AKADEMIA KALISKA

im. Prezydenta Stanisława Wojciechowskiego

ROZPRAWA DOKTORSKA

Znaczenie wywiadu opartego na otwartych źródłach (OSINT) w zapewnieniu bezpieczeństwa systemów teleinformatycznych i bezpieczeństwa osobowego

mgr inż. Krzysztof Wosiński

Napisana
pod kierownictwem naukowym:
płk. rez. dr. hab. inż. Piotra Deli

Kalisz 2023

Spis treści

STRESZCZENIE

Rozprawa doktorska pod tytułem „Znaczenie wywiadu opartego na otwartych źródłach (OSINT) w zapewnieniu bezpieczeństwa systemów teleinformatycznych i bezpieczeństwa osobowego" w zamyśle autora dotyczyła identyfikacji i oceny technik wykorzystywanych w ramach wywiadu otwartoźródłowego w celu zdobycia informacji o budowie i działaniu infrastruktury teleinformatycznej oraz informacji dotyczących sfery osobistej i biznesowej.

W ramach dysertacji przeprowadzona została analiza możliwości wykorzystania OSINT-u do ochrony infrastruktury teleinformatycznej oraz zapewnienia bezpieczeństwa osobowego i biznesowego, poprzez wyprzedzające rozpoznanie i wdrożenie zabezpieczeń przed stosowaniem wywiadu otwartoźródłowego jako jednej z metod i faz ataku na bezpieczeństwo danej organizacji bądź osoby.

Głównym przyczynkiem do napisania niniejszej dysertacji był bardzo ograniczony zasób aktualnej literatury naukowej w języku polskim w temacie samej istoty wywiadu otwartoźródłowego, jego podstaw, technik, wykorzystywanych narzędzi oraz, przede wszystkim, metod obrony przed OSINT-em, realizowanym przez osoby lub organizacje, które mogą zagrozić bezpieczeństwu osobistemu i biznesowemu.

Główny problem badawczy dysertacji określono odpowiednio do przedmiotu oraz celu badań pytaniem: *W jakim zakresie dostępne narzędzia, służące do gromadzenia informacji w ramach wywiadu otwartoźródłowego oraz techniki ich analizy wpływają na bezpieczeństwo systemów teleinformatycznych oraz bezpieczeństwo osobowe, a także jakie są możliwości obrony przed zidentyfikowanymi technikami?*

Dysertacja, nie licząc wstępu i zakończenia, składa się z pięciu merytorycznych rozdziałów: Rozdział 1 Założenia badawcze; Rozdział 2 Charakterystyka metod pozyskiwania informacji z otwartych źródeł; Rozdział 3 Analiza informacji pozyskanych z zasobów internetowych jako fundament wywiadu; Rozdział 4 Wykorzystanie wywiadu opartego na otwartych źródłach w zakresie bezpieczeństwa systemów teleinformatycznych oraz bezpieczeństwa osobowego i biznesowego; Rozdział 5 Możliwe do wprowadzenia zalecenia bezpieczeństwa w zakresie przeciwdziałania wywiadowi otwartoźródłowemu.

W rozdziale pierwszym szczegółowo omówiono przedmiot i cel badań, określono problemy oraz hipotezy badawcze, a także opisano procedury, metody, techniki oraz narzędzia wykorzystywane w ramach badań.

W drugim rozdziale scharakteryzowano metody pozyskiwania informacji z otwartych źródeł, nakreślono zakres i podział rozpoznania otwartoźródłowego, a także omówiono wybrane techniki działań OSINT-owych.

W trzecim rozdziale przedstawiono charakterystykę cyklu wywiadowczego oraz omówiono metody analizy zebranych informacji, a także scharakteryzowano błędy

poznawcze oraz wzorce w ludzkim rozumowaniu, które mogą mieć negatywny wpływ na prowadzenie poprawnego wywiadu otwartoźródłowego.

W rozdziale czwartym zidentyfikowano możliwe zagrożenia, wynikające z prowadzenia wywiadu otwartoźródłowego przeciw osobom lub organizacjom. W tym rozdziale przedstawione zostały także możliwości wykorzystania OSINT-u dla zabezpieczenia sfery osobistej, biznesowej i operacyjnej.

W piątym rozdziale opisano możliwe do wprowadzenia zalecenia bezpieczeństwa w zakresie przeciwdziałania mogącym stanowić zagrożenie technikom wywiadu otwartoźródłowego, w oparciu o istniejące normy, standardy oraz inne wytyczne, w tym zasady OPSEC i PERSEC, a także przedstawiono zasady zabezpieczeń odnoszące się do kluczowych osób w organizacjach.

Rozprawa doktorska stanowi szeroki zbiór treści, odnoszących się do technik wywiadu otwartoźródłowego oraz sposobów zabezpieczeń na nich bazujących. Rzetelna analiza literatury przedmiotu umożliwiła stworzenie pracy o charakterze zwartym i nowatorskim, wyczerpującym główne zagadnienie badawcze.

SUMMARY

The doctoral dissertation entitled “The role of open-source intelligence (OSINT) in ensuring the security of ICT systems and personal security” in the author’s intention concerned the identification and evaluation of the techniques used in the scope of open-source intelligence in order to gather the information about the construction and operation of an IT infrastructure as well as the information from a personal and business domain.

