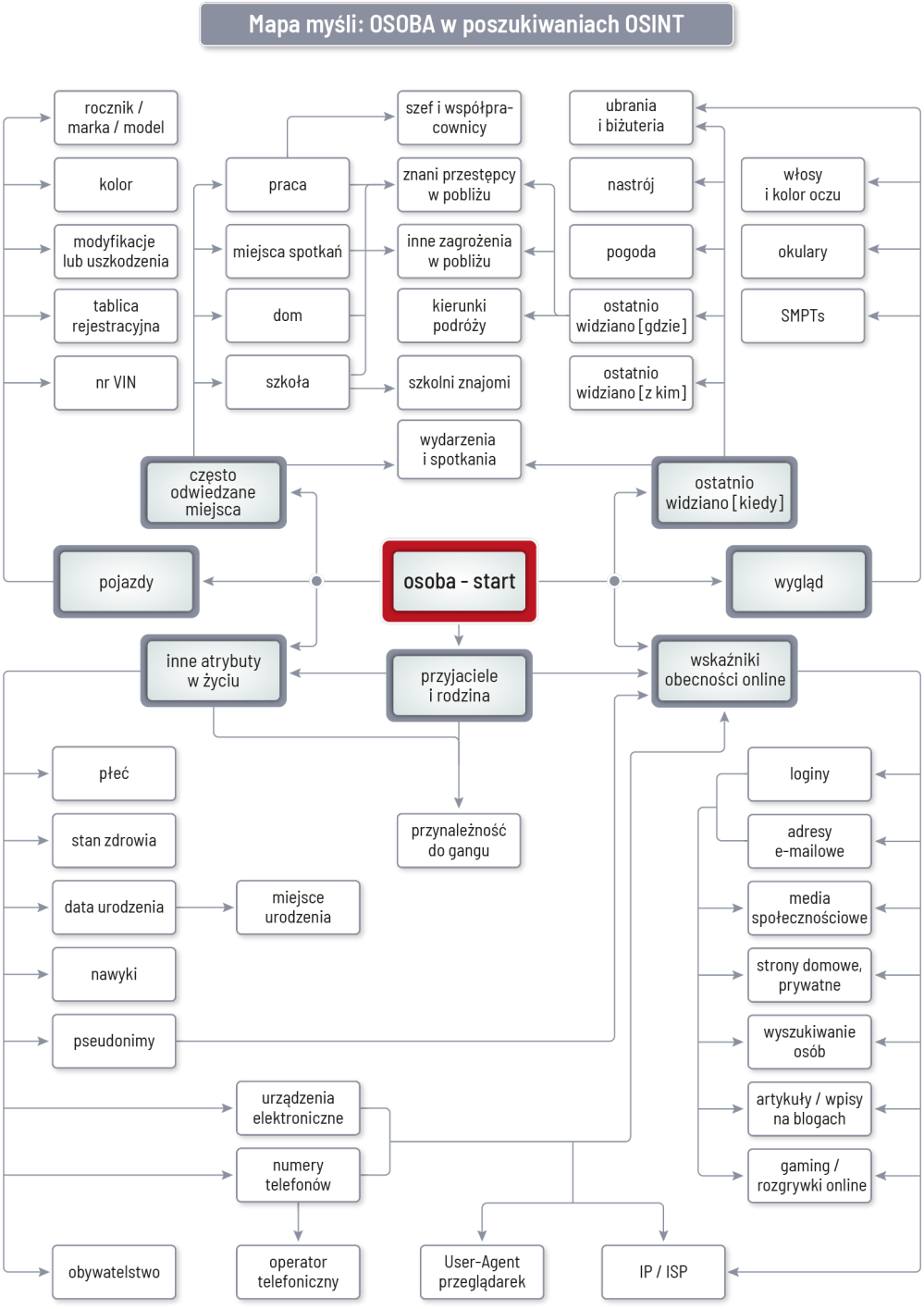
Przykładowa mapa ukierunkowana na poszukiwanie osoby została przedstawiona na rysunku 1 (autorstwa wspomnianego już użytkownika sinwindie). Z własnego doświadczenia mogę tu jeszcze wspomnieć, że ww. mapy są stworzone subiektywnie, choć do celów uniwersalnych. Nie ma złotego środka ani pewności, że zapisany domyślny plan działania sprawdzi się także przy innych poszukiwaniach. Stanowi jednak bazę i podpowiedź, pomoc, by nie przeoczyć jakiegoś obszaru poszukiwań. Ostatecznie, jako dobrą praktykę, proponuję do każdego zadania tworzyć oddzielną mapę, wyspecyfikowaną w odniesieniu do konkretnego celu, posiłkując się przy tym przykładami. Warto do tego wykorzystać interaktywny zbiór narzędzi i linków, **OSINT Framework**[18] oraz **Otwarte Źródła**[19] (dostępny w języku polskim).

Zabezpieczenie anonimowości

Definicja białego wywiadu jasno wskazuje, że polega on na gromadzeniu danych pozyskiwanych z publicznych informacji w sposób etyczny, jednak niemal zawsze jest to pewien rodzaj balansowania na krawędzi prawa. Polski kodeks karny tak definiuje pojęcie prześladowania (ang. *stalking*) w artykule 190a:

> § *§ 1. Kto przez uporczywe nękanie innej osoby lub osoby jej najbliższej wzbudza u niej uzasadnione okolicznościami poczucie zagrożenia lub istotnie narusza jej prywatność, podlega karze pozbawienia wolności do lat 3.*

Każdy, kto chce się zajmować takimi poszukiwaniami, musi mieć na uwadze ten zapis jako dotyczący działań i metod OSINT-u. Jednym ze sposobów ochrony jest możliwość ukrycia swojej lokalizacji oraz tożsamości jako osoby analizującej gromadzone dane. Pod względem lokalizacji warto skorzystać z rozwiązania, jakim jest szyfrowany tunel VPN, który w sytuacji wykrycia wskaże badanemu obiektowi inny adres IP niż ten, z którego łączymy się bezpośrednio z Internetem. Nie jest to jednak forma zapewnienia anonimowości, a jedynie przejście przez kolejny serwer.



Rysunek 1. Przykładowa mapa myśli zorientowana na zbieranie danych o osobie. Czytanie rozpoczynamy od środka – blok `osoba - start`.

Tunel VPN to forma pośredniczenia w dostępie do Internetu dla użytkownika końcowego. Istnieją dwie główne możliwości uzyskania takiego połączenia:

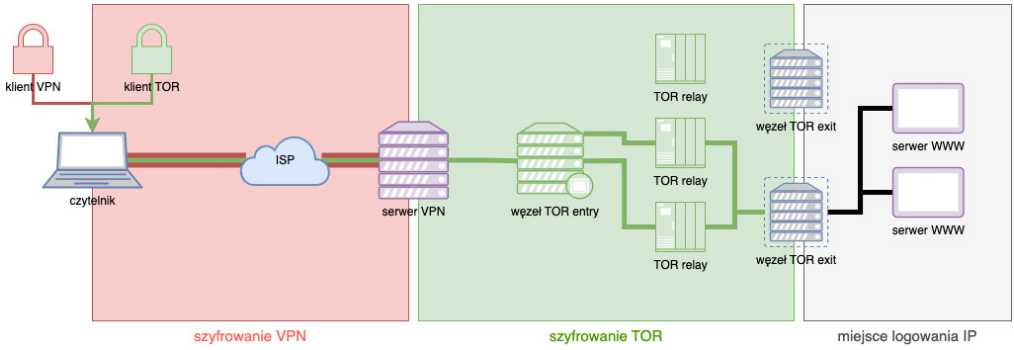
1. Znalezienie odpowiedniego serwera w wyszukiwarce po wpisaniu frazy: *anonymous proxy list* – należy jednak pamiętać, że nie jest to serwer VPN, tylko **serwer pośredniczący** (nie do końca wiadomo, kto i gdzie nim zarządza), a słowo „anonimowy" ułatwia jedynie znalezienie portali z listami adresów IP do użycia. Rozwiązanie z reguły polega na podmianie adresu IP w ustawieniach proxy przeglądarki. Zdarza się jednak, że listy proxy nie są zbyt często odświeżane – i wybrany adres może już nie działać. W tym momencie użytkownik musi ponownie wyłączyć ww. ustawienia, wybrać nowy adres i spróbować powtórnie. Istnieje sposób przyspieszenia tego procesu, czyli rozszerzenie FoxyProxy[20] do popularnych przeglądarek (rysunek 2). Za pośrednictwem definiowanej, importowanej listy z adresami poprzez jedno kliknięcie można przełączać się w ustawieniach proxy przeglądarki bez zbędnego klikania pomiędzy różnymi serwerami.

Rysunek 2. Zrzut ekranu z rozszerzenia FoxyProxy obrazujący łatwy sposób przełączania adresów proxy w przeglądarce

2. Wybór i **zakup tunelu VPN jako usługi** – istnieje bardzo wiele firm zajmujących się sprzedażą usługi **pseudo anonimowego** dostępu do Internetu. Należy jednak pamiętać, że za każdą firmą stoją ludzie, którzy są nastawieni na prowadzenie biznesu, czyli osiąganie zysków. O ile w większości przypadków sama usługa może nie budzić zastrzeżeń, to jednak w sytuacji, gdy uprawnione do tego służby zwrócą się z prośbą do właścicieli serwerów o przekazanie logów połączeń, sytuacja staje się patowa, a usługa „anonimowego" dostępu okazuje się nie do końca *incognito*. Oczywiście tylko wtedy, gdy działania OSINT-owe są prowadzone nielegalnie…

Z uwagi na fakt, że sama usługa VPN nie zapewnia subskrybentom prawdziwej anonimowości, najrozsądniejszym rozwiązaniem zwiększenia poziomu anonimowości jest **wykorzystanie usługi kombinowanej, czyli użycie VPN i sieci Tor**. W momencie korzystania wyłącznie z sieci Tor dostawca internetu ISP (ang. Internet Service Provider) może zweryfikować, czy transmisja użytkownika wykorzystuje tę sieć na zasadzie P2P (ang. *peer-to-peer*). Istnieją dwie możliwości współpracy tych technologii, w zależności od tego, która będzie użyta jako pierwsza:

1. ***Tor over VPN*** – w przypadku podłączenia się do usługi VPN jako pierwszej najpierw ruch jest szyfrowany od dostawcy VPN, a następnie przechodzi przez sieć Tor. Podstawową zaletą takiego rozwiązania jest to, że ISP nie wie, że korzystamy z sieci Tor. Także węzeł wejściowy sieci Tor nie zna naszego prawdziwego adresu IP. Dodatkowo można wchodzić na strony w adresacji *.onion, które będziemy omawiać w dalszej części tego rozdziału. Wadą tego rozwiązania jest jednak fakt, że dostawca VPN ma dostęp do naszego prawdziwego adresu IP. Również część serwerów WWW wykrywa ruch Tor i blokuje dostęp, automatycznie klasyfikując go jako podejrzany.
2. ***VPN over Tor*** – jak przedstawiono na rysunku 3, w sytuacji odwrotnej, gdy najpierw łączymy się z siecią Tor, a dopiero potem przez serwer VPN, główną zaletą rozwiązania jest fakt, że dostawca VPN nie zna naszego prawdziwego adresu IP, ale ISP widzi, że korzystamy z sieci Tor (co nie jest nielegalne). Odwiedzane strony internetowe nie wiedzą, że korzystamy z sieci Tor, więc nie blokują dostępu. Wadą tego rozwiązania jest jednak fakt, że węzeł wejściowy do sieci Tor może poznać nasz prawdziwy adres IP. Mimo to ten sposób działania wciąż wydaje się bezpieczniejszy aniżeli znajomość prawdziwego adresu IP przez firmę oferującą usługę VPN.



Rysunek 3. Schemat połączenia *VPN over Tor*

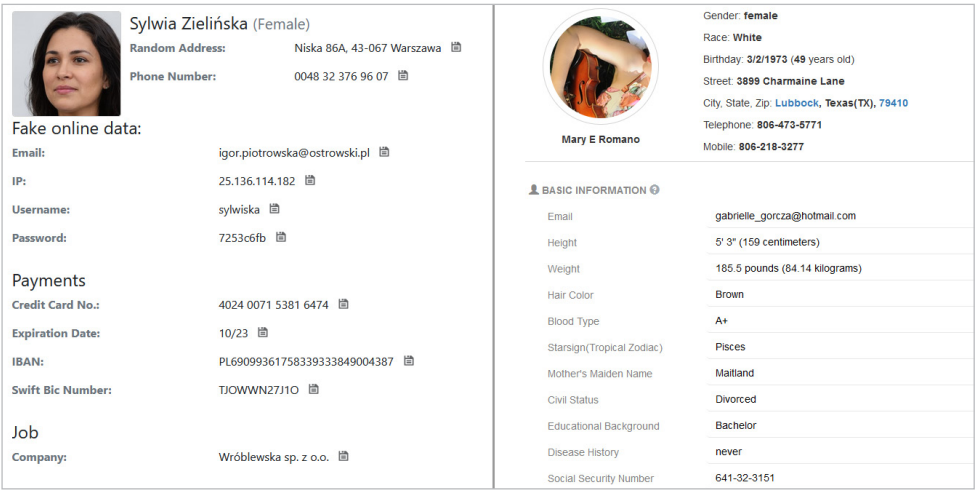
Na listingu 1 przedstawiono przykładową konfigurację połączenia wykorzystującego sieci VPN oraz Tor w systemie Linux. Warto zwrócić uwagę na znak &, którego użycie powoduje przeniesienie procesu w tło i odblokowanie konsoli do dalszej pracy.

Listing 1. Instalacja usług sieci Tor, VPN oraz proxychains wraz z konfiguracją pliku `torrc`

```
# sudo apt-get install tor proxychains openvpn
# sudo vi /etc/tor/torrc

    ORPort 9001
    ExitPolicy reject *.*
    Nickname myrelay
    RelayBandwidthRate 512KB
    RelayBandwidthBurst 1024KB
```

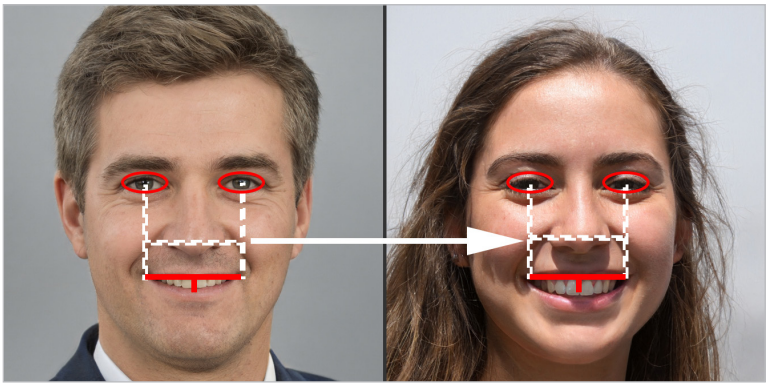
```
    SocksListenAddress 0.0.0.0:9050
    SocksPolicy accept *

# sudo sed -i 's/proxy_dns.*/proxy_dns\tsocks5\t127.0.0.1\t9050/' /etc/
proxychains.conf
# sudo service tor start
# sudo openvpn --config konfiguracja.ovpn &
# proxychains chrome
```

Jedną z form pozatechnicznych zabezpieczenia tożsamości jest stworzenie profilu tożsamości. Można skorzystać tu z wielu portali generujących fałszywe dane personalne (rysunek 4[21]). **Fałszywa tożsamość** w tym wypadku oznacza, że nie jest to kradzież tożsamości prawdziwej osoby, ale stworzenie fikcyjnej. Cel takiego działania także nie może być w konflikcie z obowiązującym prawem. Warto jednak zwrócić uwagę na fakt, że wygenerowane dane często są losowe i nie mają ze sobą związku. Jak przedstawiono na rysunku 4, adresy e-mail nie korelują z danymi personalnymi.



Rysunek 4. Porównanie danych generowanych przy użyciu narzędzi Random Name Generator i Fake Person Generator (brak korelacji między e-mailem oraz imieniem i nazwiskiem)

Istnieje też prosty sposób na wygenerowanie fikcyjnego zdjęcia – jednym z rozwiązań używanych do tego celu jest portal This Person Does Not Exist[22] lub Generated Photos[23]. W obu przypadkach wprawne oko doświadczonego analityka dostrzeże pewne wskazówki sugerujące wykorzystanie generatorów fotografii. W przypadku zdjęć z pierwszego serwisu (rysunek 5) każda wygenerowana twarz ma parę oczu usytuowaną dokładnie w tym samym miejscu grafiki. Niezależnie od płci, wieku, skręcenia profilu twarzy. Generated Photos, dzięki możliwości personalizacji cech twarzy, pozwala uzyskać lepszy efekt. Trzeba jednak bardzo uważać, by nie przesadzić, bo ktoś pozbawiony zmysłu estetycznego (lub bardzo dobrego monitora) może nie zauważyć, że postać, którą wygenerował, wygląda nienaturalnie. Ten serwis jest płatny.



Rysunek 5. Porównanie obrazów osób wygenerowanych w narzędziu This Person Does Not Exist; pozycja oczu i ust jest zawsze w tym samym miejscu

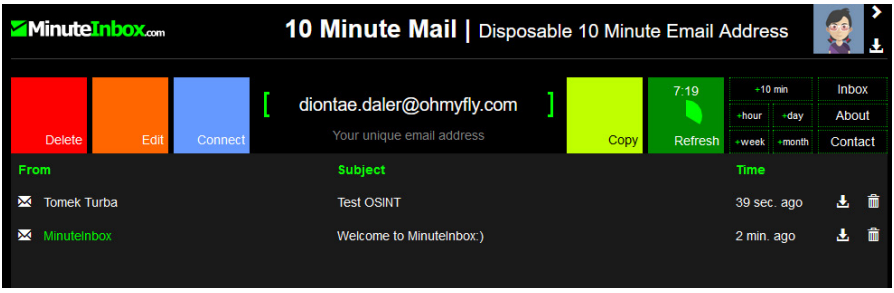
Obok wygenerowania tożsamości równie niezbędnym elementem jest **anonimowy adres e-mail**. Można go pozyskać na wiele sposobów; poniżej wskażę trzy:

1. **Kolejne konto** – można założyć zupełnie nowe konto e-mail na portalu, który oferuje takie konta za darmo w publicznym Internecie. Warto zadbać o to, by konto nie zawierało w swojej nazwie słowa, cechy, z którymi można powiązać właściciela. Jeżeli cel jest niezaawansowany technologicznie lub oczekuje konkretnego zachowania, dobrze jest skorzystać z kont znanych portali, takich jak: wp.pl, onet.pl, interia.pl, gdyż w większości przypadków nie budzą one podejrzeń. Mogą także sugerować, że właściciel konta nie jest osobą zaawansowaną technicznie. W sytuacjach o bardziej skomplikowanym charakterze można utworzyć konta w serwisach Proton Mail lub CTemplar, które oferują szyfrowanie. Szyfrowanie wiadomości e-mail polega na zabezpieczeniu jej treści przed nieautoryzowanym odczytem. Przed wysłaniem e-maila jego treść jest zaszyfrowana przy użyciu algorytmów szyfrowania, takich jak AES+{ECC,RSA}[24]. Odbiorca, który dysponuje odpowiednim kluczem deszyfrującym, może odczytać treść wiadomości. Szyfrowanie e-maili pozwala na zwiększenie bezpieczeństwa ich przesyłania i ochronę prywatności użytkownika. Takie szyfrowanie sprawdzi się w sytuacji, gdy zarówno nadawca, jak i odbiorca będą posiadali konta na ww. portalach[25]. Powszechnie stosowanym algorytmem szyfrowania jest PGP (Pretty Good Privacy)*. Minusem rozwiązania typu „kolejne konto" jest fakt, że raz utworzone, może zostać użyte wielokrotnie, do różnych zadań i celów, a nie tylko tego jednego, który wyznaczono. Przez lenistwo. Wielokrotne użycie generuje możliwość powiązania konta z miejscami, w których zostało użyte, tworząc przy tym schemat ułatwiający namierzenie właściciela. Dodatkowym problemem jest możliwość pozyskania adresu IP, z którego właściciel się loguje, z uwagi na jego dostęp – przez publiczny

* PGP, Pretty Good Privacy, to jeden z najpopularniejszych i najstarszych algorytmów szyfrowania stosowanych do zabezpieczenia e-maili. PGP jest standardem otwartym i został opracowany przez Phila Zimmermanna w 1991 roku. PGP korzysta z algorytmów szyfrowania symetrycznego (np. AES) oraz asymetrycznego (np. RSA) do szyfrowania i podpisywania wiadomości e-mail. PGP jest szeroko stosowany przez różne organizacje i indywidualne osoby do zabezpieczenia przesyłanych wiadomości e-mail.

Internet. Najlepiej więc takie konto utworzyć w portalu w sieci Tor (darknet). Teoretycznie w ten sposób zostaje wyeliminowany problem namierzania, gdyż anonimowość tej sieci zapewnia architektura *peer-to-peer*. Serwisy takie jak Proton Mail oraz CTemplar mają swoje odpowiedniki w darknecie.

2. **Konto tymczasowe z generatora** – w większości przypadków konto e-mail będzie potrzebne do rejestracji i jej potwierdzenia na potrzeby wejścia w strefę deep web* danego portalu lub w celu uzyskania możliwości użycia narzędzia wymagającego rejestracji. Takie konto będzie więc potrzebne na parę minut do zarejestrowania i aktywacji. Istnieje naprawdę dużo serwisów, które można namierzyć w wyszukiwarce po wpisaniu frazy „temporary e-mail", uzyskując w wynikach listę z wieloma domenami, które w opisie zawierają np.: „10 minute e-mail". Trudność pojawia się wtedy, gdy chcemy jednak używać takiego adresu e-mail dłużej (np. próba nawiązania interakcji z badanym obiektem). Z rozwiązaniem tego problemu przychodzi serwis MinuteInbox[26] oferujący nieodpłatnie możliwość przedłużenia konta nawet na miesiąc (rysunek 6). Niestety, tak jak w przypadku dodatkowej skrzynki omówionej powyżej, usługa nadal dostępna jest w publicznym Internecie. Warto więc podobne wyszukiwanie wykonać w sieci Tor (z oczywistych względów tymczasowości sieci nie ma sensu podawać tutaj linka, który w momencie czytania tego tekstu prawdopodobnie już dawno nie będzie aktywny).



Rysunek 6. Zrzut ekranu ze skrzynki e-mail z narzędzia MinuteInbox.com

3. **Konto „szyte na miarę"** – nie ma rozwiązania idealnego, ale tajemnicą sukcesu jest dostosowywanie narzędzi do potrzeb i celów, jakim mają służyć. W związku z tym trzeba pamiętać o podstawowej zasadzie: efekt jest wprost proporcjonalny do wysiłku i zaangażowania. Najrozsądniejszym podejściem wydaje się więc stworzenie konta e-mail w powszechnie używanym serwisie oferującym takie usługi w sieci Tor, ale z zachowaniem dewizy: jedno zadanie – jedno konto. Obowiązkowe w takiej sytuacji jest jednak stworzenie listy z dostępami do poszczególnych kont oraz zapisanie, które konto do jakiego zadania zostało przypisane.

* Deep web, głęboki Internet, to część sieci Internet, która nie jest dostępna dla standardowych wyszukiwarek internetowych, takich jak Google czy Bing. Deep web składa się z różnych stron internetowych, baz danych, usług i aplikacji, które nie są indeksowane przez wyszukiwarki i są trudno dostępne dla przeciętnego użytkownika. Więcej na ten temat w dalszej części rozdziału.

PRZEGLĄD OTWARTYCH ŹRÓDEŁ

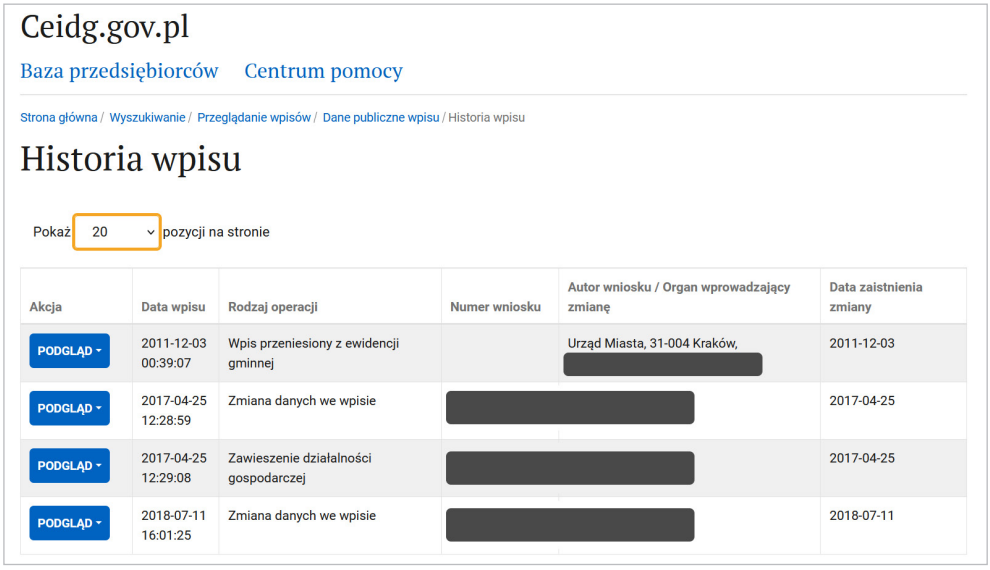
W czasie pisania tego rozdziału wszystkie wymienione źródła lub narzędzia były dostępne i dostosowane do rodzaju konkretnego zadania. OSINT podlega jednak ciągłej ewolucji. Narzędzia zmieniają się, tak jak zmienia się technologia na świecie. Zaczynają nawet powstawać firmy i modele biznesowe oparte na działaniach OSINT-owych, sprzedawanych w formie usługi[27]. Niestety – większość narzędzi staje się płatna. Niemniej jednak warto przyjrzeć się poniższej liście jako pewnego rodzaju przykładowi, który i za kilka lat może posłużyć jako fundament do poszukiwania nowszych narzędzi, opartych na podobnym sposobie procedowania i myślenia o informacji.

Imię i nazwisko, adres e-mail, adres fizyczny

Bellingcat, grupa dziennikarska specjalizująca się w analizie danych, opublikowała raport na temat metodologii działania Federalnej Służby Bezpieczeństwa (FSB) Rosji, której użyto do próby otrucia opozycyjnego polityka Aleksieja Nawalnego. Raport opiera się na analizie danych telefonicznych, geolokalizacji i innych materiałów, które wykazały, że FSB miała kontrolę nad tym, co robi polityk, od samego początku jego roli opozycjonisty. W artykule możemy przeczytać, że w krajach Europy Wschodniej nie kładzie się dużego nacisku na bezpieczeństwo danych osobowych – w przeciwieństwie do regulacji europejskich. I tak dziennikarze z Bellingcat zaledwie w kilka minut po konfiguracji bota w Telegramie i po drobnej opłacie na czarnym rynku uzyskali szczegółowe informacje o potencjalnych agentach FSB zaangażowanych w akcję, m.in. dane telefoniczne, geolokalizacyjne, profile społecznościowe i konta w portalach handlowych[28]. Wszystko to okazało się dostępne w Internecie. Powyższa historia pokazuje, jak bardzo kluczowe znaczenie ma OSINT w poszukiwaniu osób.

W polskiej rzeczywistości, w świetle regulacji prawnych w kwestii ochrony danych osobowych, tak łatwo już nie jest. W opracowaniach poświęconych białemu wywiadowi często można spotkać stwierdzenie, że warto zaopatrzyć się w starą książkę telefoniczną, która zawiera dane osobowe umieszczone tam przed istotnymi zmianami w regulacjach ochrony danych. Książka ta niegdyś była nieodpłatnie wysyłana do każdego abonenta. Niestety, z praktycznego punktu widzenia można stwierdzić, że prawdopodobieństwo odniesienia sukcesu śledczego w wypadku posługiwania się tego typu narzędziem dzisiaj jest niewielkie. Przede wszystkim należy wziąć pod uwagę fakt, że ostatnie wydania liczą już około 20 lat i dotyczą wyłącznie telefonów stacjonarnych. Oznacza to, że ówczesny trzydziestolatek, który był wpisany do ewidencji, obecnie ma mniej więcej pięćdziesiąt lat, a na dodatek prawdopodobnie zrezygnował z numeru stacjonarnego.

Dzisiaj do skutecznego odnalezienia adresu, numeru telefonu lub adresu e-mail należy wykorzystywać techniki łączone. Jeżeli mamy do czynienia z osobą, która prowadzi firmę, istnieje prawdopodobieństwo, że dane telefoniczne znajdziemy w wielu bazach, a powinniśmy zacząć od rejestru Centralnej Ewidencji i Informacji o Działalności Gospodarczej (CEIDG)[29]. Jego ciekawą funkcjonalnością, poza przekazaniem danych adresowych, jest możliwość przeglądania historycznych zmian wpisu, w których czasami znajduje się znacznie więcej danych (rysunek 7). Adresów e-mail lub numerów telefonów, których nie ma w rejestrze przedsiębiorców, można szukać w innych portalach gromadzących dane, m.in.: LinkedIn, Allegro, OLX, Vinted.



Rysunek 7. Zrzut ekranu z danymi przedsiębiorcy z przykładowymi wpisami historycznymi

Sprawa nieco bardziej się komplikuje w przypadku poszukiwania prywatnego adresu e-mail, który użytkownicy często zastrzegają lub po prostu w ogóle go nie podają. Ewentualnie mogą go podawać w postaci zakodowanej w formacie: login (at) domena. Taki zapis nie pozwoli crawlerom[30] go zaindeksować i tym samym zaprezentować jako wyniku wyszukiwania. Autorzy wyszukiwarek znają jednak tę metodę od lat i niestety nie jest ona już w żaden sposób skuteczna. Jedyny sprawdzony obecnie sposób na ominięcie indeksowania to użycie nawiasów kwadratowych, np. proton[.]me (chociaż chronią jedynie przed przypadkowym kliknięciem w złośliwy link). Warto posłużyć się tu narzędziami automatyzującymi, do których zaliczamy popularne: theHarvester[31], Maltego[32] czy SpiderFoot HX[33]. Najefektywniejszym narzędziem przeznaczonym do tego typu poszukiwań jest jednak stary, dobry mailcat[34]. Zasada jego działania (listing 2) polega na próbie odpytania danego serwera SMTP, czy login@domena istnieje poprzez wykorzystanie szczątkowej informacji z próby rejestracji, próby odzyskania hasła lub dostępu do API danego usługodawcy. Lista obsługiwanych przez narzędzie domen jest bardzo długa i zawiera większość znanych, darmowych serwerów pocztowych, także polskich (Gmail, o2, Interia, Proton Mail, iCloud itd.). Narzędzie może zostać uruchomione także za pośrednictwem sieci Tor.

Listing 2. Instalacja i uruchomienie narzędzia mailcat

```
# git clone https://github.com/sharsil/mailcat
# cd mailcat
# python3 -m pip install -r requirements/base.txt
# ./mailcat.py --tor username
# proxychains4 -q python3 mailcat.py username
```

Rozwiązanie to, w połączeniu z narzędziami z serwisów Hunter (sprawdzanie pracowników w obrębie domeny) oraz Have I Been Pwned[35] (sprawdzanie, gdzie dany adres e-mail lub numer telefonu został wykorzystany do założenia konta, które w pewnym momencie znalazło się w bazie jednego z wycieków), stanowi dosyć kompletne źródło informacji na temat cyfrowego rysopisu i profilu użytkownika.

Fotografie i metadane

Pewne stare przysłowie mówi, że jeden obraz jest wart więcej niż tysiąc słów – i jest to prawda także w świecie OSINT-u. Często zdarza się, że uda się namierzyć lokalizację podmiotu lub obiektu wyłącznie na podstawie cech jednego zdjęcia. Źródłem informacji mogą być metadane zdjęcia albo to, co faktycznie na nim widać (numer rejestracyjny, budynek, ulica, ludzie, gazeta, a nawet kolor i wzór na dywanie). W przypadku gdy fotografia została gdzieś opublikowana i nie pozbawiono jej metadanych, można wykorzystać narzędzia typu ExifTool[36], przekazujące informacje o zdjęciu w sposób zrozumiały dla przeciętnego użytkownika (listing 3).

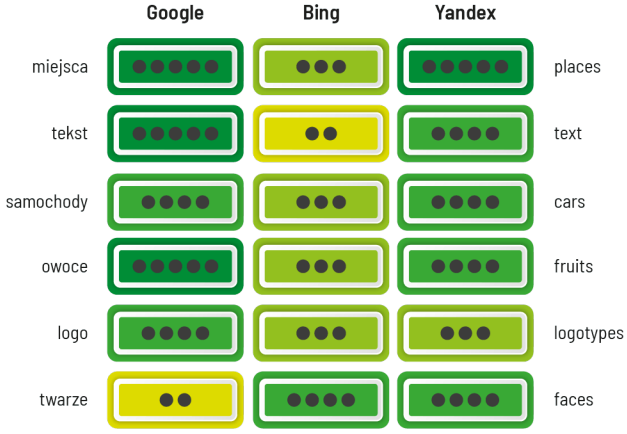
Listing 3. Wynik użycia narzędzia ExifTool na zdjęciu z zawartymi metadanymi (wynik skrócony do najważniejszych parametrów)

```
# exiftool DSCN0042.jpg\?raw\=true
ExifTool Version Number         : 11.85
File Name                       : DSCN0042.jpg?raw=true
Directory                       : .
File Size                       : 153 kB
File Modification Date/Time     : 2023:01:22 13:49:48+00:00
File Access Date/Time           : 2023:01:22 13:49:48+00:00
File Inode Change Date/Time     : 2023:01:22 13:49:48+00:00
File Permissions                : rw-r--r--
File Type                       : JPEG
File Type Extension             : jpg
MIME Type                       : image/jpeg
Exif Byte Order                 : Little-endian (Intel, II)
Image Description               : Sample image for OSINT purpose
Make                            : NIKON
Camera Model Name               : COOLPIX P6000
[…]
Modify Date                     : 2008:11:01 21:15:11
[…]
Date/Time Original              : 2008:10:22 17:00:07
Create Date                     : 2008:10:22 17:00:07
[…]
GPS Latitude Ref                : North
GPS Longitude Ref               : East
GPS Altitude Ref                : Above Sea Level
GPS Time Stamp                  : 14:57:41.37
```

```
GPS Satellites                  : 04
GPS Map Datum                   : WGS-84
GPS Date Stamp                  : 2008:10:23
[…]
GPS Date/Time                   : 2008:10:23 14:57:41.37Z
GPS Latitude                    : 43 deg 27' 52.04" N
GPS Longitude                   : 11 deg 52' 53.32" E
Circle Of Confusion             : 0.006 mm
Field Of View                   : 28.8 deg
Focal Length                    : 15.0 mm (35 mm equivalent: 70.0 mm)
GPS Position                    : 43 deg 27' 52.04" N, 11 deg 52' 53.32" E
[…]
```

Świetnym przykładem skutecznej analizy metadanych jest historia identyfikacji personelu rosyjskiej armii odpowiedzialnego za zestrzelenie samolotu malezyjskich linii lotniczych w 2014 roku. Głównym źródłem informacji dziennikarzy śledczych były zdjęcia i filmy opublikowane w mediach społecznościowych przez rosyjskich żołnierzy i jednostki wojskowe. Pomimo faktu, że portale społecznościowe z reguły wycinają metadane z zamieszczonych plików, publikujący materiały żołnierze nie zwrócili uwagi na automatycznie dodane informacje o dacie i miejscu wykonania zdjęć. Pozwoliło to na ustalenie trasy przejazdu konwoju wojskowego, który przemieszczał się w pobliżu miejsca startu rakiety. Dzięki analizie zdjęć i filmów dziennikarze z Bellingcat byli w stanie znaleźć znaczące punkty orientacyjne na mapie, takie jak: mosty, skrzyżowania drogowe czy budynki. Pozwoliły one na precyzyjne określenie miejsca, z którego startowała rakieta. Analiza metadanych zdjęć i filmów umożliwiła także identyfikację indywidualnych żołnierzy i jednostek wojskowych, które były zaangażowane w operację. Dzięki tym danym udało się ustalić, że konwój był związany z rosyjskim batalionem rakietowym, który odpowiadał za zestrzelenie MH17[37].

Co jednak w sytuacji, gdy zdjęcie nie posiada metadanych? Zgodnie z przysłowiem z pierwszego akapitu tego podrozdziału można stwierdzić, że zdjęcie samo w sobie reprezentuje już dane. Analiza tych danych bardzo często pozwala na niezwykle szczegółowe wnioskowanie. Tak było w przypadku zatrzymania dilera narkotyków przez walijską policję. U zatrzymanej osoby, podejrzanej o posiadanie narkotyków, na komunikatorze WhatsApp na smartfonie znaleziono zdjęcie tabletek ecstasy na dłoni dilera. Zdjęcie było na tyle dobrej jakości, że analitycy byli w stanie porównać linie papilarne widoczne na fotografii ze swoją bazą odcisków. Wraz z dodatkowymi metodami OSINT pozwoliło to śledczym odkryć tożsamość dilera i ją potwierdzić po zatrzymaniu go[38]. Zdjęcie bez metadanych można także sprawdzić w znanych wyszukiwarkach oferujących funkcję *Reverse Image Search* (czyli wyszukiwanie obrazem), np. takich jak: Google, Bing czy Yandex. Każda z wyszukiwarek ma odrębne cechy i każda zupełnie inaczej przedstawia skatalogowane wyniki. Cechy te zostały zbadane i pokazane na rysunku 8[39].



Rysunek 8. Ranking porównawczy wyszukiwarek oferujących funkcję *Reverse Image Search* (im więcej kropek, tym lepszy wynik)

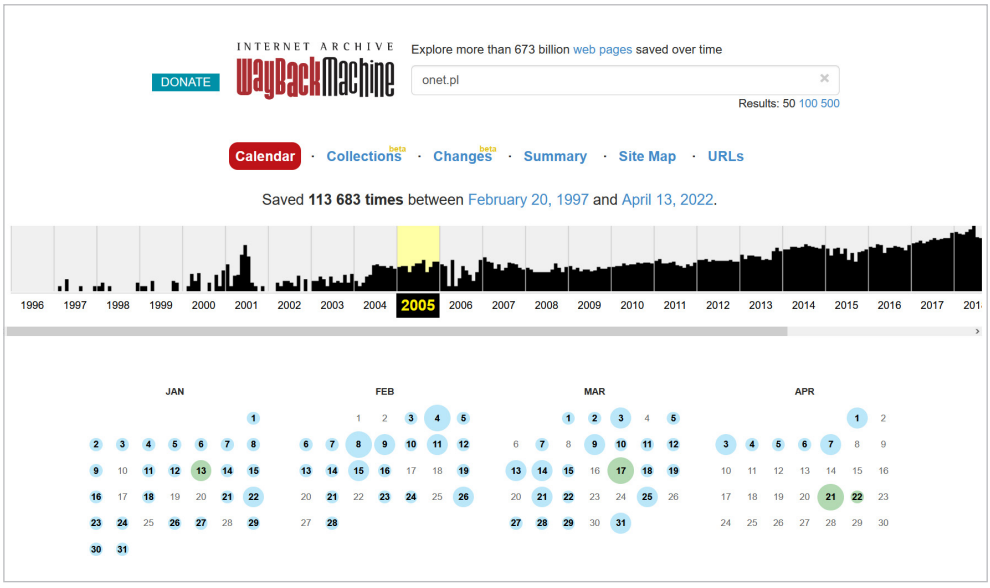
W 2021 roku wyniki przedstawione powyżej wyglądałyby zupełnie inaczej. Firma Google odrobiła jednak lekcję i wprowadziła zaawansowany mechanizm Google Lens do wyszukiwarki grafiki. Teraz znacznie łatwiej można wyszukiwać miejsca, teksty, przedmioty, lecz Google Images nadal dosyć słabo radzi sobie z rozpoznawaniem twarzy – w przeciwieństwie do wyszukiwarek Yandex i Bing. Podczas wyszukiwania wizualnego, gdzie jest możliwość zmiany wielkości zaznaczonego obszaru zdjęcia do analizy, warto zwrócić uwagę, jak bardzo wyniki mogą się różnić odnośnie do różnych ustawień obszaru. Czasami nawet niewielka zmiana w tym zakresie może mieć duży wpływ na zwracane wyniki. Warto też w tym przypadku zapisywać interesujące nas zestawy wyszukanych danych, gdyż z doświadczenia mogę powiedzieć, że niekiedy naprawdę trudno jest tak zmienić obszar zaznaczenia, aby trafić w wyniki, które jeszcze przed chwilą udało nam się uzyskać.

Archiwa

Flagowym przykładem użycia danych archiwalnych w śledztwach OSINT jest historia Międzynarodowego Konsorcjum Dziennikarzy Śledczych (International Consortium of Investigative Journalists, ICIJ), którzy wykorzystali publiczne dane finansowe do ujawnienia tzw. Panama Papers w 2016 roku. Dziennikarze ICIJ uzyskali dużą ilość danych finansowych panamskiej kancelarii prawnej Mossack Fonseca, specjalizującej się w tworzeniu spółek zagranicznych i korporacji. Zapisy, które sięgały 1970 roku, zawierały informacje o ponad 200 000 firm, fundacji i funduszy powierniczych. ICIJ użyło tych danych do śledztwa i ujawnienia wykorzystywania tych spółek przez bogatych ludzi i urzędników państwowych na całym świecie do unikania opodatkowania poprzez pranie brudnych pieniędzy i ukrywanie swoich aktywów. Śledztwo ujawniło szeroki zakres nielegalnych działań i wykazało udział w nich polityków, urzędników państwowych i najbogatszych ludzi z całego świata. To dochodzenie jest uważane za jedno z największych i najważniejszych w ostatnich latach i pokazuje, że dane archiwalne mogą być cennym źródłem informacji do odkrywania przestępstw[40].

Istną kopalnią wiedzy historycznej jest niewątpliwie archiwum Wayback Machine[41]. Po wpisaniu adresu URL w wyszukiwarce użytkownik otrzymuje możliwość obejrzenia i interakcji ze stroną z danego, historycznego punktu w wybranym czasie oferowanym przez Wayback Machine (rysunek 9). W większości przypadków na takich stronach możemy znaleźć informacje sprzed czasów RODO/GDPR, gdy nie zwracano szczególnej uwagi na udostępniane dane osobowe lub organizacyjne. Często strony takie prezentują linki do stron domowych użytkowników, które już *de facto* nie istnieją, ale ich kopia została zachowana w Wayback Machine. Można tam niejednokrotnie odnaleźć szczegółowe dane prywatne o osobie, jej rodzinie czy miejscu zamieszkania wraz z fotografiami, opisami, a nawet informacją o hobby.

W 2017 roku wspomniane konsorcjum ICIJ uzyskało dużą ilość danych finansowych od bermudzkiej kancelarii prawnej Appleby. Dane te obejmowały informacje o ponad 100 000 firm i funduszy, sięgające aż 1950 roku. Zespół ICIJ użył Wayback Machine, aby uzyskać dostęp do historycznych wersji strony internetowej kancelarii prawnej. Stara strona kancelarii zawierała informacje o jej usługach i klientach. Konsorcjum udało się ustalić, że niektórzy z nich byli zaangażowani w działalność nielegalną, taką jak unikanie opodatkowania, pranie brudnych pieniędzy i inne przestępstwa finansowe. Śledztwo, określone jako *Paradise Papers*, ujawniło, w jaki sposób bogaci ludzie i firmy korzystają ze spółek zagranicznych i z funduszy w celu ukrycia swoich aktywów oraz uniknięcia opodatkowania. Afera skutkowała naciskiem na zwiększenie przejrzystości w całym systemie finansowym[42].

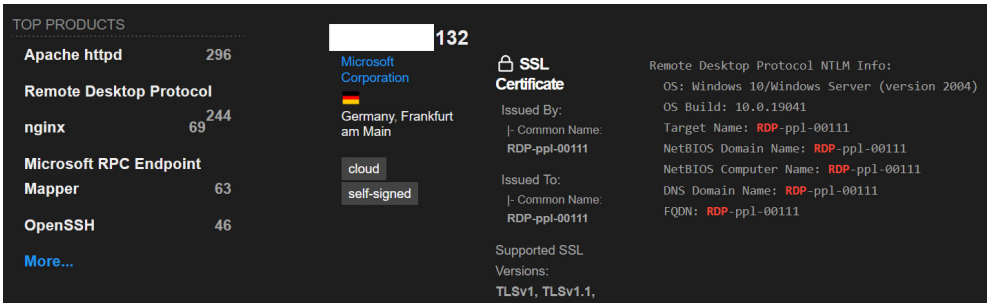


Rysunek 9. Kalendarz z możliwością wyboru migawki strony onet.pl według wybranej daty

Infrastruktura

Shodan[43] lub ZoomEye[44] to dwie specyficzne wyszukiwarki urządzeń podłączonych do Internetu, w tym IoT (Internet of Things, internet rzeczy). Za pośrednictwem zestawu filtrów podobnych do Google Dorks można przewertować ogromne bazy metadanych serwerów, takich jak: informacje nagłówka odpowiedzi serwera, wiadomość i banner MOTD[45], informacje o oprogramowaniu serwera i otwarte porty, takie jak np.: WWW (HTTP/HTTPS – porty: 80, 8080, 443, 8443), a także FTP (port 21), SSH (port 22), Telnet (port 23), SNMP (port 161), IMAP (port 143 lub [szyfrowane] 993), SMTP (port 25), SIP (port 5060) i Real Time Streaming Protocol (RTSP, port 554). To tylko niewielka grupa portów „znanych", którą można znaleźć w obszernych bazach wyszukiwarek. Tak naprawdę dzisiaj nie ma już znaczenia numer portu, bo tak samo można odnaleźć informację nt. portów wysokich, np. 31337. W przypadku wykonywania rekonesansu infrastruktury na podstawie nazwy przedsiębiorstwa lub adresu IP możemy posłużyć się tymi wyszukiwarkami w celu sprawdzenia aktualnych i historycznych danych o udostępnianych usługach. Warto także w celach defensywnych i rozpoznawczych sprawdzić w wyszukiwarce swój adres IP lub całą klasę – w przypadku organizacji. Dobrze jest również zapoznać się z podstroną Filter Reference[46], na której prezentowane są przykładowe filtry pozwalające na zawężanie wyników wyszukiwania.

Na rysunku 10 znajduje się wynik użycia przykładowego filtru `"port:3389 country:DE"`, pokazujący urządzenia zlokalizowane w Niemczech, z otwartą usługą RDP*. O tym, jak bardzo te narzędzia są istotne w kwestiach cyberbezpieczeństwa, świadczy wyciek danych 200 milionów amerykańskich wyborców z 2017 roku. Firma Deep Root Analytics, która była wspierana przez Republikański Komitet Narodowy, umieściła plik bazy danych w nieprawidłowo skonfigurowanym kontenerze S3 Amazon. Okazało się, że baza jest dostępna publicznie i można było ją odnaleźć, używając kwerendy: `"amazon s3 bucket" filetype:sql` w serwisie Shodan[47].

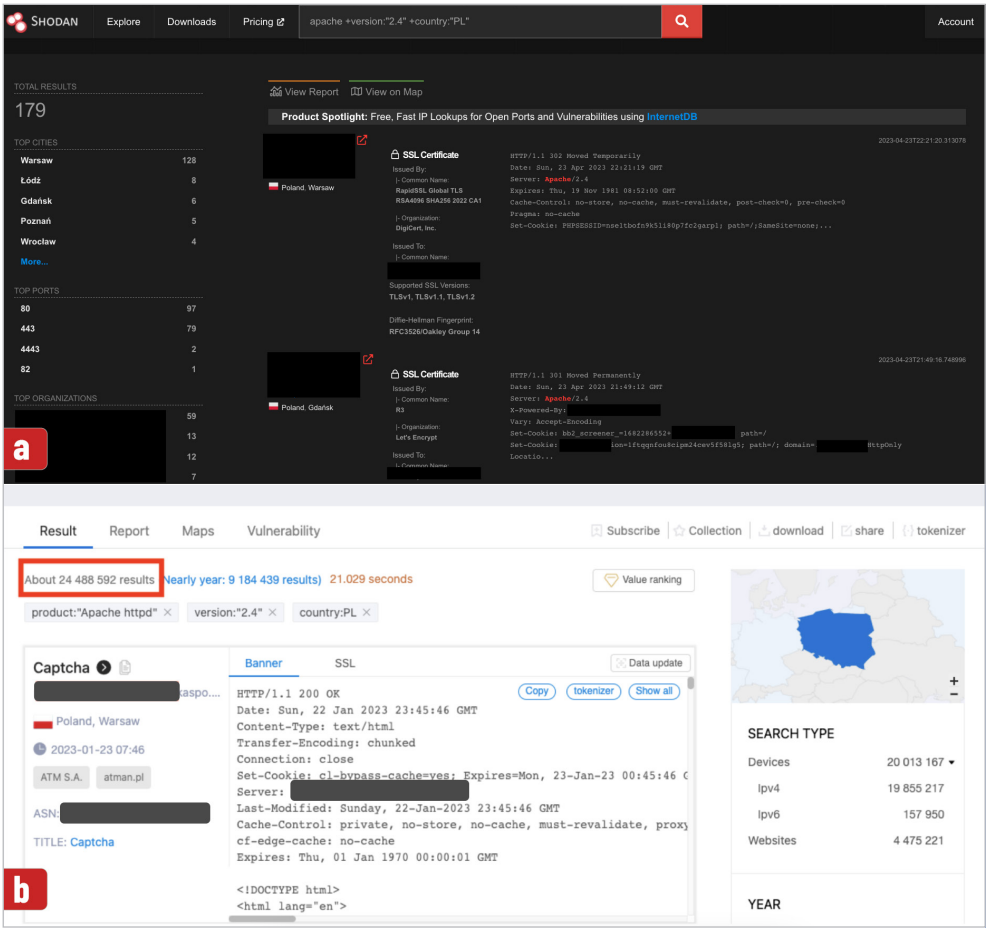


Rysunek 10. Zrzut ekranu z wyszukiwarki Shodan prezentujący znaleziony serwer z wykorzystaniem filtra na port 3389 oraz kraju DE

ZoomEye jest podobnym narzędziem do Shodana, jednak istnieją pewne różnice między nimi, wskazujące, że to drugie rozwiązanie może być nieco lepsze – w zależności od badanej sytuacji. ZoomEye oferuje bardziej zaawansowany sposób wyszukiwania za

* MOTD (*message of the day*) to komunikat powitalny serwera.

pomocą wielu kryteriów, takich jak: porty, protokoły, słowa kluczowe, adresy IP, zakres dat. Jednocześnie zwracane wyniki wydają się bardziej szczegółowe. Narzędzie to ma jednak dosyć duże ograniczenia wyświetlania wyników w wersji darmowej – nie działa ono zbyt precyzyjnie i liczba odnalezionych urządzeń jest przekłamana, bo w domyślnym wyszukiwaniu zawiera także wszystkie dane historyczne pasujące do zapytania. Samo narzędzie pozwala w wersji darmowej na wyświetlenie 10 000 wyników[48]. W zależności od zadania zarówno Shodan, jak i ZoomEye mogą mieć swoje odrębne zastosowania. Porównanie wyników dosyć wąskiego zapytania z filtrem odnoszącym się do serwera WWW Apache w wersji 2.4 z Polski zostało przedstawione na rysunku 11 – filtr `product:"Apache httpd" +version:2.4 +country:PL`.



Rysunek 11. Porównanie liczby wyników wyszukiwania za pomocą narzędzi Shodan (a) i ZoomEye (b)

Geolokalizacja

W 2018 roku pojawiły się informacje, że lokalizacje baz wojskowych i tras patrolowych były widoczne na globalnej mapie ciepła aplikacji Strava[49] (rysunek 12). Aplikacja służy do monitorowania aktywności fizycznej użytkowników i rozpowszechniania informacji na ten temat na portalach społecznościowych (np. monitorowanie trasy biegacza lub rowerzysty). Ten rodzaj informacji udostępnianej przez aktywnych fizycznie żołnierzy był jawny na portalu Strava dla każdego, w tym dla potencjalnych przeciwników. W odpowiedzi na te obawy Strava niezwłocznie wprowadziła funkcję pozwalającą użytkownikom na ograniczenie widoczności ich aktywności.

Rysunek 12. Korelacja mapy bazy wojskowej i kont żołnierzy, które były dostępne do podglądu w portalu sportowym Strava

Geolokalizacja jest określeniem dowolnego medium elektronicznego podłączonego do publicznego Internetu w obrębie pewnego rejonu geograficznego wyznaczonego za pomocą długości i szerokości geograficznej. W przypadku posiadaczy smartfonów, smartwatchy lub urządzeń podobnej klasy po włączeniu usługi lokalizowania każde miejsce, do którego się udadzą właściciele tych gadżetów elektronicznych, zostanie zarejestrowane i może zostać nałożone na mapę w odpowiednich systemach, np.: Google Maps, Apple Maps, Yandex Maps, Baidu Maps.

Wyróżnia się trzy **metody geolokalizacji**:

1. **Geolokalizacja na podstawie danych z urządzenia (trilateracja)** – zbieranie danych przy wykorzystaniu modułu GPS wbudowanego w popularne urządzenia typu: smartfon, smartwatch, smartband, laptop, nawigacja oraz całą gamę urządzeń IoT. Metoda trilateracji polega na określeniu położenia urządzenia na podstawie sygnałów od trzech lub więcej nadajników (bazowych stacji komórkowych BTS lub satelitów GPS). Proces polega na tym, że urządzenie jest w stanie namierzyć odległość od co najmniej trzech nadajników, dzięki temu może określić swoje położenie. W tym celu urządzenie (np. smartfon) odbiera sygnały od nadajników

i mierzy czas potrzebny na dotarcie sygnału do urządzenia. Z tego pomiaru czasu i znając prędkość rozchodzenia się sygnału, urządzenie jest w stanie obliczyć odległość do każdego z nadajników. Następnie tworzy trójkąt na podstawie tych odległości i lokalizacji i określa swoje położenie. Im więcej nadajników jest dostępnych, tym dokładniejsze będzie położenie. Dużą zaletą tej metody względem pozostałych jest możliwość określenia położenia w trudnym terenie.

2. **Metoda analizy danych o sieci** – polega na określeniu położenia urządzenia na podstawie danych o sieciach bezprzewodowych (takich jak Wi-Fi lub Bluetooth) w pobliżu tego urządzenia. Kiedy urządzenie jest połączone z siecią Wi-Fi, automatycznie wyszukuje i przechowuje informacje o dostępnych sieciach w pobliżu. Te informacje zawierają nazwę sieci (SSID), adres fizyczny MAC i poziom sygnału. Analiza danych polega na ich porównaniu z danymi o sieciach znajdujących się w bazie danych z informacjami o sieciach na danym obszarze. Metoda jest mniej dokładna niż trilateracja, gdyż opiera się na analizie dostępnych sieci Wi-Fi, których stan może się ciągle zmieniać. Metoda jest często używana w aplikacjach do nawigacji, dostępności przyjaciół w „pobliżu" i monitorowania floty.
3. **Analiza danych o ruchu** – polega na określeniu położenia urządzenia na podstawie danych o ruchu, takich jak: prędkość, kierunek, akceleracja itp. Dane te mogą być pobierane z różnych źródeł, takich jak: akcelerometr, żyroskop, GPS czy moduł kompasowy. Są one analizowane przez algorytm, który jest w stanie określić położenie na podstawie ruchu. Przykładowo jeżeli urządzenie porusza się z prędkością samochodu w mieście (50 km/h) na północny zachód, to algorytm może określić, że urządzenie jest przesuwane w kierunku północno-zachodnim. Metoda jest często wykorzystywana w aplikacjach do monitorowania ruchu, aktywności fitness i śledzenia pojazdów.

Geolokalizacja w OSINT może służyć do określenia położenia geograficznego różnych celów badań, takich jak: osoby, budynki, pojazdy, organizacje. Może być używana do przeprowadzenia następujących czynności:

a. **śledzenia ruchu** – w celu identyfikacji potencjalnych zagrożeń lub przyzwyczajeń;
b. **analizy danych** – analiza z różnych źródeł, takich jak: zdjęcia satelitarne, filmy z dronów lub dane z mediów społecznościowych (mapki);
c. **monitorowania infrastruktury krytyczne**j;
d. **tworzenia map**, które pomagają w identyfikacji i analizie danych geograficznych w analizie trendów, np.: liczby przestępstw, ilości zanieczyszczeń lub określenia skupisk ludzi;
e. **działań marketingowych** – określenia lokalizacji potencjalnych klientów, wspierających planowanie kampanii reklamowych i kierowanie nimi.

Na rysunku 13 przedstawiono zrzut ekranu z narzędzia Google Earth. Jest to bardziej zaawansowana wersja narzędzia Google Maps, ale bez możliwości nawigowania trasy. Pozwala na przeglądanie trójwymiarowej mapy Ziemi z różnymi warstwami danych (zdjęcia satelitarne, modele 3D miast), a także z informacjami historycznymi.

Rysunek 13. Interfejs aplikacji Google Earth

W tabeli 4 zestawiono natomiast popularne narzędzia geolokalizacyjne wraz z ich cechami indywidualnymi.

Tabela 4. Przykładowe narzędzia geolokalizacyjne i ich cechy charakterystyczne dla OSINT

NAZWA MAPY	INFORMACJE ISTOTNE DLA OSINT
Bing Maps https://www.bing.com/maps	► położenie geograficzne różnych celów (budynki, pojazdy, organizacje, osoby) ► różne warstwy danych (mapy satelitarne, modele 3D, informacje o trasach komunikacji miejskiej) ► mapping (tworzenie map do własnych celów)
Google Maps https://www.google.com/maps https://timeline.google.com/maps	► położenie geograficzne różnych celów (budynki, pojazdy, organizacje, osoby) ► różne warstwy danych (mapy satelitarne, modele 3D, informacje o trasach komunikacji miejskiej) ► mapping (tworzenie map do własnych celów) ► większy zasięg danych ► intuicyjny interfejs użytkownika ► dużo opcji wyszukiwania: restauracje, hotele, sklepy itp. ► narzędzia analityczne, np. obciążenie ruchem drogowym, dane demograficzne ► widok Street View ► wyznaczanie odległości ► dostępność w wielu językach

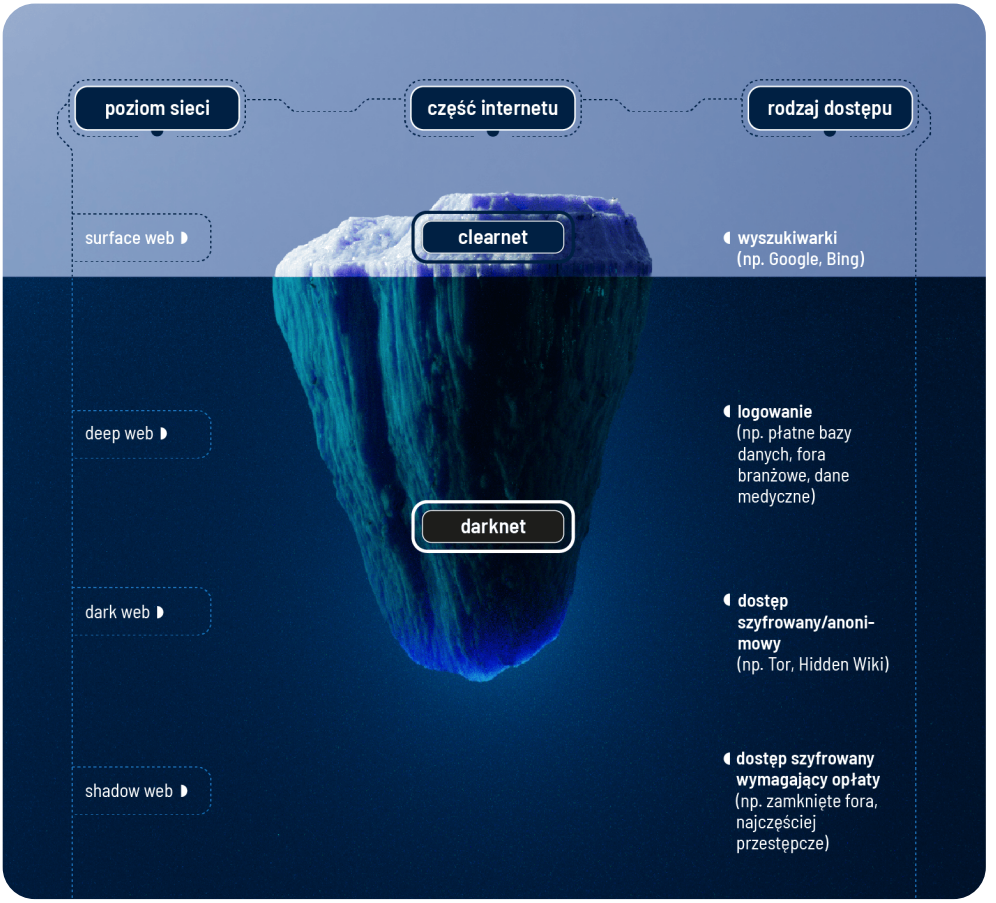
NAZWA MAPY	INFORMACJE ISTOTNE DLA OSINT
Yandex Maps https://yandex.eu/maps	► położenie geograficzne różnych celów (budynki, pojazdy, organizacje, osoby) ► różne warstwy danych (mapy satelitarne, modele 3D, informacje o trasach komunikacji miejskiej) ► mapping (tworzenie map do własnych celów) ► większy zasięg danych ► intuicyjny interfejs użytkownika ► dużo opcji wyszukiwania: restauracje, hotele, sklepy itp. ► narzędzia analityczne, np. obciążenie ruchem drogowym, dane demograficzne ► pochodna funkcji Street View ► mniejszy zasięg danych, głównie Rosja i inne kraje WNP ► ruch transportu publicznego na żywo (tylko Federacja Rosyjska i Białoruś) ► wyznaczanie odległości ► interfejs użytkownika głównie w języku rosyjskim ► część ekosystemu aplikacji Yandex
Baidu Maps https://map.baidu.com	► położenie geograficzne różnych celów (budynki, pojazdy, organizacje, osoby) ► różne warstwy danych (mapy satelitarne, modele 3D, informacje o trasach komunikacji miejskiej) ► mapping (tworzenie map do własnych celów) ► intuicyjny interfejs użytkownika ► dużo opcji wyszukiwania: restauracje, hotele, sklepy itp. ► narzędzia analityczne, np. obciążenie ruchem drogowym, dane demograficzne ► niższa rozdzielczość niż Google Maps (ze względu na priorytetyzację obszaru Chin) ► część ekosystemu aplikacji Baidu
Google Earth https://earth.google.com/web	► zaawansowana wersja Google Maps z widokiem satelitarnym oraz dokładną wizualizacją 3D obiektów ► narzędzia do mierzenia odległości, kątów, powierzchni ► zestaw archiwalnych widoków satelitarnych, obejmujących nawet widok starych przedwojennych zdjęć lotniczych
Map Channels https://www.mapchannels.com	► narzędzie umożliwiające nałożenie na siebie kilku obrazów map (np. widok satelitarny i kartograficzny) i synchronizację poruszania się na obu jednocześnie
Zoom Earth https://zoom.earth oraz Windy https://www.windy.com	► mapa warunków pogodowych na żywo
F4 Map https://demo.f4map.com	► bardzo dokładne odzwierciedlenie obiektów 3D ► bardzo dokładne odzwierciedlenie obiektów zielonych, aktualnych remontów, murów, ścian, linii energetycznych ► prezentacja na żywo warunków pogodowych oraz czasu (np. określenie padania cieni oraz odbicia promieni od tafli wody)
Mapillary https://www.mapillary.com/app	► podobnie jak F4 Map, ale w wersji satelitarnej z materiałami fotograficznymi i wideo udostępnionymi przez użytkowników z kamer samochodowych

NAZWA MAPY	INFORMACJE ISTOTNE DLA OSINT
WikiMapia http://wikimapia.org	► mapa z podziałem na obszary gminne, powiatowe, obiekty – na podstawie danych zebranych z portalu Wikipedia
Satellites Pro https://satellites.pro	► dokładna mapa satelitarna z warunkami pogodowymi
Open Street Map https://www.openstreetmap.org	► darmowa alternatywa względem Google Maps Platform[50] ► dane dostarczane i aktualizowane przez społeczność ► dostępność w wielu językach ► wsparcie dla platform mobilnych oraz programów GIS ► możliwość integracji z własną aplikacją ► lokalizacja planowanych budów ► dokładna pozycja drzew
Travel With Drone https://travelwithdrone.com	► mapa pokazująca multimedia (fotografie, filmy) stworzone za pośrednictwem cywilnych jednostek UAV, ► informacje pobierane z YouTube na podstawie danych geolokalizacyjnych udostępnionych w materiale
Snap Map https://map.snapchat.com	► mapa krótkich filmów z komunikatora Snapchat wykorzystująca zebrane dane geolokalizacyjne
Heavy AI Tweet Map https://www.heavy.ai/demos/tweetmap	► mapa tweetów oraz popularnych hashtagów stworzona na podstawie danych geolokalizacyjnych
One Million Tweet Map https://onemilliontweetmap.com	► mapa tweetów z możliwością filtrowania dat i zbierania dodatkowych statystyk
MapHub https://maphub.net/	► możliwość zaplanowania własnej podróży ► możliwość określenia własnej mapy (oznaczanie, fotografowanie, notatki, przygotowanie trasy)
FlightRadar24 https://www.flightradar24.com	► możliwość weryfikacji jawnego ruchu lotniczego z aktywnym systemem ADS-B
PeakFinder https://www.peakfinder.org PeakVisor https://peakvisor.com	► wizualizacja pasm górskich oraz wysokości n.p.m. do analizy widoczności obiektów
GeoPortal https://www.geoportal.gov.pl	► polski rządowy projekt zawierający m.in. ewidencję gruntów, numery działek

Narzędzia zaprezentowane w tabeli 4 to zaledwie część możliwości, jakie dają nam otwarte źródła. Większość tych linków pojawia się w projekcie GeoHack[51], czyli zmodyfikowanej wersji źródła map Wikipedii, przeznaczonej do wykonywania prostych zmian HTML i dostarczania linków do różnych serwisów mapowych w jednym, zebranym miejscu.

Dane z deep webu i dark webu

Omawianie różnych poziomów Internetu, różnice pomiędzy deep i dark webem oraz clear- i darknetem zwykle opiera się na metaforze góry lodowej dryfującej po oceanie informacji (rysunek 14).



Rysunek 14. Struktura Internetu

Niewielka część nad poziomem wody – to publiczny Internet, czyli clearnet, a jego zawartość, dostępna dla wszystkich przeglądarek, określana jest też jako *surface web*. To, co jest niewidoczne i znajduje się „pod wodą", to tzw. sieć głęboka (ang. *deep web*) z różnymi poziomami w głąb coraz mniej zdefiniowanych obszarów. Czym różni się ona od zawartości „zwykłego" Internetu? Tym, że dostęp do danych nie jest możliwy dla robotów indeksujących, ponieważ np. wymaga uwierzytelnienia po stronie użytkownika lub dostęp do treści tam umieszczonych wymaga rozwiązania CAPTCHA. Zwykle z tego powodu bazy te nie są indeksowane w wyszukiwarkach, ponieważ roboty indeksujące nie mają dostępu do zawartości portalu. Są to np. strony bibliotek, portali naukowych, portali specjalistycznych, ale także… portali społecznościowych! Ze względu na ochronę prywatności większość z nich nie pozwala na indeksowanie danych użytkowników w wyszukiwarkach (chyba

że użytkownik sam wyrazi na to zgodę). Dobrym przykładem portalu znajdującego się w deep webie jest MyHeritage[52], czyli drzewa genealogiczne, lub baza publikacji naukowych IEEE[53]. Portale znajdujące się na poziomie internetowego deep webu zwykle wymagają rejestracji użytkownika. Jeśli więc planujemy zrobić w nich OSINT-owy rekonesans, warto skorzystać z wcześniej utworzonego anonimowego profilu.

Przemieszczając się coraz bardziej w głąb internetowej góry lodowej, w odmętach wirtualnego oceanu, dotrzemy do części mrocznej i niezbadanej. Często określa się ją jako darknet, nazywając w ten sposób sieć, której zawartość należy często do dark webu, świata treści związanych z przestępstwami, nadużyciami, terroryzmem, przemocą. W poszukiwaniach OSINT-owych, zanim wejdziemy w te obszary Internetu, należy więc wyrobić sobie świadome ich rozróżnienie. Najważniejsze różnice między dark i deep webem zostały zebrane w formie tabeli 5.

Tabela 5. Różnice pomiędzy deep webem i dark webem

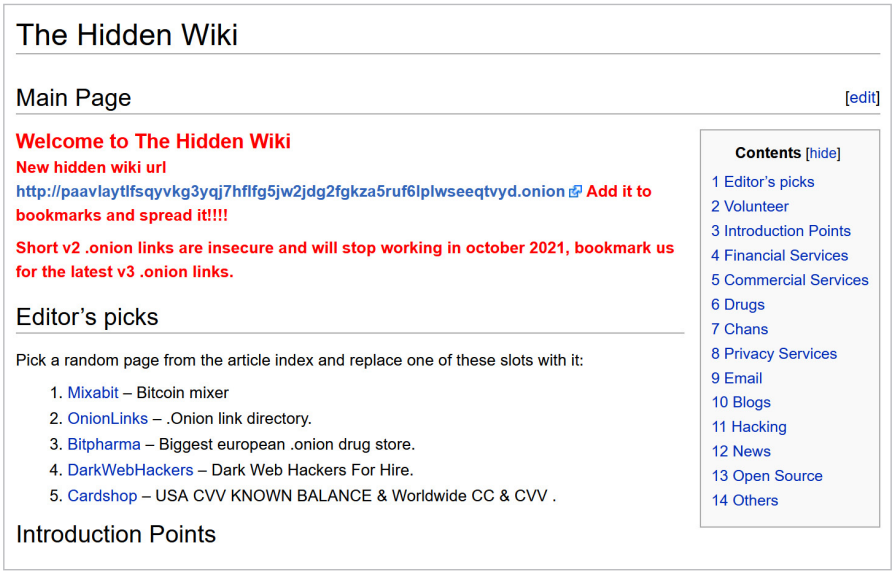
LP.	DEEP WEB	DARK WEB
1.	część Internetu, która nie jest indeksowana przez standardowe i popularne wyszukiwarki	zawartość darknetu, która jest dostępna dzięki możliwości wykorzystania technologii komunikacji (np. sieci Tor) oraz specjalnej przeglądarki i konfiguracji
2.	część Internetu wymagająca dodatkowego dostępu, np. w postaci rejestracji i logowania w zamkniętym portalu	zawartość dostępna jedynie poprzez dostęp do szyfrowanych sieci lub konfiguracji P2P pomiędzy użytkownikami
3.	dostęp nie jest zależny od przeglądarki i technologii	dostęp jest zależny od konkretnej przeglądarki i technologii, np: Tor, I2P, Freenet
4.	wielkość deep webu jest mierzalna ze względu na jej publiczny charakter	wielkość jest niemierzalna i podlega ciągłej zmianie (ekspansji)
5.	bezpieczeństwo jest zależne od powszechnych regulacji i standardów związanych z cyberbezpieczeństwem	bezpieczeństwo nie jest regulowane i przeglądanie zawartości może być niebezpieczne (obecność złośliwych skryptów, XSS-ów, brak certyfikatów bezpieczeństwa itp. zagrożenia)

Sieć darknet wykorzystuje często komunikację P2P, chociaż nie zawsze musi tak być, gdyż najbardziej znany przykład darknetu to sieć Tor, w której zastosowano routing cebulowy[54] (szyfrowanie wielowarstwowe).

W największym uproszczeniu: **sieć Tor** zapewnia użytkownikowi anonimowość oraz uniemożliwia analizę jego ruchu sieciowego dzięki połączeniu użytkownika z trzema węzłami („cebulami"), zanim dotrze do serwera docelowego. Mroczną nazwę darknet zawdzięcza treściom, które się tam znajdują. Ze względu na wysoki stopień anonimowości klienta i serwerów w darknecie można znaleźć wszystko to, co byłoby namierzalne i uznane za nielegalne w publicznym Internecie: sklepy z narkotykami i bronią, zlecenia czynów i usług karalnych, nielegalne fora dyskusyjne, niezgodne z prawem publikacje i nielegalnie rozpowszechniane utwory autorskie. Jest to także niezwykła kopalnia wiedzy z zakresu OSINT. Większość współczesnych teorii spiskowych, faktycznych ataków i zmian sytuacji geopolitycznej oraz – przede wszystkim – nowych podatności, w tym również typu 0-day, ma pierwszą publikację w darknecie.

Surfowanie po sieci Tor należy zacząć od pobrania specjalnej przeglądarki umożliwiającej korzystanie z węzłów tej sieci i udostępnienia użytkownikowi interfejsu SOCKS, np. Tor Browser[55]. Trzeba jednak pamiętać, że poruszanie się po darknecie nie jest w żaden sposób kontrolowane. Nie zobaczymy tam komunikatu informującego o ciasteczkach. Wejście na dowolną stronę może skończyć się uruchomieniem złośliwego skryptu. Dlatego warto skonfigurować dodatek NoScript[56] do przeglądarki, by kontrolować, co i kiedy jest uruchamiane podczas przeglądania stron. Poruszanie się po sieci nie jest proste ze względu na trudność w konstrukcji adresu *.onion. Do połowy roku 2021 istniały krótsze adresy *.onion w wersji v2[57], jednak ze względu na potencjalne niebezpieczeństwo zostały one wyłączone i zastąpione 56-znakowymi adresami v3, mogącymi wyglądać np. tak: `http://paavlaytlfsqyvkg3yqj7hflfg5jw2jdg2fgkza5ruf6lplwseeqtvyd.onion`.

Taki adres warto zapisać, gdyż tak naprawdę jest to punkt startowy w przeglądaniu darknetu. Niestety nie ma gwarancji, że za kilka miesięcy ten adres będzie jeszcze aktywny. Sama strona The Hidden Wiki[58], której adres w sieci Tor podano powyżej, przedstawia mało czytelne linki *.onion uporządkowane w skatalogowanej postaci z podziałem na konkretne obszary zainteresowań, np.: anonimowa skrzynka e-mail, usługi finansowe, wyszukiwarki czy fora darknetowe (rysunek 15). Co ciekawe, można tam znaleźć także oficjalną stronę Tor Centralnej Agencji Wywiadowczej USA, czyli CIA (Central Intelligence Agency).



Rysunek 15. Zrzut ekranu ze strony tytułowej portalu The Hidden Wiki w darknecie

W sieci Tor nie działają wyszukiwarki, ale można używać wyszukiwania operatorami za pośrednictwem wyszukiwarki DuckDuckGo[59] lub Google w przeglądarce Tor Browser po clearnecie. Istnieje zestaw narzędzi do prowadzenia działań białego wywiadu w sieci Tor. Do najbardziej przydatnych zalicza się TorBot oraz OnionScan. TorBot[60] służy do zbierania informacji o podmiocie i działa jak robot indeksujący dane. Jego publicznym

odpowiednikiem jest wspomniany już theHarvester. Narzędzie OnionScan[61] jest weryfikatorem **anonimowości** serwera, który został uruchomiony w sieci Tor. Jednocześnie przekazuje ono ogromną ilość danych cząstkowych połączenia, jeżeli poufność serwera nie jest na właściwym poziomie.

Jako przykład ciekawej historii OSINT związanej z darknetem warto wspomnieć śledztwo AlphaBay. Był to jeden z największych rynków darknetowych w sieci Tor, który był wykorzystywany do działań nielegalnych, takich jak: handel narkotykami, sprzedaż broni i pranie brudnych pieniędzy. FBI użyło połączenia technik OSINT, w tym monitorowania online, analizy mediów społecznościowych i pobierania danych, aby zgromadzić informacje o rynku, sprzedawcach i kupujących. Po miesiącach śledztwa FBI udało się zidentyfikować administratora strony, Alexandra Cazesa, który został aresztowany w Tajlandii, a później znaleziony martwy w swojej celi podczas oczekiwania na ekstradycję do Stanów Zjednoczonych. FBI również przejęło serwery i stronę internetową oraz aresztowało inne osoby związane z serwisem. To śledztwo było ważnym sukcesem organów ścigania, ponieważ zlikwidowało jeden z największych rynków darknetowych i zakłóciło ogromny procent nielegalnej działalności. Jest to przykład, w jaki sposób techniki OSINT mogą zostać wykorzystane do śledzenia działalności przestępczej w darknecie i jak mogą być używane w połączeniu z innymi technikami, takimi jak monitorowanie online i analiza mediów społecznościowych, aby zbierać informacje i identyfikować osoby zaangażowane w nielegalną działalność.

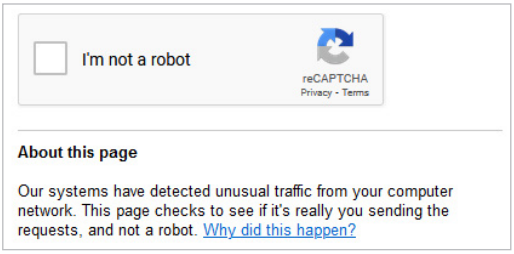
Google Hacking Database

Narzędzi OSINT jest bardzo wiele. Niektóre z nich tak naprawdę nie zostały stworzone do działań wywiadowczych, a jednak mogą być używane do tego celu. Oczywistym przykładem są wyszukiwarki, które formalnie służą jedynie do wyszukiwania, ale czyż OSINT nie jest właśnie tego typu działaniem: wyszukiwaniem danych dostępnych w sposób otwarty? Współczesne wyszukiwarki internetowe dają możliwość stosowania dodatkowych kryteriów podczas wyszukiwania. Taką funkcją jest m.in. Google dorking (lub Google Hacking, bo tak dzisiaj potocznie nazywa się ten sposób używania wyszukiwarki Google, a same zapytania określa się jako Google Dorks), czyli umiejętne formułowanie zapytań do przeglądarki, aby udostępniła dane, które mogą być potrzebne. Problemem jest możliwość dostępu do danych, do których przeglądania użytkownik nie jest uprawniony (w sensie etycznym bądź prawnym).

Google Hacking miał swoje początki w 2002 roku. Zaledwie dwa lata później powstało pierwsze narzędzie służące do proaktywnej obrony przed szpiegowaniem, wykorzystujące Google Dorks: SiteDigger[62], i zaraz potem Google Hack Honeypot[63]. W wypadku alternatywnych wyszukiwarek, jak np.: Yandex, Bing czy DuckDuckGo, użycie operatorów również jest możliwe. Występują różnice w konstruowaniu samych zapytań, ale zazwyczaj są one niewielkie i operują w obrębie tych samych zapisów logicznych George'a Boole'a. Większość narzędzi do precyzyjnego przeszukiwania zasobów Internetu opiera się na automatyzacji procesu wyszukiwań właśnie z użyciem wbudowanych klasyfikatorów i operatorów stron oraz danych. Istnieje nawet strona zbierająca i kategoryzująca „dorki" na podstawie zapytań publikowanych przez użytkowników: GHDB[64] (Google Hacking Database).

Google Dorks to potężne narzędzie. Umiejętnie wykorzystane daje efekty w postaci uzyskania pożądanych danych, należy jednak pamiętać o aspektach prawnych i etycznych takich działań, które oprócz tego, że są monitorowane przez właściciela wyszukiwarki, to jednocześnie mogą stanowić podstawę do podjęcia kroków prawnych wobec osoby, która z naruszeniem prawa i prywatnych pobudek takie informacje wyszukuje.

W tabeli 6 zebrano najpopularniejsze „dorki" z podziałem na charakter wywiadowczy, jakiemu służą. Szary i czarny wywiad są działaniami nielegalnymi i nie należy ich stosować bez odpowiedniego pozwolenia ani wiedzy w celu innym niż edukacyjny. Zastosowanie jakiegokolwiek z „dorków", którego cel jest nielegalny, wiąże się z konsekwencjami i powoduje pojawienie się odpowiedniego komunikatu (rysunek 16).



Rysunek 16. Zrzut ekranu z ostrzeżeniem Google po aktywności użytkownika rozpoznanej jako podejrzana

Tabela 6. Przykładowe zastosowanie operatorów Google Dorks

SKŁADNIA	OPIS	CEL
`site:example.com "Jan Kowalski"`	wyszukiwanie informacji o Janie Kowalskim w domenie example.com	biały
`cache:example.com`	wyszukiwanie zarchwizowanej przez Google wersji strony example.com	biały
`inurl:twitter @jankowalski`	wyszukiwanie konta „jankowalski" w portalu Twitter ze względu na składnię Twitter ID, która musi się znaleźć w adresie URL	biały
`related:sekurak.pl`	wyświetlenie listy stron podobnych do zadanej	biały
`link:hackingparty.pl`	wyświetlanie stron zawierających link do żądanej strony	biały
`site:example.com "*jan*kowalski*@example.com"`	wyszukiwanie wariantów adresu e-mail (znak asterysk *) w domenie example.com na stronie example.com	biały
`site:example.com intext:e-mail`	wyszukiwanie słów „e-mail" na stronie example.com	biały
`site:linkedin.com Firma "prezes zarzadu"`	wyszukiwanie w portalu LinkedIn informacji, kto zajmuje stanowisko prezesa zarządu w firmie	biały
`#bezpieczenstwo`	wyszukiwanie portali i stron, na których został użyty hashtag "bezpieczenstwo"	biały

SKŁADNIA	OPIS	CEL
`~it~security`	użycie znaku tyldy (~) powoduje wyszukanie słów podobnych synonimicznie lub pokrewnych, np. „cybersecurity", „itsec"	biały
`war around(3) Ukraine`	operator around(x) zaprezentuje strony zawierające dwa słowa, np.: „war" i „Ukraine", także te rozdzielone od siebie trzema innymi słowami	biały
`site:example.com albert -einstein`	przeszukiwanie strony example.com w celu odnalezienia słowa „albert" z wyłączeniem wyników słowa „einstein"; znak (minus -) jest wykluczeniem z wyszukiwania	biały
`5 pln to eur`	włączenie dodatkowej funkcjonalności wyszukiwarki Google, będącej przelicznikiem aktualnej wartości kursu wymiany walut, np. pomiędzy PLN a EUR	biały
`filetype:xls site:example.com`	przeszukwanie strony example.com pod kątem odnalezienia plików arkusza kalkulacyjnego Microsoft Excel w formacie XLS	szary
`ext:xls intext:wynagrodzenie`	Przeszukiwanie dowolnej strony pod kątem odnalezienia plików zawierających rozszerzenie *.xls oraz słowo „wynagrodzenie"	szary
`site:example.com intitle:index.of`	przeszukiwanie strony example.com pod kątem słabo zabezpieczonej struktury katalogów serwera WWW, który w tytule wyświetla komunikat „Index of"	szary
`site:example.com inurl:admin`	przeszukiwanie strony example.com pod kątem słowa „admin" w strukturze URL, np. dostęp do strony logowania do panelu administracyjnego	szary
`site:example.com filetype:bak`	przeszukiwanie strony example.com pod kątem istnienia niezabezpieczonych kopii zapasowych na serwerze z rozszerzeniem *.bak	szary
`site:linkedin.com "CEO" (☏ OR ✆) +"Krakow"`	przeszukiwanie portalu LinkedIn pod kątem stanowiska „CEO" oraz lokalizacji „Kraków" z wykorzystaniem emotikonów	szary
`"tresc tweeta" -site:https://twitter.com`	przeszukiwanie cytatów o treści „treść tweeta" dla portali innych niż Twitter	szary
`"Index of" +passwords`	przeszukiwanie dowolnej strony pod kątem istnienia dostępu do plików z poziomu serwera WWW, które zawierają słowo „passwords" w nazwie	czarny
`filetype:bak inurl:passwd|shadow|htusers|htaccess`	przeszukiwanie dowolnej strony pod kątem istnienia kopii plików: passwd, shadow, htusers lub htaccess jako niezabezpieczonej kopii z rozszerzeniem *.bak	czarny
`intitle:"index of" inurl:ftp`	przeszukiwanie dowolnej strony pod kątem istnienia dostępu do serwera FTP przez przeglądarkę	czarny
`inurl:/proc/self/cwd`	przeszukiwanie dowolnego serwera WWW, który został błędnie skonfigurowany lub jego zasady zabezpieczeń zostały wcześniej naruszone	czarny

SKŁADNIA	OPIS	CEL
`filetype:log username putty`	przeszukiwanie dowolnej strony pod kątem przechwycenia pliku *.log, w którym znajduje się ciąg znaków „username" oraz „putty", celem przechwycenia klucza SSH	czarny
`inurl:top.htm inurl:currenttime`	przeszukiwanie dowolnej strony pod kątem przechwycenia publicznego dostępu do kamer D-Link, które zawierają w adresie słowo „currenttime" oraz plik „top.htm"	czarny
`intitle:"Weather Wing WS-2"`	przeszukiwanie dowolnej strony pod kątem publicznego dostępu do urządzenia pogodowego Meteo System WS-2	czarny
`inurl:zoom.us/j and intext:"scheduled for"`	przeszukiwanie dowolnej strony ze spotkania komunikatora Zoom	czarny
`intitle:"Index of" wp-admin`	przeszukiwanie strony, która jest zarządzana przez CMS WordPress oraz której panel administracyjny znajduje się pod domyślną ścieżką, która w adresie URL ma „/wp-admin"	czarny
`inurl:phpmyadmin`	przeszukiwanie dowolnej strony, która w adresie ma ciąg „phpmyadmin", tj. usługa phpMyAdmin jest dostępna w Internecie	czarny
`inurl:Dashboard.jspa intext:"Atlassian Jira Project Management Software"`	przeszukiwanie dowolnej strony, która w adresie ma ciąg związany z publicznym dostępem do usługi zarządzania oprogramowaniem Jira	czarny
`allintitle:"poufny dokument" filetype:docx site:gov`	przeszukiwanie dowolnej strony w domenie *.gov, która zawiera pliki o rozszerzeniu *.docx o tytule zawierającym słowo „poufny dokument"	czarny
`inurl:index.php?id=`	przeszukiwanie dowolnej strony, która została napisana w języku PHP i przekazuje dane w parametrze id za pomocą metody GET	czarny
`filetype:sql intext:password|pass|passwd intext:username intext:"INSERT INTO `users` VALUES"`	przeszukiwanie dowolnej strony zawierającej plik o rozszerzeniu *.sql, który zawiera słowa: „password", „pass", „passwd", „username" lub komendę DML języka SQL z przechwyceniem momentu wprowadzania hasła do bazy danych	czarny
`site:s3.amazonaws.com "confidential" OR "top secret"`	przeszukiwanie kontenerów S3 usługi Amazon Web Services pod kątem istnienia dokumentów zawierających słowo „confidential" lub „top secret"	czarny

Powyższa przykładowa lista może wydać się przerażająca, jednak taka jest rzeczywistość. Każdego dnia wysyłanych jest setki tysięcy podobnych zapytań realizowanych w ramach rekonesansu jako pierwsza faza do ułożenia wektora ataku. Na kanwie powyższych działań powstały takie serwisy, jak wspomniany już Shodan oraz ZoomEye, które – poza faktem bycia kopalnią wiedzy o podatnościach – służą także prewencyjnemu uświadamianiu ludziom, jak często i jak bardzo ich środowiska są niezabezpieczone pod względem poufności, integralności i dostępności do danych. Dlatego niezwykle ważne jest analizowanie bezpieczeństwa własnych danych pod tym kątem, by sprawdzać, w jaki sposób są reprezentowane i indeksowane za pośrednictwem wyszukiwarek. Istnieje

możliwość zabezpieczenia odpowiednich stron i odcięcia robotom indeksującym dostępu do nich. Każdy serwer WWW może być zaopatrzony w plik `robots.txt` (tabela 7), który powinien być honorowany przez pająki indeksujące.

Tabela 7. Przykładowe zapisy w pliku `robots.txt`

KONFIGURACJA ROBOTS.TXT	OPIS
`User-agent: *` `Disallow: /admin/`	dla dowolnego bota, nie zezwalaj na indeksowanie strony `/admin/`
`User-agent: Googlebot` `Disallow: /priv/plik.bak`	dla bota Google, nie zezwalaj na indeksowanie pliku `plik.bak` ze strony `/priv/`*
`User-agent: *` `Disallow: /*.xls$/`	dla dowolnego bota, nie zezwalaj na indeksowanie plików o rozszerzeniu *.xls

Oczywiście zapisy z tabeli 7 nie powodują zabezpieczenia dostępu do zasobów, a jedynie nieindeksowanie ich w wynikach wyszukiwania. Jest to jednak pierwszy krok do prewencji i uniknięcia zostania „bohaterem" serwisu Shodan lub ZoomEye. Legalne używanie Google Dorks potrafi w sposób zdecydowany przyspieszyć działania białego wywiadu, lecz w większości przypadków nadal szybsze będą narzędzia, które mogą działać multi- i interoperacyjnie.

NARZĘDZIA DO POZYSKIWANIA INFORMACJI O PODMIOCIE

Z technik OSINT-owych w ostatnich latach wyłoniła się gałąź zwana SOCMINT, czyli Social Media Intelligence, przegląd źródeł społeczności internetowych. To właśnie wszelkiego rodzaju media społecznościowe stały się dominującym źródłem pozyskiwania informacji o podmiocie.

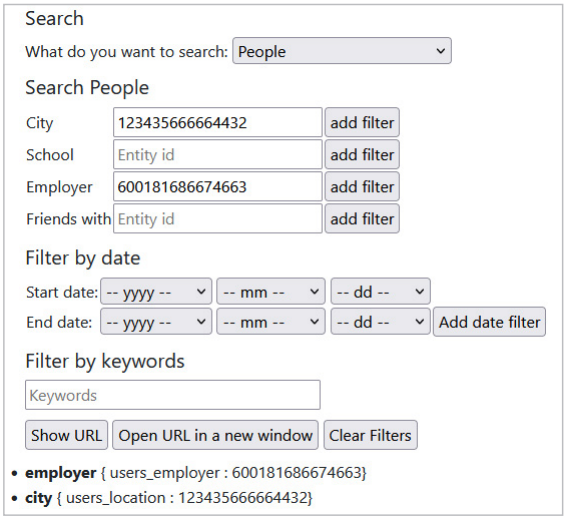
Natura ludzka jest tak skonstruowana, że 80% osób korzystających z mediów społecznościowych pragnie się pochwalić, gdzie jest, co robi, jakie są ich osiągnięcia. Pozostałe 20% nie jest tym zainteresowane, jednak nadal muszą mieć konto na portalach społecznościowych, inaczej, zgodnie z regulaminami, nie mogliby przeglądać tego, co inni tak ochoczo udostępniają. Wszystkich natomiast łączy jedno: ciekawość. Chcemy wiedzieć jak najwięcej. O koledze, sąsiedzie, celebrycie, a może o starej znajomej sprzed lat. Platformy społecznościowe stały się więc swoistymi skarbnicami dla służb – ale także dla cywili – do przetwarzania danych Big Data[65].

SOCMINT jest niezwykle cennym źródłem informacji i istnieje bardzo wiele narzędzi wspierających proces zbierania danych w tym środowisku. W naszej części świata podstawowymi portalami społecznościowymi są: Facebook, Twitter, LinkedIn oraz Instagram – i tu skupimy się na nich jako na obszarach do penetracji w białym wywiadzie. W przypadku młodszych użytkowników prym wiodą: TikTok, Snapchat, Discord oraz Twitch.

* Warto jednak nadmienić, że nie jest to dzisiaj żadna forma ochrony. Atakujący lub narzędzia automatycznie przeszukujące zawartość portalu natychmiast odczytają zawartość `robots.txt` w celu odnalezienia takich ścieżek.

Facebook

Trzy miliardy kont w 2023 roku[66] – największy obecnie portal społecznościowy to Facebook, a największym narzędziem wspierającym działania OSINT-owe jest wbudowana w ten portal zaawansowana wyszukiwarka – Graph[67] (wcześniej nazywana Facebook Directory). Narzędzie to umożliwia wyszukiwanie postów, osób, relacji, miejsc, reklam, przedmiotów na sprzedaż, treści multimedialnych oraz wyszukiwań krzyżowych, np. czy użytkownik danego profilu odwiedził miejscowość X. Niestety ze względu na ciągłe kontrowersje związane z łamaniem polityki prywatności wyszukiwarka ta ulega stałym modyfikacjom. Dlatego korzystanie z dostępnych narzędzi automatyzujących proces wyszukiwania, jak np. Facebook Graph Searcher[68] od Intelligence X lub – alternatywnie – Facebook Search od użytkownika o nicku s0wdust[69], które powstało po tym, jak Facebook ograniczył możliwości narzędzia Graph, bywa kłopotliwe. Jeżeli autorzy nakładki nadążają za wprowadzanymi w portalu zmianami, to uzyskane wyniki będą satysfakcjonujące, jednak bywa z tym różnie. Dodatkowo trzeba znać identyfikator danego profilu, by wprowadzić go w pożądane pole. W czasie, w którym powstawał ten rozdział, aktualne identyfikatory można znaleźć samodzielnie, analizując kod źródłowy stron profilu i wyszukując parametry `pageID`, `userID` dla użytkowników indywidualnych, `groupID` dla grup, np.:
`"props":{"pageID":"600181686674663","tabName":"tab_home","ref":""}`
… lub korzystając z darmowego narzędzia online Lookup-ID[70] (rysunek 17).



Rysunek 17. Zrzut ekranu z narzędzia użytkownika s0wdust – Facebook Search

Kłopoty spowodowane ciągłymi zmianami w Graphie, które utrudniają wyszukiwanie, można ominąć, rozpoznając technologię stojącą za sposobem wyszukiwania tego narzędzia. Mechanizmy wyszukiwawcze wykorzystywane w narzędziu Facebook Graph są nakładkami modyfikującymi URL zapytania, których składnia zawsze wygląda tak samo, bazuje na technologii JSON oraz kodowaniu Base64:
`facebook.com/search/<KATEGORIA>/?q=<CEL>&epa=FILTERS&filters=<FILTRY>`

W tabeli 8 przedstawiono kategorie, które można wykorzystać w linku URL, poszukując konkretnego celu.

Tabela 8. Rodzaje kategorii w narzędziu Facebook Graph

KATEGORIA	RODZAJ WYSZUKIWANIA
/top/	przeszukaj TOP content
/posts/	przeszukaj posty publiczne
/people/	poszukaj ludzi
/photos/	poszukaj zdjęć
/videos/	poszukaj wideo
/pages/	przeszukaj strony (*pages*)
/places/	poszukaj miejsc
/marketplace/	przeszukaj ogłoszenia lokalne
/groups/	przeszukaj grupy
/events/	przeszukaj wydarzenia

Jeżeli chcemy wyszukać osoby np. o danych Jan Kowalski, należałoby skonstruować zapytanie w następujący sposób (zamieniając spację na `%20` w adresie URL):
`facebook.com/search/people/?q=Jan%20Kowalski&epa=FILTERS&filters=`

Samo zapytanie złożone po słowie `filters=` wymaga znajomości składni struktury JSON. Na przykład jeżeli chcemy sprawdzić lokalizacje, w których przebywał dany użytkownik, należałoby utworzyć takie zapytanie do struktury, gdzie `109212626870260` to `userID`:

```
{"city":"{\"name\":\"users_location\",\"args\":\"109212625870260\"}"}
```

Zapytania zgodnie ze standardem struktury JSON można ze sobą łączyć, np. dodając pole zakładu pracy zatrudniającego danego użytkownika:

```
{"city":"{\"name\":\"users_location\",\"args\":\"109212625870260\"}","emp
loyer":"{\"name\":\"users_employer\",\"args\":\"104458163837\"}"}
```

Pozostałe przykładowe zapytania zostały przedstawione w tabeli 9.

Następnie, wykorzystując np. narzędzie CyberChef[71], wybieramy To Base64, aby zakodować powyższą strukturę w Base64 i połączyć z fragmentem adresu po słowie `&filters=`, otrzymując w ten sposób poprawne zapytanie:

```
facebook.com/search/people/?q=Jan%20Kowalski&epa=FILTERS&filters=eyJja
XR5Ijoie1wibmFtZVwiOlwidXNlcnNfbG9jYXRpb25cIixcImFyZ3NcIjpcIjEwOTIxMjY
yNTg3MDI2MFwifSIsImVtcGxveWVyIjoie1wibmFtZVwiOlwidXNlcnNfZW1wbG95ZXJcI
ixcImFyZ3NcIjpcIjEwNDQ1ODE2MzgzN1wifSJ9
```

Jeżeli takie zapytanie gdziekolwiek zobaczymy, istnieje też możliwość prostego zdekodowania go po wybraniu w CyberChef dekodowania From Base64.

Tabela 9. Najbardziej popularne zapytania strukturalne w narzędziu Facebook Graph

KATEGORIA	ZAPYTANIE	CZEGO DOTYCZY
/top/	{"rp_chrono_sort":"{\"name\":\"chronosort\", \"args\":\"\"}"}	najnowsze, najbardziej popularne treści
/top/	{"rp_author":"{\"name\":\"merged_public_posts\", \"args\":\"\"}"}	najbardziej popularne treści od autora
/top/	{"rp_author":"{\"name\":\"author_friends_feed\", \"args\":\"\"}"}	najbardziej popularne treści publikowane przez znajomych
/top/	{"rp_author":"{\"name\":\"author\",\"args\":\"ID\"}"}	najbardziej popularne treści od autora
/top/	{"rp_group":"{\"name\":\"group_posts\",\"args\":\"ID\"}"}	najbardziej popularne treści z grupy ID
/top/	{"rp_location":"{\"name\":\"location\",\"args\":\"ID\"}"}	najbardziej popularne treści z lokalizacji ID
/posts/	{"rp_author":"{\"name\":\"author_friends_feed\", \"args\":\"\"}"}	posty znajomych
/posts/	{"rp_author":"{\"name\":\"author\",\"args\":\"ID\"}"}	posty autora
/posts/	{"rp_group":"{\"name\":\"group_posts\",\"args\":\"ID\"}"}	posty z grupy ID
/posts/	{"rp_location":"{\"name\":\"location\",\"args\":\"ID\"}"}	posty z lokalizacji ID
/people/	{"city":"{\"name\":\"users_location\",\"args\":\"ID\"}"}	wyszukiwanie ludzi w lokalizacji ID
/people/	{"school":"{\"name\":\"users_school\",\"args\":\"ID\"}"}	wyszukiwanie ludzi na podstawie ID szkoły
/people/	{"employer":"{\"name\":\"users_employer\", \"args\":\"ID\"}"}	wyszukiwanie ludzi na podstawie ID pracodawcy
/people/	{"friends":"{\"name\":\"users_friends_of_people\", \"args\":\"ID\"}"}	wyszukiwanie znajomych na podstawie ID użytkownika
/photos/	{"rp_author":"{\"name\":\"author\",\"args\":\"ID\"}"}	wyszukiwanie zdjęć na podstawie ID użytkownika
/photos/	{"rp_group":"{\"name\":\"group_posts\",\"args\":\"ID\"}"}	wyszukiwanie zdjęć na podstawie ID grupy
/photos/	{"rp_location":"{\"name\":\"location\",\"args\":\"ID\"}"}	wyszukiwanie zdjęć z ID lokalizacji
/videos/	{"videos_source":"{\"name\":\"videos_live\", \"args\":\"\"}"}	wyszukiwanie aktualnie streamowanych materiałów wideo
/videos/	{"rp_location":"{\"name\":\"location\",\"args\":\"ID\"}"}	wyszukiwanie materiałów wideo na podstawie ID lokalizacji
/pages/	{"category":"{\"name\":\"pages_category\", \"args\":\"1006\"}"}	wyszukiwanie miejsc i lokalnych atrakcji (ID=1006)
/pages/	{"category":"{\"name\":\"pages_category\", \"args\":\"1013\"}"}	wyszukiwanie stron firmowych, organizacji i instytucji (ID=1013)

KATEGORIA	ZAPYTANIE	CZEGO DOTYCZY
/pages/	{"category":"{\"name\":\"pages_category\", \"args\":\"1009\"}"}	wyszukiwanie stron produktów (ID=1009)
/pages/	{"category":"{\"name\":\"pages_category\", \"args\":\"1007,ID\"}"}	wyszukiwanie stron artystów i osób publicznych (ID=1007)
/pages/	{"category":"{\"name\":\"pages_category\", \"args\":\"1019\"}"}	wyszukiwanie stron na temat rozrywki (ID=1019)
/pages/	{"category":"{\"name\":\"pages_category\", \"args\":\"2612\"}"}	wyszukiwanie stron na temat konkretnego wydarzenia (ID=2612) lub społeczności

Twitter

Twitter jest portalem społecznościowym skupiającym się wokół przekazywania krótkich wpisów informacyjnych. Podstawową ideą przyświecającą twórcom tego portalu było jak najszybsze publikowanie skróconych informacji (depesza, telegram) i ich rozpowszechnianie. Twitter to miejsce, w którym konta ma większość polityków, osób publicznych, celebrytów oraz zwykłych obserwujących ich ludzi.

Jednym z przykładów spektakularnego wykorzystania narzędzi OSINT było odnalezienie osoby zamieszanej w porwanie i morderstwo w 2014 roku w Meksyku. 43 studentów z Ayotzinapa zostało porwanych i ślad po nich zaginął[72]. W wyniku śledztwa ujawniono, że zostali oni schwytani przez miejscowych gangsterów i przekazani skorumpowanej policji, a następnie zabici i spaleni w krematorium. Organizacje pozarządowe i dziennikarze użyli narzędzi OSINT do zbierania informacji o zdarzeniu. W szczególności skupili się na analizie relacji między różnymi podmiotami i na monitorowaniu działań władz. Przy użyciu danych, głównie z Twittera, oraz innych źródeł publicznie dostępnych, udało się zebrać wiele informacji, które pozwoliły na ujawnienie kłamstw i manipulacji ze strony rządu. W wyniku tego śledztwa ujawniono, że rząd Meksyku próbował zataić fakty związane z porwaniem i morderstwem studentów, a także manipulować dowodami. Dzięki narzędziom OSINT udało się ujawnić te nieprawidłowości i zwrócić uwagę międzynarodową na sytuację w Meksyku. To przyczyniło się do zwiększenia nacisku na rząd i na podjęcie działań zmierzających do poprawy sytuacji w kraju.

Ten przykład pokazuje, jak narzędzia OSINT mogą pomóc w walce z korupcją i naruszeniami praw człowieka. Dzięki nim można gromadzić dowody, które pozwalają na ujawnienie nieprawidłowości i wykorzystać je do wzmocnienia działalności organizacji pozarządowych, mediów i innych podmiotów, które dążą do poprawy sytuacji społeczno-politycznej w kraju.

Twitter stanowi obecnie ogromne źródło wiedzy. Na pewno jest to miejsce, w którym większość informacji ma swój początek, a także prędkość jej przekazywania jest zdecydowanie większa niż w pozostałych portalach społecznościowych.

Wbudowana w portal wyszukiwarka ma możliwość używania operatorów zaawansowanych przedstawionych w tabelach 10 i 11[73]. Podstawowym linkiem wyszukiwania jest zapis: `https://twitter.com/search?q=<TREŚĆ>&src=typed_query&f=<FILTR>`

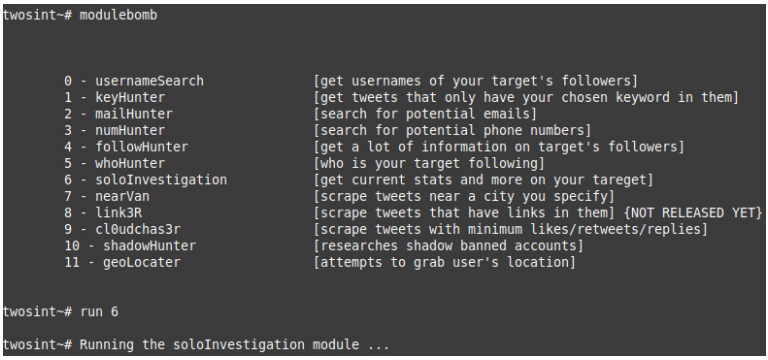
Tabela 10. Operatory wyszukiwania treści w portalu Twitter

OPERATOR Z TREŚCIĄ	OPIS
`@uzytkownik`	zwraca tweety wspominające określonego użytkownika, np. @tturba
`from:SEINT_pl`	zwraca tweety utworzone przez określonego użytkownika, np. SEINT_pl
`to:sajdoor`	zwraca tweety skierowane do określonego użytkownika oraz „odpowiedzi do:"
`since:2023-02-24`	zwraca tweety utworzone po konkretnej dacie w formacie YYYY-MM-DD (również uwzględniając tę datę)
`until:2023-06-30`	zwraca tweety utworzone przed konkretną datą w formacie YYYY-MM-DD (również uwzględniając tę datę)
`-http`	operator negacji "-", np.: nie pokazuj tweetów z linkami do zewnętrznych portali
`lang:pl`	zwraca tweety wyłącznie w wybranym języku, zapisanym w postaci dwuznaków
`near:"Krakow, PL" within:10km`	zwraca tweety utworzone w pobliżu wybranej lokalizacji oraz w promieniu, np. 10 km od niej

Tabela 11. Filtry wyszukiwania w portalu Twitter

FILTR	OPIS
`filter:live`	najnowsze tweety
`filter:images`	zdjęcia
`filter:links`	linki
`filter:media`	zdjęcia lub wideo
`filter:news`	linki do materiałów źródłowych
`filter:verified`	tweety zweryfikowanych użytkowników
`filter:videos`	materiały wideo

Do jednych z najpopularniejszych narzędzi OSINT-owych dla portalu Twitter zaliczano twosint[74] będący nakładką na bibliotekę Twint[75] (zaawansowany projekt scrapingowy). Wraz z przejęciem firmy przez Elona Muska w 2022 roku nastąpiło jednak wiele zmian programistycznych w portalu, które ostatecznie doprowadziły do zamknięcia kilku narzędzi białowywiadowczych dla Twittera, także Twinta. Twosint (rysunek 18) jako nakładka w obecnej formie nie współpracuje z Twitterem. Autor biblioteki Twint stworzył jednak nowy scraper o nazwie Twint Zero[76], który współpracuje z aktualną wersją portalu.



Rysunek 18. Możliwości archiwalnego narzędzia twosint

Biblioteka Twint Zero pozwala na gromadzenie danych bez wykorzystywania API i logowania. Jej przewagę nad innymi stanowi brak ograniczeń w zakresie pobierania danych. API pozwala maksymalnie na scraping 3200 tweetów. W przypadku korzystania z samej biblioteki Twint Zero (w tej chwili bez używania twosint) niezbędna jest podstawowa znajomość języka Go. W poprzedniej wersji biblioteki Twint niezbędna była wiedza z zakresu języka Python. Przykładowy archiwalny kod dla biblioteki Twint znajduje się w listingu 4 (wyszukiwanie słowa "great" we wpisach konta "realDonaldTrump"):

Listing 4. Przykładowa konfiguracja biblioteki Twint Zero

```
import twint

# Configure
c = twint.Config()
c.Username = "realDonaldTrump"
c.Search = "great"

# Run
twint.run.Search(c)
```

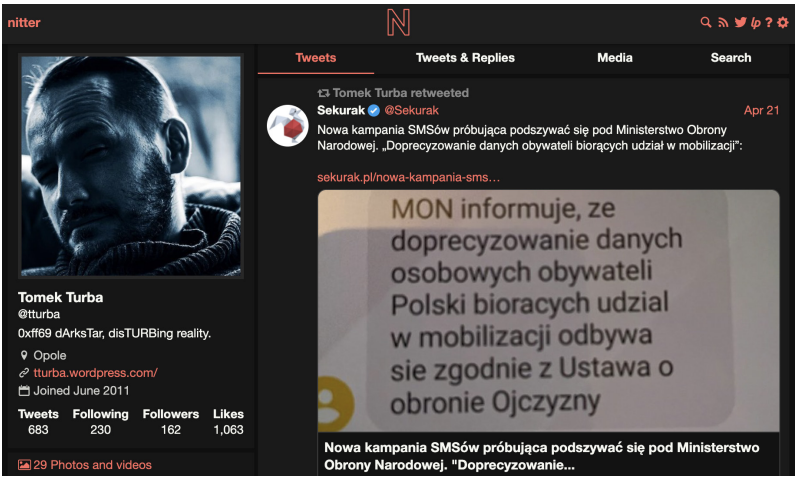
Poniżej (listing 5) przedstawiony został sposób instalacji obu narzędzi wraz z koniecznymi zależnościami.

Listing 5. Instalacja biblioteki Twint Zero oraz jej uruchomienie

```
# apt-get install golang
# git clone https://github.com/twintproject/twint-zero
# cd twint-zero
# go mod init twint-zero
# go mod tidy
```

Uruchomienie wyszukiwania i scrapowania odbywa się poprzez wpisanie komendy:
`# go run main.go -Query <QUERY> -Instance <INSTANCE> -Format <FORMAT>`

W polu QUERY należy wpisać operator(-y) treści z tabel 10 i 11. Dodatkowo niezbędne jest podanie instancji pośredniczącej, by nie zostać zbanowanym na Twitterze za pobieranie zbyt dużych ilości informacji. Przykładem takiej instancji jest nitter.net* (rysunek 19), która służy jako aplikacja *front-end* i *open source* dla Twittera. W przypadku parametru FORMAT należy wskazać format wyświetlanych wyników: plik .csv lub .json.



Rysunek 19. Zrzut ekranu z narzędzia nitter.net

Przykładowe zapytanie zostało przedstawione poniżej wraz z wynikami (rysunek 20):

```
# go run main.go -Query sekurak -Instance nitter.simpleprivacy.fr -Format json
```

Rysunek 20. Wynik pracy narzędzia Twint Zero

* Pełna lista aktywnych instancji dostępna jest pod adresem: zedeus, *nitter*, GitHub, https://github.com/zedeus/nitter/wiki/Instances

Mimo że Twitter obecnie zmienia się praktycznie każdego dnia, wszyscy OSINT-owcy mają nadzieję, że autor narzędzia twosint stworzy aktualizację wykorzystującą bibliotekę Twint Zero, gdyż aktualnie parsowanie dużej liczby wyników w formie tekstu jest niewygodne, co przedstawiono na rysunku 20.

Użytkownik Twittera, poza funkcjonalnością publikacji postów i interakcji z nimi, ma też możliwość tworzenia list obserwacyjnych agregujących dane ze względu na temat lub ciekawe dla niego treści. Dostęp do takich list uzyskuje się dzięki fragmentowi `/lists/` podanemu w adresie URL. W związku z tym można wykorzystać „dork", który w wyszukiwarce Google wyświetli listę zainteresowań:

```
site:twitter.com inurl:lists <słowo kluczowe>
```

Ten sposób przeszukiwania może posłużyć do późniejszej obserwacji jednej z takich list – zgodnie z zadanymi słowami kluczowymi. Wejście w interakcję z listą poprzez jej obserwowanie z zalogowanego konta zapewni autorowi listy informację o powstaniu subskrypcji. Z uwagi na prowadzenie (w większości przypadków) działań operacyjnych w formie pasywnej rekomendowanym sposobem pracy jest użycie narzędzia do kopiowania listy do innego, zalogowanego, konta. Twitter List Copy[77], narzędzie przeznaczone do takich zadań, wymaga logowania z dowolnego konta, do którego lista zostanie następnie skopiowana.

W tabeli 12 przedstawiono pozostałe warte uwagi narzędzia, które są związane z głęboką analizą danych dostępnych w portalu Twitter.

Tabela 12. Porównanie funkcjonalności narzędzi online do analizy zasobów Twittera

NARZĘDZIE	OPIS
twXplorer https://twxplorer.knightlab.com	► wymaga logowania ► możliwość wyszukiwania w wielu językach ► możliwość zapisania migawki wyników
TruthNest https://app.truthnest.com	► wymaga logowania ► przyjazny interfejs ► możliwość eksportu wyników do dokumentu PDF
Spoonbill https://spoonbill.io	► wymaga logowania ► możliwość śledzenia zmian na wskazanym koncie

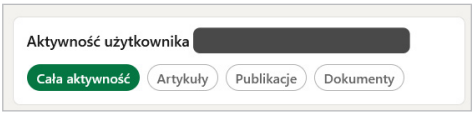
LinkedIn

LinkedIn jest portalem społecznościowym skupiającym się wokół środowiska pracy. Profile użytkowników są wypełniane miejscami zatrudnienia wraz z doświadczeniem zawodowym, certyfikatami, umiejętnościami oraz etapami dotychczasowej nauki. Same firmy również mają profile społecznościowe. Portal jest też niezwykle przydatny w poszukiwaniach SOCMINT ze względu na jego specyfikę. Z reguły osoby w takim portalu podają pełne imię, nazwisko, miejsce pracy oraz swoją aktualną fotografię. Aby przeglądać takie dane, należy mieć zarejestrowane konto lub właściciele analizowanych kont muszą zgadzać się na dostęp otwarty. Jednak nawet kiedy te kryteria są spełnione, kliknięcie w profil użytkownika zdradza mu nasze intencje w postaci powiadomienia „Jan Kowalski oglądał Twój profil". Zależy to jednak od skonfigurowanych ustawień profili po

obu stronach. Można wprowadzić takie ustawienia, że będzie widoczny tylko komunikat: „<stanowisko> przeglądał Twój profil", gdzie <stanowisko> to np. rekruter IT. Jest to parametr wpływający także na nasz sposób otrzymywania informacji. Im większe ograniczenia dostępu ma nasz profil, tym mniej informacji do nas dociera. Konta Premium nie mają tego ograniczenia.

W momencie pisania tego rozdziału LinkedIn oferował także statystyczny zbiór aktywności użytkownika na portalu. Formalnie jest to funkcjonalność dostępna tylko dla znajomych właściciela danego profilu lub za pośrednictwem aplikacji mobilnej. Jednak... URL tej funkcjonalności jest stały i ma następującą postać:
`https://www.linkedin.com/in/<nazwa profilu>/detail/recent-activity/`

Oznacza to, że każdy, kto zna nazwę profilu użytkownika, dopisując w adresie URL powyższy przyrostek, może sprawdzić aktywność użytkownika (jeżeli ten nie blokuje dostępu do profilu). Przedstawiono to na rysunku 21.



Rysunek 21. Dostęp do statystyk aktywności użytkownika po spreparowaniu linku

Instagram

Jednym z kluczowych narzędzi do pozyskiwania danych z platformy Instagram jest Osintgram dostępny na platformie GitHub[78]. Narzędzie to pozwala na zebranie informacji takich jak: adresy e-mail osób śledzących profil, krzyżowy przegląd kont wchodzących w interakcję z kontem, częściowo – numerów telefonów (w przypadku rejestracji w portalu za pośrednictwem numeru telefonu) oraz – co najważniejsze – metadanych publikowanych zdjęć. Pełna lista możliwości znajduje się w tabeli 13, a na rysunku 22 przedstawiono przykład pozyskania danych.

Instalacja oprogramowania Osintgram (listing 6) sprowadza się do sklonowania repozytorium i uruchomienia interpretera języka Python 3 wraz z poleceniem.

Listing 6. Instalacja i uruchomienie narzędzia Osintgram

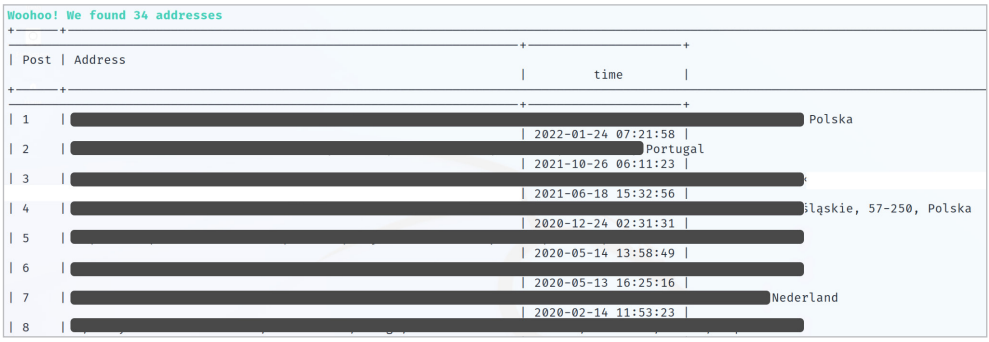
```
# git clone https://github.com/Datalux/Osintgram.git
# cd Osintgram/
# sudo apt-get install -y python3-pip python3-venv
# sudo python3 -m venv venv
# sudo source venv/bin/activate
# sudo pip install -r requirements.txt
# vi config/credentials.ini
```

W pliku `credentials.ini` należy umieścić login i hasło profilu Instagram, który będzie wykorzystywany do uruchomienia narzędzia Osintgram. Następnie narzędzie można uruchomić z interpretera języka Python:
`# sudo python3 main.py <target username> {--command <command>}`

Tabela 13. Lista komend narzędzia Osintgram

PARAMETR WYWOŁANIA	OPIS KOMENDY
addrs	pokaż wszystkie zarejestrowane adresy przypisane do zdjęć danej osoby
captions	pokaż tytuły zdjęć danej osoby
comments	pokaż wszystkie komentarze postów danej osoby
followers	pokaż profile użytkowników obserwujących dany profil
followings	pokaż użytkowników obserwowanych przez daną osobę
fwersemail	pokaż adresy e-mail użytkowników obserwujących dany profil
fwingsemail	pokaż adresy e-mail użytkowników obserwowanych przez dany profil
fwersnumber	pokaż numery telefonów użytkowników obserwujących dany profil
fwingsnumber	pokaż numery telefonów użytkowników obserwowanych przez daną osobę
hashtags	pokaż hashtagi używane przez dany profil
info	pokaż informacje zbiorcze o danym profilu
likes	pokaż łączną liczbę polubień postów na danym profilu
mediatype	pokaż określony typ postów (zdjęcia lub wideo)
photodes	pokaż opisy zdjęć na danym profilu
photos	pobierz zdjęcia z danego profilu do folderu lokalnego
propic	pobierz zdjęcie profilowe danej osoby
stories	pobierz relacje foto/wideo z danego profilu
tagged	pokaż listę użytkowników oznaczonych przez daną osobę
wcommented	pokaż listę użytkowników, którzy skomentowali zdjęcia danej osoby
wtagged	pokaż listę użytkowników, którzy oznaczyli cel

Przykładowe uruchomienie narzędzia w celu poszukiwania adresów użytkownika testuser: `# sudo python3 main.py testuser --command addrs`

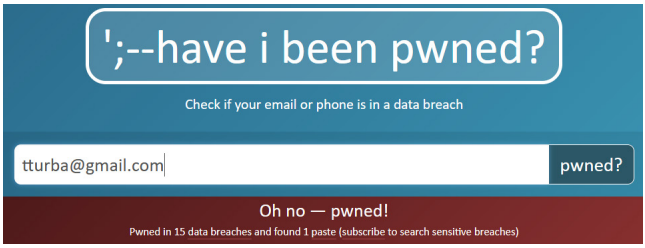


Rysunek 22. Metadane adresowe wyszukane przy użyciu narzędzia Osintgram

Inne źródła

Sama możliwość analizy konta w portalach społecznościowych w przypadku innych narzędzi niż te już omówione pozwala na uzyskanie różnych interesujących danych szczątkowych, np. o korzystaniu przez danego użytkownika z innych portali. TikTok posiada np. wbudowaną funkcję linkowania profilu użytkownika z jego kontem na Instagramie, na którym można zazwyczaj znaleźć znacznie więcej treści, ciekawych z OSINT-owego punktu widzenia.

Jednym z bardziej znanych i najłatwiejszych narzędzi do pozyskania informacji o powiązanych kontach jest serwis Have I Been Pwned (rysunek 23) lub narzędzie zbudowane na podobnych założeniach – Profil3r.



Rysunek 23. Zrzut ekranu ze strony haveibeenpwned.com

Instalacja i konfiguracja Profil3ra (listing 7) odbywa się zgodnie z instrukcją zamieszczoną w portalu GitHub[79].

Listing 7. Instalacja i uruchomienie narzędzia Profil3r

```
# sudo pip3 install PyInquirer jinja2 bs4
# sudo git clone https://github.com/MrNonoss/Profil3r-docker.git
# cd Profil3r/
# sudo python3 setup.py install
# sudo python3 profil3r.py -p john doe
```

Na rysunku 24 przedstawiono fragment drzewa danych przechwyconych za pomocą narzędzia Profil3r.

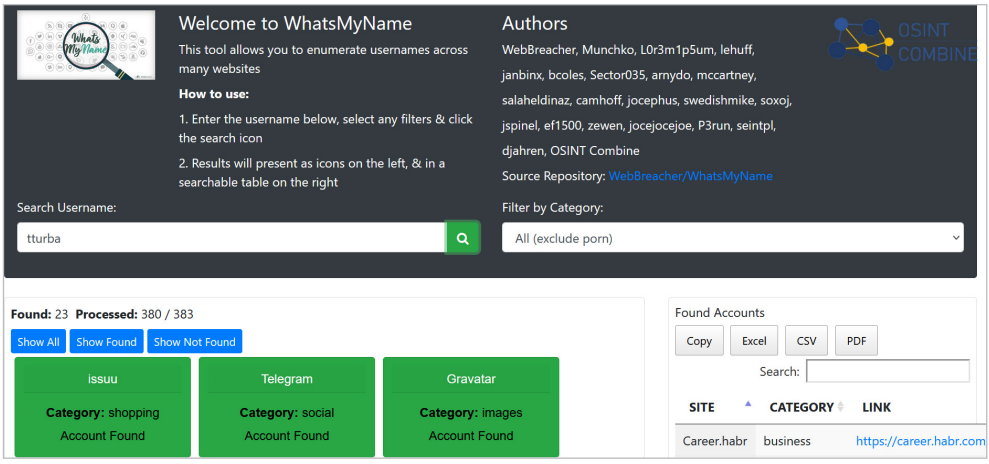
O ile serwis Have I Been Pwned porównuje istnienie kont z wyciekami danych, to serwis WhatsMyName[80] enumeruje nazwę użytkownika na wielu portalach, sprawdzając, czy konto istnieje, bez faktu powstania wycieku. Obecnie jest to najbardziej popularne narzędzie do OSINT-owego sprawdzania istnienia kont (rysunek 25).

KOMBAJNY DO POZYSKIWANIA INFORMACJI O ORGANIZACJI

Poza ręcznym pozyskiwaniem informacji narzędzia automatyzujące i przyspieszające analizę stały się podstawowym uzupełnieniem pracy analityka, śledczego i agenta. Prezentowane poniżej narzędzia są jedynie przykładami. Istnieje ich znacznie więcej, a ich wybór zależy od osobistych preferencji użytkowników. Warto jednak poznać ich jak najwięcej, by umieć wykorzystać indywidualny potencjał każdego z nich. Z powodzeniem można stosować je także wtedy, gdy celem poszukiwań jest konkretny człowiek.

```
└── HACKERNEWS ✔
    └── https://news.ycombinator.com/user?id=johndoe
    |   ├── Creation Date : March 21, 2009
    |   ├── Karma : 1

└── PATREON ✔
    └── https://www.patreon.com/johndoe

└── XVIDEOS ✔
    └── https://www.xvideos.com/profiles/johndoe

└── TWITTER ✔
    └── https://twitter.com/johndoe
    |   ├── Full Name : —
    |   ├── Username : @johndoe
    |   ├── Bio : dude, wait … what ?
    |   ├── Tweets : 21254
    |   ├── Following : 140
    |   ├── Followers : 534
    |   ├── Likes : 57185

└── SOUNDCLOUD ✔
    └── https://soundcloud.com/johndoe

└── SMULE ✔
    └── https://smule.com/johndoe
```

Rysunek 24. Fragment drzewa danych przechwyconych za pomocą narzędzia Profil3r

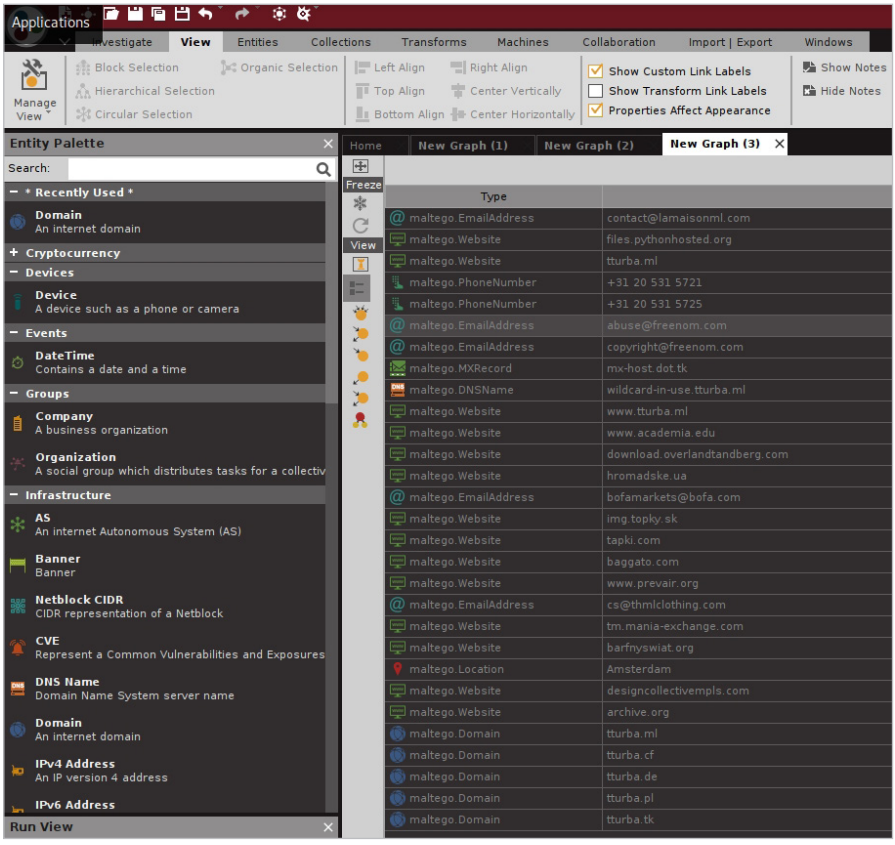


Rysunek 25. Zrzut ekranu z narzędzia WhatsMyName

Maltego

Maltego to jedno z największych narzędzi Open Source Intelligence służące do gromadzenia danych oraz tworzenia graficznych połączeń i relacji między nimi. Jest ono powszechnie wykorzystywane zarówno przez zwykłych ludzi (osoby prywatne, reporterzy śledczy), jak i przez profesjonalistów zajmujących się zawodowo zagadnieniami wywiadowczymi. Model licencjonowania uwzględnia wersję CE (Community Edition) – za darmo, po zarejestrowaniu konta. Najnowsza wersja zawiera tryb komunikacji Stealth, który w pewnym aspekcie ogranicza komunikację z adresu IP inicjatora poszukiwań, gdyż nie pobiera danych z Internetu (chociaż należy uważać, ponieważ nie cały ruch może zostać zablokowany).

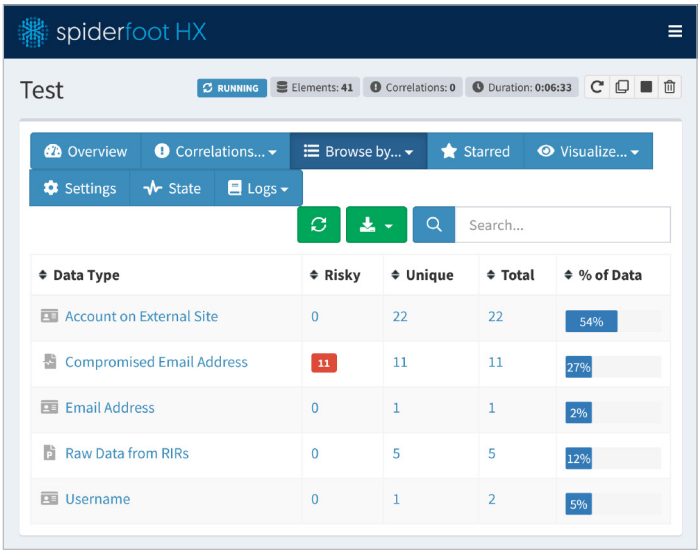
Obsługa programu polega na uzupełnieniu danych wstępnych – wybraniu encji (np.: domena, e-mail, dane osobowe, adres IP, adres portfela kryptowaluty itd.) – i transformat poszukiwawczych. Narzędzie jest cały czas rozwijane i stało się swoistym kombajnem, który po kilku kliknięciach zwraca ogromną pulę informacji do analizy i łączenia. Na rysunku 26 przedstawiono zrzut ekranu ze znalezionych wyników transformat dla encji domeny. Widok tabeli można przełączać symultanicznie z widokiem grafu (w postaci drzewa, gwiazdy lub siatki).



Rysunek 26. Zrzut ekranu zebranych danych z transformat w widoku listy narzędzia Maltego

SpiderFoot

Dobrą alternatywę i uzupełnienie Maltego stanowi SpiderFoot, który jest jeszcze większym kombajnem. Łączy on cechy narzędzi do poszukiwania podmiotów, danych o organizacji z funkcjonalnościami narzędzi do badania infrastruktury. Wczesna wersja z 2005 roku była dostępna jedynie jako oprogramowanie do zainstalowania pod Linuxem. Wraz z postępem technologicznym i pojawieniem się mody na technologie online powstała wersja SpiderFoot HX (rysunek 27) dostępna przez przeglądarkę. Podobnie jak w przypadku Maltego, SpiderFoot HX jest również dostępny w wersji darmowej po rejestracji konta.



Rysunek 27. Zrzut ekranu ze skanu wykonanego narzędziem online SpiderFoot HX

theHarvester

theHarvester[81] jest bardzo prosty w użyciu, ale potężny i skuteczny w działaniu. To narzędzie przeznaczone do stosowania na wczesnych etapach testu penetracyjnego. Jest też powszechnie używanym narzędziem defensywnym obrazującym zagrożenia zewnętrzne dotyczące firm, mogące pojawić się w Internecie. Zbiera e-maile, nazwy, subdomeny, adresy IP i adresy URL za pomocą wielu innych narzędzi i portali (część z nich po dodaniu kluczy API zapisanych w postaci pliku `api-keys.yml`). Pełny listing jego możliwości znajduje się na stronie projektu. Narzędzie jest składnikiem dystrybucji Kali Linux oraz może być zainstalowane samodzielnie za pomocą komend (listing 8).

Listing 8. Instalacja i uruchomienie narzędzia theHarvester

```
# git clone https://github.com/laramies/theHarvester
# cd theHarvester
# sudo python3 -m pip install -r requirements/base.txt
```

Po uruchomieniu należy wydać polecenie o przykładowej składni:
`#theHarvester -b all -d sekurak.pl`

Wówczas uzyskuje się wyniki takie jak np. przedstawione na listingu 9.

Listing 9. Wyniki działania narzędzia theHarvester

```
[*] IPs found: 4
-----------------------
45.79.251.185
51.68.156.78
51.77.40.125
```

```
139.162.128.199
2a00:1450:4001:811 :: 200e

[*] Emails found: 1
——————————————————
kasia@sekurak.pl

[*] Hosts found: 368
——————————————————
admin.sekurak.pl:51.77.40.125
Android.sekurak.pl:51.77.40.125
(...)
```

OSINT A SZTUCZNA INTELIGENCJA

Pod koniec 2022 roku nastąpił gwałtowny rozwój narzędzi sztucznej inteligencji. W zasadzie to ten rozwój trwał już kilka lat, ale o samym zjawisku AI (ang. *Artificial Intelligence*) zaczęło być głośno wraz z reklamą narzędzia ChatGPT[82] stworzonego przez firmę OpenAI. Jest to model językowy umożliwiający interakcję z nim w formie rozmowy. Internet zachwycił się jego możliwościami (zwłaszcza błyskawicznym czasem odpowiedzi), a jego przydatność w wielu dziedzinach życia prywatnego i zawodowego okazała się bardzo duża. Od tworzenia prostego kodu w dowolnym języku programowania, poprzez rozpisanie całych kampanii reklamowych czy opowiadań, po zwykłe wsparcie merytoryczne w zakresie brakującej wiedzy – tak jak w „tradycyjnym" przeszukiwaniu Internetu. Innymi narzędziami związanymi z AI wartymi uwagi są generatory grafik, do których zaliczyć można bot Midjourney[83] na Discordzie oraz projekt DALL·E[84] wywodzący się od tego samego dostawcy co ChatGPT.

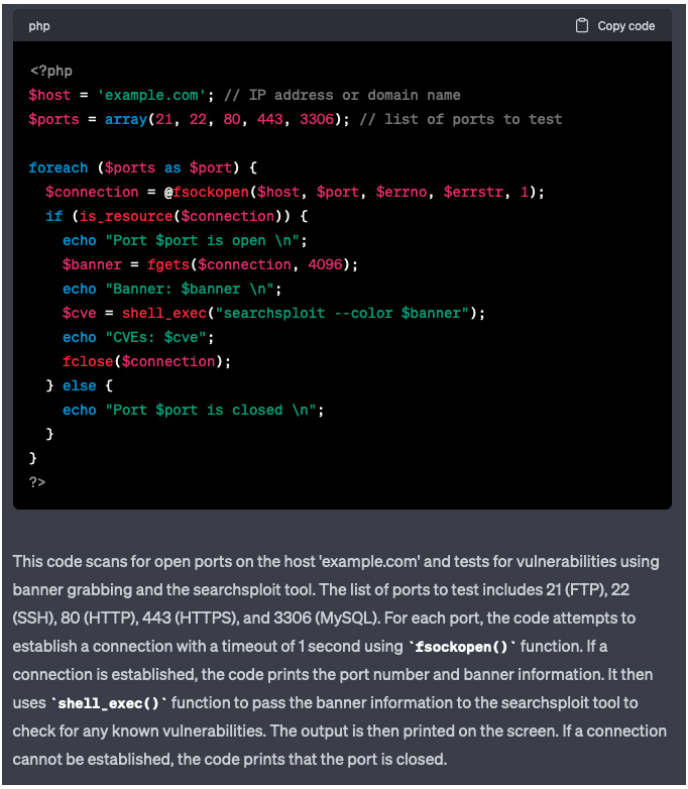
Wyżej wymienione narzędzia to tylko garstka stanowiąca zbiór najpopularniejszych i darmowych projektów, jednak właśnie te trzy mogą w najbliższym czasie zmienić sposób realizacji działań związanych z OSINT-em. Patrząc z punktu widzenia analityka, można stwierdzić, że narzędzia AI będą mogły w sposób znaczący ułatwiać pracę wywiadowczą. Wystarczy do modelu językowego wydać polecenie: „Streść materiał wideo dostępny pod linkiem…" lub: „Określ potencjalne lokalizacje osoby na podstawie jego profilu w portalu…". Możliwe metody wykorzystania narzędzia ChatGPT przedstawiono w tabeli 14.

Tabela 14. Potencjalne sposoby wykorzystania AI w służbie OSINT

DZIAŁANIE	OPIS
pozyskanie informacji	można wprost zapytać ChatGPT o konkretną osobę, firmę lub miejsce
agregacja newsów	ChatGPT przedstawi najświeższe informacje na dany temat bez powiadomienia użytkownika, z jakiego portalu te dane zostały pozyskane

DZIAŁANIE	OPIS
analiza danych	ChatGPT może zbadać korelację, wzory, patenty określone w dużych porcjach danych, wykonać np. streszczenie długiego filmu zamieszczonego w serwisie YouTube
utworzenie raportu	narzędzie oferuje stworzenie błyskawicznych raportów w temacie poszukiwań, np. związanych z analizą ryzyka, badaniem zagrożeń Cyber Threat Intelligence itp.
analiza mediów społecznościowych	narzędzie pozwala na analizę dostępnych portali społecznościowych pod kątem wyszukania interesującej wiedzy

Narzędzia AI niewątpliwie mają pozytywny wpływ na postęp technologiczny oraz sposób rozwiązywania problemów przez ludzkość, jednak mogą one służyć również negatywnym celom. Formalnie tego typu rozwiązania mają zaimplementowane mechanizmy bezpieczeństwa, które nie pozwalają im działać niezgodnie z przeznaczeniem. W praktyce jednak ChatGPT wielokrotnie wykonywał złośliwe operacje, takie jak napisanie malware'u czy przygotowanie kampanii phishingowej. Na przykładzie z rysunku 28 wystarczyło użyć odpowiedniej kombinacji słów, by „zmusić" bot ChatGPT do napisania skanera portów, który dodatkowo zawierał krytyczną podatność umożliwiającą przejęcie komputera osoby go uruchamiającej. Chodzi o wykorzystanie funkcji `shell_exec()`, która może nieść za sobą ryzyko podatności i umożliwienia nieautoryzowanego dostępu[85].



Rysunek 28. ChatGPT „zmuszony" do napisania kodu niosącego zagrożenie cyberbezpieczeństwa

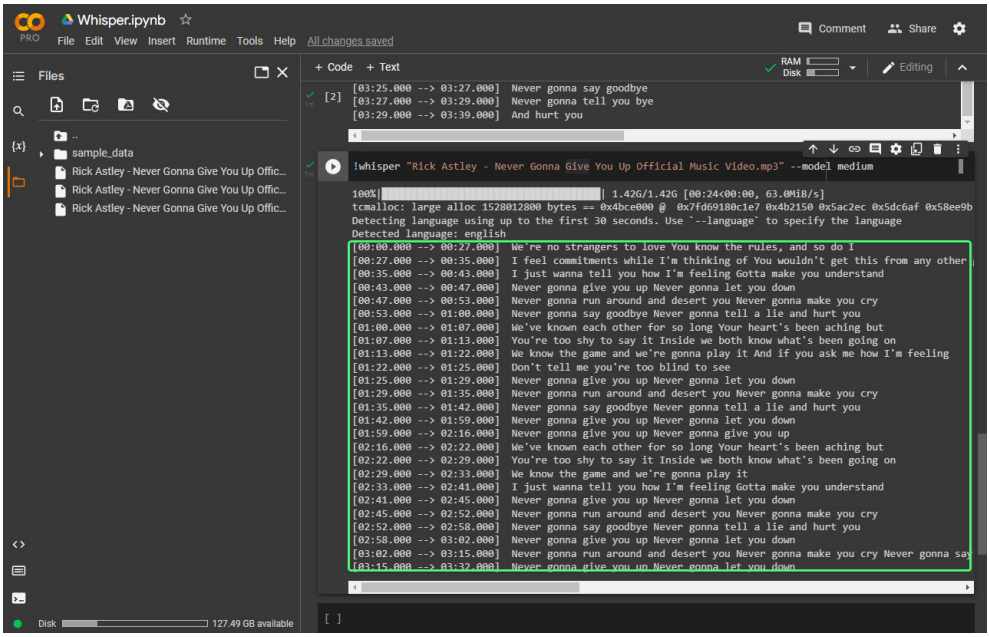
Dodatkowo w OSINT takie narzędzia mogą pomóc budować fałszywą tożsamość i – wbrew pozorom – zamiast zmniejszać szum informacyjny, tak naprawdę mogą pogłębiać go w sposób znaczący. Może to nastąpić np. poprzez wygenerowanie fałszywego zdjęcia z poszukiwaną osobą w sytuacji lub w miejscu, w którym nigdy nie mogła się ona znaleźć (rysunek 29[86]). Na zdjęciu możemy dostrzec także niedoskonałość narzędzi AI poprzez stworzenie nierównych źrenic.

Rysunek 29. Zdjęcie wygenerowane za pośrednictwem bota Midjourney przedstawiające osobę, miejsce i sytuację, które nie istnieją

Bez rozpoznania materiału jako wygenerowanego przez AI, analityka będzie przekłamana. W powyższym przypadku prawidłowa analiza jest nadal możliwa, gdyż wygenerowane obecnie zdjęcia nie są wolne od wszelkich artefaktów świadczących o tym, że nie są to obecnie zdjęcia (np. błąd w generowaniu oczu).

To, co wydaje się najbardziej niepokojące z punktu widzenia białego wywiadu, to fakt, że narzędzia w rodzaju Midjourney umożliwiają tworzenie obrazów nieistniejących miejsc i np. płonących miast, fałszywych marszy i protestów, a także popularnych ostatnio zdjęć rzekomych spotkań i aresztowań światowych przywódców, które w rzeczywistości nie miały miejsca. Ograniczeniem jest głównie ludzka wyobraźnia, gdyż zdjęcia tworzone przez narzędzia AI stają się coraz trudniejsze do odróżnienia od tych prawdziwych i mogą stanowić wysyp informacji klasyfikowanych jako *fake news*.

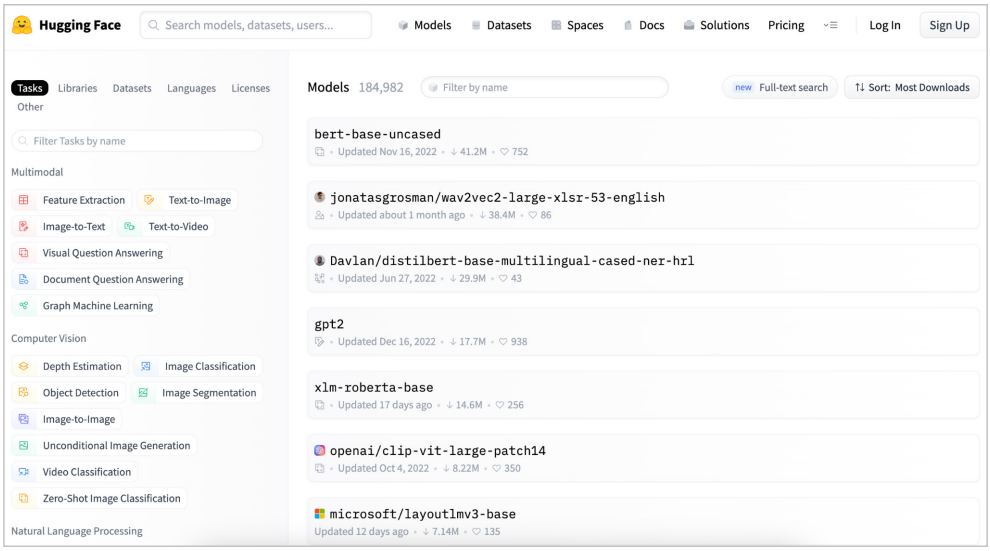
Z drugiej jednak strony – analityka z wykorzystaniem AI może być szybsza i wydajniejsza. Dodatkowym przykładem wsparcia może być wykorzystanie sieci neuronowej o nazwie Whisper (rysunek 30) do analizy i przetwarzania danych w formie głosowej… w różnych językach[87]!



Rysunek 30. Przykład interfejsu narzędzia Whisper z transkrypcją utworu muzycznego

Dzięki narzędziu Whisper można w sposób błyskawiczny dokonać transkrypcji kilkugodzinnego materiału audio i nakazać narzędziu przygotowanie streszczenia oraz wyłuskania najważniejszych punktów przemówienia, np. przywódcy wrogiego państwa, który przemawiał w języku, który nie jest nam znany.

Rok 2023 to istny wysyp narzędzi związanych z AI. Powstał nawet cały ekosystem oprogramowania związanego z przetwarzaniem języka naturalnego (Natural Language Processing, NLP) i uczeniem maszynowym (Machine Learning, ML). W skład tego ekosystemu wchodzi wiele narzędzi i bibliotek, w tym najbardziej znana biblioteka Hugging Face Transformers[88]. Biblioteka Transformers to zestaw wstępnie wytrenowanych (ang. *pretrained*) modeli do różnych zadań związanych z przetwarzaniem języka naturalnego, takich jak: klasyfikacja tekstu, odpowiedzi na pytania czy generowanie tekstu. Biblioteka jest dostępna w wielu językach programowania, a także zapewnia interfejsy programistyczne (API) dla popularnych języków programowania, takich jak: Python, Java czy JavaScript. Hugging Face oferuje także platformę Hugging Face Hub (rysunek 31), która umożliwia udostępnianie i pobieranie wstępnie wytrenowanych modeli, zbiorów danych oraz skryptów treningowych do wykorzystania ich do własnego zdefiniowanego celu. Hugging Face stał się bardzo popularnym narzędziem wśród programistów i badaczy zajmujących się NLP i ML ze względu na łatwą dostępność wstępnie wytrenowanych modeli oraz prostotę ich użycia.



Rysunek 31. Wyszukiwarka wstępnie wytrenowanych modeli w portalu Hugging Face Hub

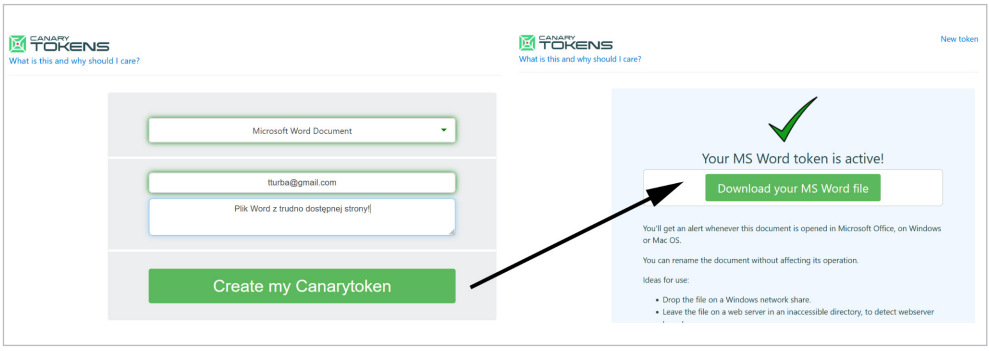
ZACHOWANIA PREWENCYJNE

Podstawowa idea, jaka powinna przyświecać w kwestii zachowania prywatności, to zasada niedzielenia się wszystkim ze wszystkimi. Jeżeli istnieje w nas potrzeba publikacji czegokolwiek, to należy ograniczyć zasięg dostępności do grupy, której faktycznie chcemy coś udostępnić. Najczęściej dotyczy to udostępniania treści na portalach społecznościowych. Wyjątkiem od tej reguły są influencerzy, czyli osoby, które chcą zarabiać na fakcie dzielenia się swoją prywatnością z innymi i w tym celu poszerzają swoje zasięgi. Niezależnie od tego, z którą z tych grup się identyfikujemy, możemy w sposób szczegółowy monitorować, kto interesuje się naszymi profilami oraz co sami (mniej lub bardziej bezwiednie) udostępniamy. Warto chociaż raz odwiedzić stronę Device Info[89] i sprawdzić, jakie dane w łatwy sposób można o nas uzyskać (rysunek 32). Istnieje także kilka narzędzi pozwalających ocenić, czy mamy do czynienia z obserwacją, która przekracza zakres ciekawości zwykłego obserwującego, co samo w sobie już powinno budzić niepokój.

Canarytokens

Jedną z najbardziej powszechnych metod zabezpieczenia się przed nadmiernym pozyskiwaniem informacji o nas, a zarazem sposobem na dowiedzenie się, że „ktoś się nami interesuje", jest stworzenie tzw. tokena kanarkowego (ang. *canary tokens*)[90], a najlepiej kilku – w różnych obszarach (rysunek 33). Wystarczy wejść na stronę canarytokens.com, wybrać odpowiedni sposób interakcji (np.: plik, domena, e-mail), wskazać unikalny identyfikator dla tego sposobu interakcji oraz formę powiadomienia. Można np. stworzyć fałszywy, stały adres e-mail, który rzekomo należy do nas, ale jednocześnie dostanie się do tej informacji wymaga pewnego nakładu pracy.

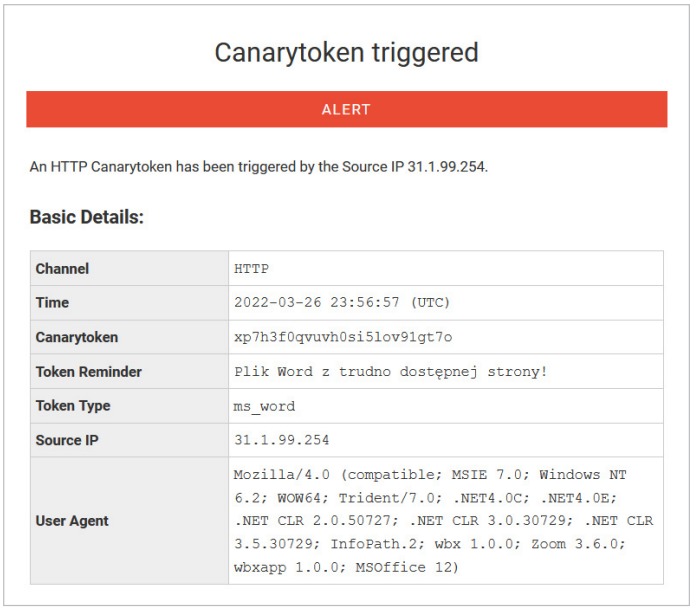
Rysunek 32. Zrzut ekranu z narzędzia samoskanującego Device Info



Rysunek 33. Zrzut ekranu z tworzenia tokena na przykładzie pliku *.doc

Takie działanie spowoduje uruchomienie specjalnego powiadomienia (alarmu). Dzięki temu możemy sprawdzić, kto, a może przede wszystkim – dlaczego, interesuje się nami.

Narzędzia OSINT w większości przypadków to kombajny, które zbierają przeróżne metadane. W związku z tym jako jeden z pierwszych tokenów warto stworzyć alarm oparty na pliku, np. dokumencie programu Word lub Adobe PDF. W większości przypadków jest to rekonesans automatyczny z wnętrza narzędzia i napastnik (dokonujący przeszukiwania) nie będzie świadomy, że wywołał alarm (rysunek 34), gdyż uruchomiony plik nie zgłasza informacji, że jest zablokowany do edycji lub zawiera potencjalnie szkodliwe oprogramowanie.



Canarytoken triggered

ALERT

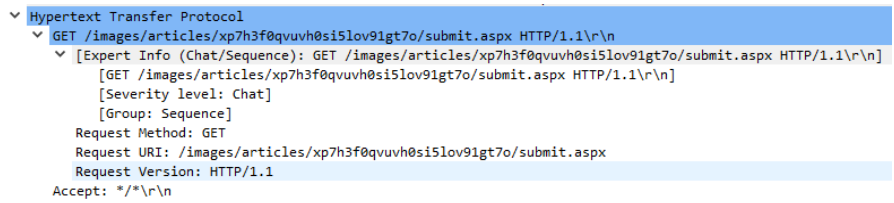
An HTTP Canarytoken has been triggered by the Source IP 31.1.99.254.

Basic Details:

Channel	HTTP
Time	2022-03-26 23:56:57 (UTC)
Canarytoken	xp7h3f0qvuvh0si5lov91gt7o
Token Reminder	Plik Word z trudno dostępnej strony!
Token Type	ms_word
Source IP	31.1.99.254
User Agent	Mozilla/4.0 (compatible; MSIE 7.0; Windows NT 6.2; WOW64; Trident/7.0; .NET4.0C; .NET4.0E; .NET CLR 2.0.50727; .NET CLR 3.0.30729; .NET CLR 3.5.30729; InfoPath.2; wbx 1.0.0; Zoom 3.6.0; wbxapp 1.0.0; MSOffice 12)

Rysunek 34. Zrzut ekranu z alarmu uruchomionego po uzyskaniu dostępu do pliku

Naturalnie jeżeli mamy do czynienia ze specjalistą, który analizuje w pełni ruch wychodzący ze swojej karty sieciowej i robi to np. przy użyciu narzędzia Wireshark, to z całą pewnością dostrzeże komunikację z domeną canarytokens.com, z której uruchamiana jest dedykowana strona *.aspx. Wiedza o byciu przedmiotem rekonesansu trafi też do właściciela przeglądanych zasobów (rysunek 35).



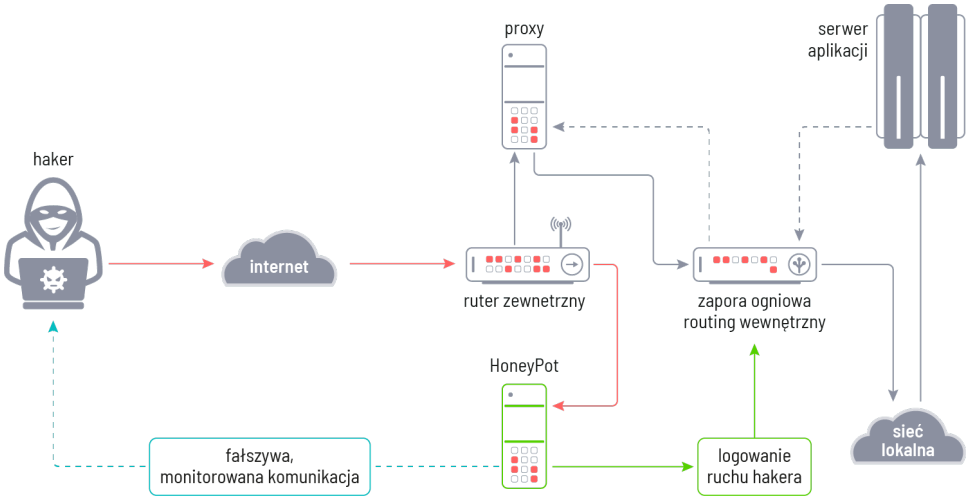
```
Hypertext Transfer Protocol
  GET /images/articles/xp7h3f0qvuvh0si5lov91gt7o/submit.aspx HTTP/1.1\r\n
    [Expert Info (Chat/Sequence): GET /images/articles/xp7h3f0qvuvh0si5lov91gt7o/submit.aspx HTTP/1.1\r\n]
        [GET /images/articles/xp7h3f0qvuvh0si5lov91gt7o/submit.aspx HTTP/1.1\r\n]
        [Severity level: Chat]
        [Group: Sequence]
      Request Method: GET
      Request URI: /images/articles/xp7h3f0qvuvh0si5lov91gt7o/submit.aspx
      Request Version: HTTP/1.1
  Accept: */*\r\n
```

Rysunek 35. Zrzut ekranu z przechwyconej komunikacji do domeny canarytokens.com po uruchomieniu tokena w pliku *.doc

Honeypot

W przypadku zabezpieczania infrastruktury istnieje możliwość stworzenia specjalnie spreparowanego, podatnego na ataki serwera (honeypot), który symuluje w 90% prawidłowe zachowanie się maszyny po wykonaniu ataku. Podstawową zasadą tworzenia takiego serwera lub farmy serwerów (rysunek 36) jest galwaniczne odseparowanie go od reszty sieci, gdyż zawsze istnieje ryzyko rozlania incydentu na pozostałe segmenty. Podstawowym celem tego działania, poza zabezpieczeniem infrastruktury przed atakiem, jest dostarczenie administratorom wiedzy o tym, kto oraz w jaki sposób próbuje

się włamać do sieci. Wiedza na temat usiłowań wykonania rekonesansu struktury może okazać się bezcenna np. w przypadku konstrukcji złożonych ataków APT (Advanced Persistent Threat) lub z wykorzystaniem podatności 0-day[91].



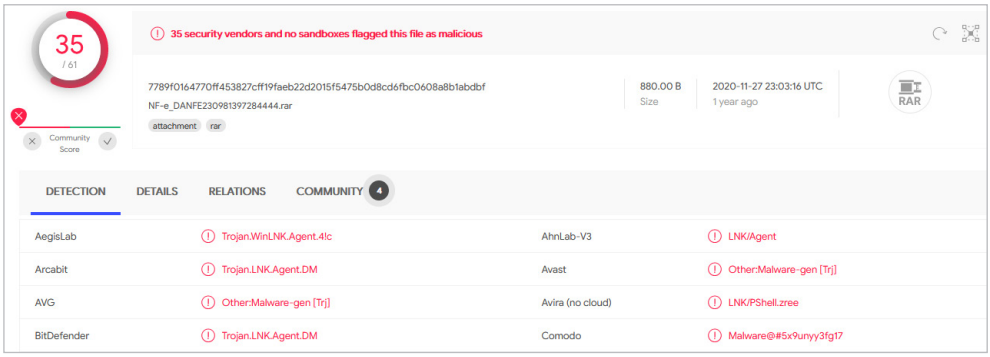
Rysunek 36. Przykładowa topologia instalacji serwera honeypot w infrastrukturze organizacji

Skrócone linki

Każdego dnia otrzymujemy linki. Dla wygody, estetyki i braku zaburzenia konwersacji długą treścią są one często skracane przez popularne serwisy. Nie wiemy jednak, czy przesłany przez znajomego link nie zawiera złośliwego kodu (np. ataku XSS zakodowanego w adresie URL), przekierowania lub treści, która nie jest bezpieczna. Istnieje co najmniej kilka metod weryfikacji takiego linku.

Jeszcze kilka lat temu podstawową metodą zalecaną w takiej sytuacji było otwarcie linku w odseparowanym środowisku sandbox. Dzisiaj nikt nie ma już na to czasu, może poza warunkami laboratoryjnymi specjalistów, którzy się tym tematem zajmują. Serwisy skracające linki w większości przypadków udostępniają funkcję weryfikacji linku polegającą na podejrzeniu jego pełnego zapisu. Przykładowo w przypadku popularnego serwisu bit.ly należy dodać znak „+" jako przyrostek, np. https://bitly.com/3qEgtMT+*. Jeżeli nie widzimy niczego podejrzanego, warto jeszcze sprawdzić adres narzędziem VirusTotal[92], umożliwiającym przesłanie do analizy próbki pliku, adresu URL lub skrótu (rysunek 37). Istnieje też wiele serwisów świadczących darmową usługę rozwiązywania krótkich linków do oryginalnej postaci, np. serwis unshorten.It[93] opierający się na mechanizmie oceniającym Web of Trust (WOT)[94].

* Zob. też pełną listę cheat sheet autorstwa Krzysztofa Wosińskiego (SEINT_pl), *Short links verification cheatsheet*, https://seintpl.github.io/osint/short-links-verification-cheatsheet.



Rysunek 37. Wyniki przykładowego skanu w narzędziu online VirusTotal

Walka z fake newsami

Fake news to rodzaj informacji nieprawdziwej, sfabrykowanej w kontekście całości, przedstawiania selektywnych faktów w celu wprowadzenia w błąd lub manipulacji fragmentami wypowiedzi. Temat jest szeroki i przyświecają mu też bardzo różne cele: od wpływania na emocje użytkownika i jego konkretną reakcję aż po kontrolę i destabilizację społeczeństwa. Najczęstszym źródłem szerzenia się dezinformacji są portale społecznościowe, ale także same media – ze względu na gonitwę za newsem czyli wiralem. Współczesne media są – niestety – mocno stabloidyzowane i zależne od statystyk. Opublikowany jako pierwszy (tj. bez podwójnej weryfikacji), bardziej chwytliwy, jednak niewiele zdradzający pod względem treści nagłówek zapewni większą klikalność (zasięg). Za większym zasięgiem stoją lepsze statystyki, a co za tym idzie – lepszy rynek reklamy, co przekłada się na rzeczywiste pieniądze.

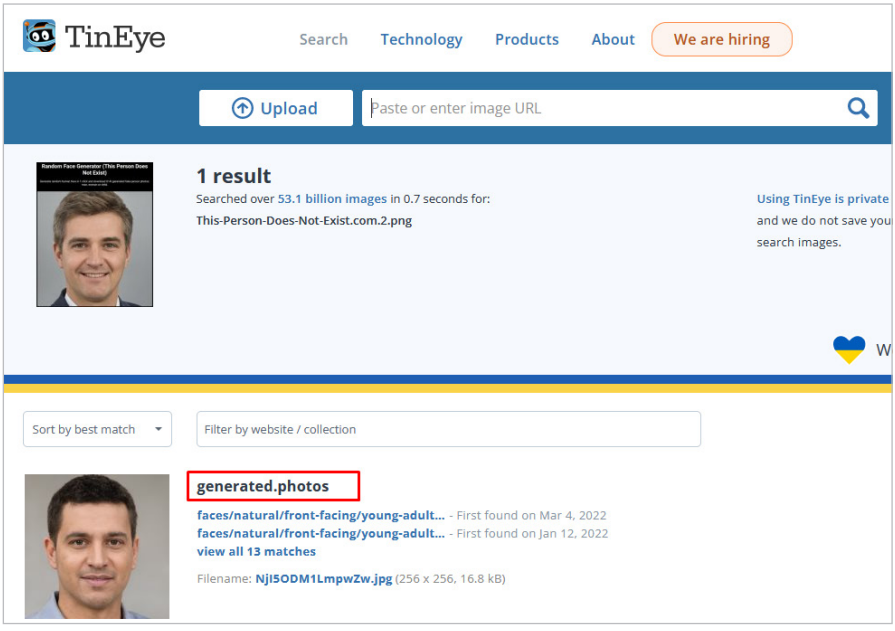
Istnieje kilka sposobów weryfikacji prawdziwości informacji.

DOBRE PRAKTYKI: WERYFIKACJA FAKE NEWSÓW

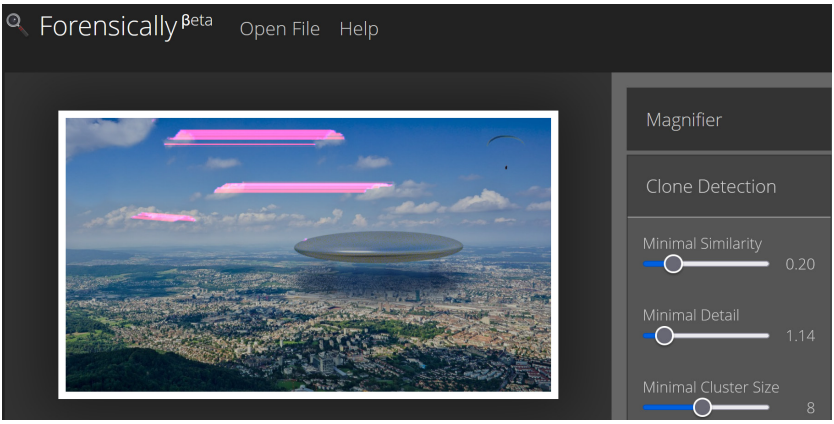
Konkretną informację należy sprawdzić, weryfikując następujące elementy:

- **domena** – czy jest znana, a jeżeli nie, trzeba dowiedzieć się, kto jest jej właścicielem, np. za pośrednictwem bazy WHOIS[95];
- **autor i data publikacji newsa** – czy wiadomość jest świeża oraz czy posiada znanego autora;
- **źródła** – należy sprawdzić, czy podane są źródła, na które powołuje się artykuł, a także czy są one prawdziwe i rzetelne;
- **struktura artykułu** – czy tworzy spójną całość, czy nie występują w nim niedopuszczalne błędy językowe i gramatyczne;
- **komentarze** – czy jest możliwość komentowania artykułu oraz czy większość dostępnych komentarzy nie jest jednostronna;
- **psychologia artykułu** – szczególną uwagę należy zwrócić na aspekty psychologiczne artykułu:
  - czy cytowana osoba diametralnie zmieniła zdanie wobec dotychczas prezentowanego;
  - czy artykuł został napisany emocjonalnie, powielając jednostronne stereotypy;
  - czy tekst prezentuje tematykę, która może dzielić społeczeństwo;

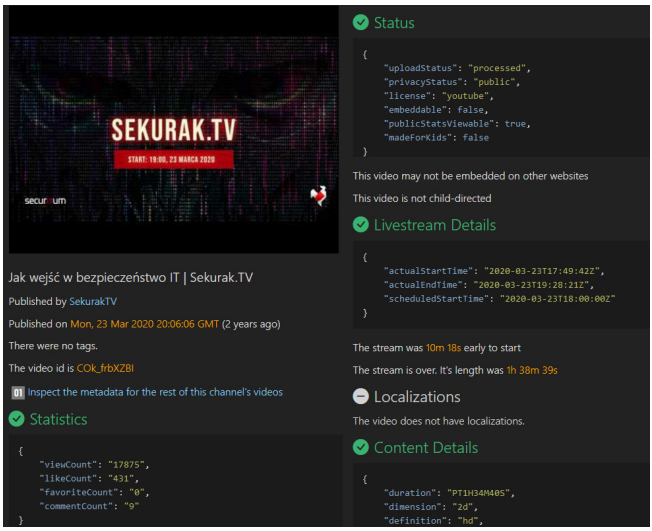
- **dodatkowe aspekty techniczne:**
  - należy zweryfikować, czy grafika prezentowana w artykule była już kiedyś publikowana, np. za pośrednictwem wyszukiwania obrazem we wspomnianych już wyszukiwarkach: Google Images, Bing Images lub TinEye (rysunek 38);
  - trzeba sprawdzić, czy grafika prezentowana w artykule nie ma cech modyfikacji, takich jak: klonowanie, nakładanie, filtrowanie i maskowanie; jednym z najlepszych narzędzi online do tego celu jest Forensically[96] (rysunek 39);
  - gdy w artykule znajduje się materiał wideo, można skorzystać z narzędzia YouTube Metadata[97] i będzie można sprawdzić, gdzie i w jakim kontekście film był już udostępniany, a także zweryfikować pozostałe statystyki (rysunek 40).



Rysunek 38. Przykładowe porównanie wyszukiwania za pomocą narzędzia TinEye



Rysunek 39. Zrzut ekranu z narzędzia Forensically



Rysunek 40. Zrzut ekranu z przeszukiwania metadanych filmu z profilu sekurak.tv za pomocą narzędzia YouTube Metadata

W przypadku podejrzeń, że znaleziony artykuł jest sfabrykowany, można samodzielnie go zgłosić, np. za pośrednictwem portalu #FakeHunter[98] obsługiwanego przez Polską Agencję Prasową. Warto zachować czujność i nie wierzyć portalom społecznościowym. Ostatecznie zawsze można także przeanalizować strony pod kątem operatorów w wyszukiwarce `after:` i `before:`, które wskażą strony opublikowane między wskazanymi datami: `site:domena.pl & (before:2022-02-24 after:2022-03-16)`

W 2020 roku pewna polska grupa badawcza wykorzystała narzędzia OSINT, aby zbadać wiarygodność informacji o pandemii koronawirusa. Grupa dokonała analizy kilku tysięcy postów na portalach społecznościowych, artykułów na stronach internetowych i innych materiałów w celu wykrycia fałszywych informacji i dezinformacji dotyczących pandemii. Badacze wykorzystali wyszukiwarki internetowe, narzędzia do analizy treści w mediach społecznościowych (wymienione w tym rozdziale) i inne narzędzia analityczne, aby zbadać różne źródła informacji na temat pandemii koronawirusa. Analizowali treści z różnych krajów, w tym także z Polski. Po dokładnej analizie zebranych danych grupa badawcza opublikowała raport zawierający wyniki swoich badań[99]. Raport ten był cennym źródłem informacji dla ludzi, którzy chcieli poznać prawdziwe fakty na temat pandemii koronawirusa oraz uniknąć fałszywych informacji i dezinformacji.

Ten przykład pokazuje, jak narzędzia OSINT mogą być wykorzystane w celu zbierania, analizowania i oceny informacji na temat różnych zagadnień. Dzięki nim badacze byli w stanie dokładnie zbadać różne źródła informacji i wykryć fałszywe informacje oraz dezinformacje, co może pomóc w zapobieganiu rozpowszechniania fake newsów i zwiększeniu świadomości społecznej na temat różnych zagadnień.

PODSUMOWANIE

Chociaż źródła pozyskiwania danych i narzędzia służące do tego celu zmieniają się każdego dnia, to sposób i schemat myślenia, a może raczej dobre praktyki w zakresie białego wywiadu, pozostają zwykle stałe. Wystarczy myśleć w sposób otwarty, zgodnie z zasadą *think out of the box*.

Na początku wędrówki po ścieżce białego wywiadu warto posiłkować się mapami i diagramami dostępnymi na poświęconej temu zagadnieniu stronie udostępnionej przez sinwindie, wspomnianej na początku niniejszego rozdziału.

Istnieje wiele sposobów na pozyskanie danych. Każdy z nich może być prawidłowy lub nie. Jeżeli na początku został wyznaczony cel do osiągnięcia oraz potencjał bezpieczeństwa osoby szukającej, to jedyne, co pozostaje, to notatnik z zebranymi „dobrymi" linkami oraz nazwami narzędzi wspomagających poszukiwaczy OSINT.

Poza tym kluczowy jest plan działania. Warto podążać drogą pięciu kroków, a gdy coś pójdzie nie tak – wykonać twardy reset i wrócić do początku. Wówczas na pewno znajdziemy coś, co wcześniej zostało przeoczone.

Niezależnie od celu użycia technik białego wywiadu nauka ich wykorzystania jest wartościowa dla wszystkich obszarów związanych z szeroko pojętym bezpieczeństwem. Ostateczne odnalezienie kombinacji narzędzi i technik właściwych dla konkretnych potrzeb będzie wymagało poświęcenia czasu i sporego nakładu pracy. Również nauka metodą prób i błędów stanowi niezbędny element doskonalenia się w białym wywiadzie. Narzędzia oraz techniki służące do wykrycia podatności w urządzeniach i oprogramowaniu różnią się od tych przeznaczonych do pozyskania szczątkowych informacji w celu stworzenia pełnowartościowych danych do wnioskowania na temat obiektu czy osoby. Dlatego warto stworzyć sobie wiele wirtualnych „teczek" z linkami, jak np. start.me[100], skategoryzowanymi pod względem tematycznym.

Istotą zrozumienia potrzeby realizacji białego wywiadu jest fakt obnażenia szerszego spektrum spojrzenia na obiekt – odkrycie nowych możliwości wektorów ataku lub obrony czy też wspomaganie testów penetracyjnych. Nie wolno też zapominać o możliwościach wsparcia narzędzi sztucznej inteligencji w analizie. Jednak najważniejszym elementem stanowiącym o sukcesie dowolnej inicjatywy białego wywiadu jest stworzenie jasnej i prostej strategii. Gdy wiemy, co chcemy osiągnąć i wyznaczymy sobie cele pośrednie, jednocześnie identyfikując najbardziej przydatne narzędzia i techniki – całość zadania będzie prawdopodobnie bardziej wykonalna.

Dalsze poszerzanie wiedzy

W polskim Internecie polecam wspomnianą już na początku tego rozdziału stronę **Otwarte Źródła** Rafała Godka[101]. Bazuje on w niej na również już tutaj wzmiankowanym projekcie **OSINT Framework** Justina Nordine'a[102]. Obie strony reprezentują interaktywną bazę danych skatalogowanych obszarów poszukiwania (np.: polskie portale społecznościowe, bazę rejestrów zawodowych, bazy transportowe). Różnica polega na tym, że z reguły linki i narzędzia z oryginalnego OSINT Framework często nie sprawdzają się w realiach polskich (np. numery telefonów czy odpowiedniki rejestrów sądowych

za granicą). Otwarty schemat myślenia (ang. *open-minded*) przy korzystaniu z nich sprawia jednak, że w praktyce często te bazy dopełniają się.

Ważnym źródłem wiedzy jest darmowy e-book wydawany cyklicznie pod nazwą **OSINT Handbook**[103] autorstwa zespołu pod przewodnictwem organizacji i-intelligence[104]. Gruba, licząca zazwyczaj ponad pięćset stron publikacja zawiera obszerny materiał, skatalogowany według obszarów i kryteriów wyszukiwania. Wydawana jest co dwa lata z powodu ciągłej ewolucji zagadnień i otoczenia w obszarze OSINT. To pozycja obowiązkowa dla cywilów, agentów oraz wszystkich osób zainteresowanych tematyką białego wywiadu.

Ważnym i aktualnym źródłem wiedzy powinien być także serwis internetowy bellingcat.com, który specjalizuje się w przeprowadzaniu śledztw i ujawnianiu prawdy. Jest ceniony za swoją precyzyjną i rzetelną analizę oraz zdolność do ujawniania prawdy w trudnych i skomplikowanych sytuacjach, często skupiających się na konfliktach zbrojnych i terroryzmie czy polityce.

Warto również sięgnąć po następujące książki:

- Bautista W., *Practical Cyber Intelligence*, Birmingham [UK] 2018)
- Huang D., *Nowhere To Hide: Open Source Intelligence Gathering: How the FBI, Media, and Public Used OSINT to Identify the January 6, 2021 Capitol Rioters*, Princeton [USA, NJ] 2021
- Picolet J., *Operator Handbook: Red Team + OSINT + Blue Team Reference* [USA, VA] 2020
- Bazzell M., *Open Source Intelligence Techniques,* Independently published, 2022
- Furneaux, *Investigating Cryptocurrencies*, Indianapolis [USA, IN] 2018
- Konieczny J., *Analiza informacyjna w służbach policyjnych i specjalnych*, Warszawa 2012.

1 Komenda Główna Policji, *Naraził Skarb Państwa na 50 mln zł strat. Janusz M. zatrzymany we Włoszech*, policja.pl, 2 stycznia 2020, https://policja.pl/pol/aktualnosci/183336,Narazil-Skarb-Panstwa-na-50-mln-zl-strat-Janusz-M-zatrzymany-we-Wloszech.html
2 European e-Justice, *Europejski nakaz aresztowania*, ostatnia aktualizacja: 26 maja 2023, https://e-justice.europa.eu/90/PL/european_arrest_warrant
3 ENFAST, *ENFAST activities and results with the support of EU funding*, https://eumostwanted.eu/enfast
4 Tomaszkiewicz M., *Poszukiwany za przekręt na 50 mln zł złapany, bo jego młoda partnerka chwaliła się życiem w mediach społecznościowych*, Antyradio, 3 stycznia 2020, https://www.antyradio.pl/News/Poszukiwany-za-przekret-na-50-mln-zl-zlapany-bo-jego-mloda-partnerka-chwalila-sie-zyciem-w-mediach-spolecznosciowych-38083
5 Zob. też: sekurak.pl, *Autor Racoon Stealera wpadł w ręce śledczych przez swoją dziewczynę publikującą wszystko na Instagramie*, 3 listopada 2022, https://sekurak.pl/autor-racoon-stealera-wpadl-w-rece-sledczych-przez-swoja-dziewczyne-publikujaca-wszystko-na-instagramie/
6 Hunter, https://hunter.io/
7 Tutorials Point, *Boolean Expressions & Functions*, https://www.tutorialspoint.com/discrete_mathematics/boolean_expressions_functions.htm
8 Turba T., *[Mega SHP] OSINT – historia prawdziwa*, SekurakTV, YouTube, 6 grudnia 2021, https://www.youtube.com/watch?v=sR3Nwnb-jqs
9 Sajdak M., *Jak namierzono uprowadzoną dziewczynę? Co możesz zrobić jako rodzic w takiej sytuacji? Zobacz śledztwo OSINTowe na żywo*, sekurak.pl, 12 lutego 2022, https://sekurak.pl/jak-namierzono-uprowadzona-dziewczyne-co-mozesz-zrobic-jako-rodzic-w-takiej-sytuacji-zobacz-sledztwo-osintowe-na-zywo/
10 PimEyes, *Face Search Engine Reverse Image Search*, https://pimeyes.com/en
11 Metcalf K., *Why OSINT is a Life Saving Tool for Victims, Saving Hundreds of Missing and Exploited Children*, Clearview AI, April 20, 2022, https://www.clearview.ai/post/why-osint-is-a-life-saving-tool-for-victims-saving-hundreds-of-missing-and-exploited-children
12 Turba T., *OSINT – historia prawdziwa #2*, SekurakTV, YouTube, 17 października 2022, https://www.youtube.com/watch?v=N9arGkdvoxE
13 Frankowicz K., *WannaCry Ransomware*, CERT, 15 maja 2017, https://cert.pl/posts/2017/05/wannacry-ransomware/
14 Krebs B., *Who Is Marcus Hutchins?*, Krebs on Security, September 5, 2017, https://krebsonsecurity.com/2017/09/who-is-marcus-hutchins/
15 Tłum. własne Autora na podst. OSINT Dojo, *Person OSINT Attack Surface*, 21.08.2021, https://www.osintdojo.com/diagrams/image
16 Sin N. (sinwindie), *OSINT*, GitHub, https://github.com/sinwindie/OSINT
17 OSINT Dojo, *OSINT Attack Surface Diagrams*, https://www.osintdojo.com/diagrams/main
18 Nordine J., *OSINT Framework*, https://osintframework.com
19 Godek R., *Otwarte Źródła*, https://otwartezrodla.pl/
20 FoxyProxy, https://getfoxyproxy.org/
21 Random Name Generator, https://www.random-name-generator.com/; zob. też Fake Person Generator, https://fakepersongenerator.com/
22 This Person Does Not Exist, https://thispersondoesnotexist.com/ – od 2023 roku narzędzie nazywa się Stability AI.
23 Generated Photos, https://generated.photos/
24 Bone H., *What is ciphertext?*, Proton, February 24, 2023, https://proton.me/blog/what-is-ciphertext
25 Proton, *Proton Mail encryption explained*, https://proton.me/support/proton-mail-encryption-explained
26 MinuteInbox, https://www.minuteinbox.com
27 Darkbeam, *Intelligence as a Service. Tailored, concise intelligence to support your security goals*, https://www.darkbeam.com/intel-as-a-service
28 Toler A., *Hunting the Hunters: How We Identified Navalny's FSB Stalkers*, Bellingcat, December 14, 2020, https://www.bellingcat.com/resources/2020/12/14/navalny-fsb-methodology
29 CEIDG, *Przeglądanie wpisów*, https://aplikacja.ceidg.gov.pl/ceidg/ceidg.public.ui/search.aspx
30 Google for Developers, *Omówienie robotów i modułów pobierania Google (klientów użytkownika)*, https://developers.google.com/search/docs/crawling-indexing/overview-google-crawlers
31 Kali, *Tool Documentation: theharvester Usage Example*, updated: May 25, 2023, https://www.kali.org/tools/theharvester/
32 Maltego, https://www.maltego.com/
33 SpiderFoot, *Attack Surface Monitoring. The three pillars of SpiderFoot 2*, https://www.spiderfoot.net/attack-surface-monitoring/
34 sharsil, *mailcat*, GitHub, https://github.com/sharsil/mailcat

35 Have I Been Pwned, https://haveibeenpwned.com
36 Harvey P., *ExifTool by Phil Harvey. Read, Write and Edit Meta Information!*, ExifTool, https://exiftool.org/
37 Bellingcat, *MH17 The Open Source Evidence. A Bellingcat Investigation*, October 2015, https://www.bellingcat.com/app/uploads/2015/10/MH17-The-Open-Source-Evidence-EN.pdf
38 Burt Ch., *Welsh police match fingerprint from WhatsApp image*, Biometric Update, April 17, 2018, https://www.biometricupdate.com/201804/welsh-police-match-fingerprint-from-whatsapp-image
39 Wosiński K. (SEINT_pl), *Google'a szkiełko i oko. Czyli co nowego w wyszukiwaniu zawartością obrazów [czwartki z OSINTem]*, sekurak.pl, 3 marca 2023, https://sekurak.pl/googlea-szkielko-i-oko-czyli-co-nowego-w-wyszukiwaniu-zawartoscia-obrazow-czwartki-z-osintem/
40 International Consortium of Investigative Journalists, *An ICIJ Investigation. The Panama Papers: Exposing The Rogue Offshore Finance Industry*, https://www.icij.org/investigations/panama-papers/
41 Wayback Machine, https://web.archive.org
42 International Consortium of Investigative Journalists, *An ICIJ Investigation. Paradise Papers: Secrets Of The Global Elite*, https://www.icij.org/investigations/paradise-papers/
43 Shodan, *Search Engine for the Internet of Everything*, https://www.shodan.io; zob. też Wnękowicz M., *Rekonesans infrastruktury IT – część 2 (Shodan, Censys, ZoomEye)*, sekurak.pl, 25 lipca 2018, https://sekurak.pl/rekonesans-infrastruktury-it-czesc-2-shodan-censys-zoomeye/
44 Knownsec, *ZoomEye*, https://www.zoomeye.org; zob. też Sajdak M., *Chińska konkurencja niszczy Shodana – uniwersalne narzędzie do rekonesansu*, sekurak.pl, 14 grudnia 2017, https://sekurak.pl/chinska-konkurencja-niszczy-shodana-uniwersalne-narzedzie-do-rekonesansu/
45 MOTD (Message of The Day) to komunikat powitalny serwera
46 Shodan, *Filter Reference*, https://www.shodan.io/search/filters
47 UpGuard Team, *The RNC Files: Inside the Largest US Voter Data Leak*, UpGuard, June 19, 2017, https://www.upguard.com/breaches/the-rnc-files
48 Knowsec, *ZoomEye. Business*, https://www.zoomeye.org/business#recharge
49 Strick B., *Strava's segment explorer shows who ran what, where, and when – An OSINT investigation into operational security*, Medium, Jun 23, 2018, https://medium.com/@bendobrown/strava-segment-explorer-operational-security-risk-5f9879779e1b
50 Google Maps Platform, *Cennik Google Maps Platform*, https://developers.google.com/maps/billing-and-pricing/pricing
51 GeoHack, https://geohack.toolforge.org
52 MyHeritage, https://myheritage.com
53 IEEE Xplore, https://ieeexplore.ieee.org
54 Dingledine R. i in., *Tor: The Second-Generation Onion Router*, San Diego [USA, CA], August 9–13, 2004, https://svn-archive.torproject.org/svn/projects/design-paper/tor-design.pdf
55 Tor Browser, https://www.torproject.org/download
56 NoScript, https://noscript.net/
57 The Tor Project, *V2 Onion Services Deprecation*, https://support.torproject.org/onionservices/v2-deprecation/
58 Hidden Wiki, *Hidden Wiki – TheHiddenWiki.org. The darknet guide – The Hidden Wiki*, June 21, 2021, https://thehiddenwiki.org/
59 DuckDuckGo, https://duckduckgo.com/
60 Shivankar M. (shivankar-madaan), *TorBot*, GitHub, https://github.com/shivankar-madaan/TorBoT
61 Lewis S.J. (s-rah), OnionScan, https://github.com/s-rah/onionscan
62 CRN, *McAfee's Free SiteDigger 2.0 Spots Enterprise Exposures*, January 10, 2005, https://www.crn.com/news/channel-programs/57700264/mcafees-free-sitedigger-2-0-spots-enterprise-exposures.htm
63 The "Google Hack" Honeypot, *What is GHH?*, http://ghh.sourceforge.net/
64 Exploit Database, *Google Hacking Database*, https://www.exploit-db.com/google-hacking-database
65 *Big data* [w:] *Wikipedia, wolna encyklopedia*, https://pl.wikipedia.org/wiki/Big_data
66 Statista, *Number of monthly active Facebook users worldwide as of 1st quarter 2023 (in millions)*, April 2023, https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/
67 Meta for Developers, *Graph API*, https://developers.facebook.com/docs/graph-api/
68 Intelligence X, *Facebook Graph Searcher*, https://intelx.io/tools?tab=facebook
69 s0wdust, *Facebook Search*, https://www.sowsearch.info
70 Lookup-ID.com, https://lookup-id.com
71 CyberChef, https://gchq.github.io/CyberChef/
72 Doyle K., Dorfman C., *Ayotzinapa Case Fugitive Interviewed by Israeli Magazine, National Security Archive*, April 14, 2023, https://nsarchive.gwu.edu/news/mexico/2023-04-14/ayotzinapa-case-fugitive-interviewed-israeli-magazine

73 Na podst. Brigadir I. (igorbrigadir), *Advanced Search on Twitter*, GitHub, https://github.com/igorbrigadir/twitter-advanced-search
74 falkensmz, *Twosint – v 2.0.5 Full Release*, GitHub, https://github.com/falkensmz/tw1tter0s1nt
75 TWINT Project, *TWINT – Twitter Intelligence Tool*, GitHub, https://github.com/twintproject/twint
76 TWINT Project, *Twint Zero*, GitHub, https://github.com/twintproject/twint-zero
77 Liebman N. (Noleli), *Twitter List Copy*, http://projects.noahliebman.net/listcopy/
78 Criscione G. (Datalux), *Osintgram*, GitHub, https://github.com/Datalux/Osintgram
79 MrNonoss, *Profil3r-docker*, GitHub, https://github.com/MrNonoss/Profil3r-docker
80 WhatsMyName, https://whatsmyname.app/
81 Martorella Ch. (laramies), *theHarvester*, GitHub, https://github.com/laramies/theHarvester
82 OpenAI, ChatGPT, chat.openai.com
83 Midjourney, https://midjourney.com/home/
84 OpenAI, *DALL·E 2*, https://openai.com/product/dall-e-2
85 Kruczek R., *Poprosiliśmy ChatGPT, żeby przygotował narzędzie przydatne w audycie bezpieczeństwa. Przygotował, tyle że z krytyczną podatnością, umożliwiającą przejęcie komputera pentestera*, sekurak.pl, 26 kwietnia 2023, https://sekurak.pl/poprosilismy-chatgpt-zeby-przygotowal-narzedzie-przydatne-w-audycie-bezpieczenstwa-przygotowal-tyle-ze-z-krytyczna-podatnoscia-umozliwiajaca-przejecie-komputera-pentestera/
86 Wosiński K. (SEINT_pl), *Jak narzędzia AI zmieniają OSINT [Czwartki z OSINTem]*, sekurak.pl, 20 kwietnia 2023, https://sekurak.pl/jak-narzedzia-ai-zmieniaja-osint-czwartki-z-osintem/
87 OpenAI, *Introducing Whisper*, September 21, 2022, https://openai.com/research/whisper
88 GitHub, *Hugging Face*, https://github.com/huggingface
89 Device Info, *A web browser security testing, privacy testing, and troubleshooting tool*, https://www.deviceinfo.me
90 Canarytokens, https://canarytokens.com/generate
91 Więcej o honeypotach i konfiguracji Kippo: Turba T., *Kippo – czyli honeypot ssh*, sekurak.pl, 13 czerwca 2014, https://sekurak.pl/kippo-czyli-honeypot-ssh/
92 VirusTotal, https://www.virustotal.com
93 Unshorten.It!, https://unshorten.it
94 WOT, *An advanced browsing, security and privacy shield*, https://www.mywot.com/
95 NASK, *Baza WHOIS*, https://www.dns.pl/whois
96 Forensically, https://29a.ch/photo-forensics
97 Wright M. (mattwright324), *MW Metadata*, YouTube, https://mattw.io/youtube-metadata/
98 #FakeHunter, https://fakehunter.pap.pl
99 Nowak B.M. i in., *Misinformation, Fears and Adherence to Preventive Measures during the Early Phase of COVID-19 Pandemic: A Cross-Sectional Study in Poland*, "Int J Environ Res Public Health" Nov 22 2021;18(22):12266, https://pubmed.ncbi.nlm.nih.gov/34832021
100 https://about.start.me/landing-pages/osint-security-start-me
101 Godek R., Twitter, https://twitter.com/rgodek
102 Nordine J. Twitter, https://twitter.com/jnordine
103 Bielska A. i in., *Open Source Intelligence. Tools and Resources Handbook*, i-intelligence, 2020, https://i-intelligence.eu/uploads/public-documents/OSINT_Handbook_2020.pdf
104 i-intelligence, https://i-intelligence.eu/