Within this dissertation, an analysis was performed to examine the possibility of using OSINT to protect IT infrastructure and to ensure both personal and business security, by an advance reconnaissance and the implementation of security measures protecting against using open-source intelligence as one of the methods and phases of an attack on organisational or personal safety.

The main reason for writing this Ph.D. thesis was the extremely limited current Polish academic literature on the topic of the matter of open-source intelligence itself, its basics, techniques, tools used, and most of all the methods for defending against the use of OSINT, performed by individuals or organisations, that may pose a threat to personal and business security.

The main research problem of the dissertation, according to the subject and the purpose of the research, is determined by the question: *In what extent the available tools for gathering information during the open-source intelligence process and the techniques for analysing this information influence the security of ICT systems and personal security, and what are the abilities to defend against the identified techniques?*

The dissertation, excluding the introduction and conclusion, consists of five substantive chapters: Chapter 1 Methodological assumptions; Chapter 2 The characteristics of open-source information gathering methods; Chapter 3 The analysis of information acquired from the Internet as the foundation of intelligence; Chapter 4 Using open-source intelligence in the scope of ICT systems security as well as personal and business security; Chapter 5 Security recommendations possible to use against open-source intelligence.

In the first chapter the subject and purpose of the dissertation, research problems and work hypotheses were discussed in detail. Research procedures, methods, techniques and tools were described.

In the second chapter methods for acquiring information from open sources were characterised, the scope and division of open-source intelligence were presented, and selected OSINT techniques were discussed.

The third chapter presents the characteristics of the intelligence cycle, the methods for the analysis of gathered information as well as it describes human reasoning patterns and cognitive biases, which may have a negative impact on performing proper open-source intelligence.

In the fourth chapter the threats were identified, which may result from open-source intelligence performed against individuals or whole organisations. Also, the possibilities of using OSINT for personal, business, and operational security were presented.

The fifth chapter describes the security measures, based on existing standards and other guidelines, including the OPSEC and PERSEC rules, possible to implement to defend against open-source intelligence techniques. Also, this chapter presents the security of the high value individuals in organisations.

The doctoral dissertation is a broad collection of content, related to open-source intelligence techniques and security measures basing on them. A thorough analysis of the subject literature enabled creating a compact and innovative work, that exhausts the main research problem.

WSTĘP

Tematyka OSINT-u, czyli wywiadu otwartoźródłowego (od ang. *Open-Source Intelligence*) była obecna w działaniach militarnych i wywiadowczych od zawsze, jednak dopiero rozwój możliwości rozpoznania, w tym Internetu sprawił, że dzisiaj wykorzystywana jest ona bardzo szeroko, także w obszarze biznesu i bezpieczeństwa osobistego. Rozwój usług internetowych, wkraczających praktycznie w każdą dziedzinę życia ludzkiego sprawił, że temat wyszukiwania i analizy ogólnodostępnych informacji stał się przedmiotem zainteresowania wielu instytucji i osób, niezależnie od ich profilu czy legalności celów działania. Dodatkowo, przykłady śledztw, głównie dziennikarskich, związanych z konfliktami czy łamaniem praw człowieka, spopularyzowały termin „OSINT". Jednak jak pisze Michael Bazzell, w swojej książce „Open Source Intelligence Techniques. Resources for Searching and Analyzing Online Information": „[...] OSINT może mieć różne znaczenie dla różnych osób. Oficjalnie jest zdefiniowany jako rozpoznanie, wypracowane na podstawie publicznie dostępnych informacji, które są zbierane, przetwarzane i rozpowszechniane w określonym czasie dla określonego odbiorcy na potrzebę odniesienia się do konkretnego wymagania wywiadowczego. Dla CIA może to oznaczać informacje pozyskane z zagranicznych przekazów medialnych. Dla prawnika może to znaczyć dane, pozyskane z oficjalnych dokumentów rządowych, które są publicznie dostępne. Dla większości osób są to powszechnie dostępne materiały, pozyskane z Internetu."[1]

Częstym błędem w określaniu czym właściwie jest OSINT, jest jego spłycanie wyłącznie do wyszukiwania danych w Internecie, często z ograniczeniem do wyszukiwarki Google. Równie ważna jest archiwizacja danych w celu ich późniejszej odpowiedniej analizy i zaraportowania, gdyż dopiero informacja, której nadano kontekst może być uznawana za dane wywiadowcze. Jak wskazuje Bazzell „znalezienie dostępnych za darmo informacji w Internecie nie stanowi ostatniego kroku analizy OSINT-owej. [...] Niezależnie czy dane są pozyskiwane w ramach śledztwa, weryfikacji danych czy identyfikacji pracowników zagrażających organizacji, konieczne jest udokumentowanie wszystkich znalezisk. Nie można polegać jedynie na informacji w sieci, która może w pewnym momencie stamtąd zniknąć. Strona internetowa może

[1] M. Bazzell, *Open Source Intelligence Techniques. Resources for Searching and Analyzing Online Information*, Eighth Edition, 2021, s.5.

zostać zdjęta lub informacje mogą zostać usunięte. Dlatego należy zachowywać wszystkie wartościowe informacje po ich wyszukaniu”[2].

Problem odpowiedniej analizy pozyskanych już informacji stanowi wyzwanie dla analityków, którzy powinni być odpowiednio wykształceni nie tylko w zakresie merytorycznym, ale także w zakresie błędów poznawczych i niedoskonałości ludzkiego rozumowania, które może mieć znaczący wpływ na wyniki uzyskane w ramach pracy analitycznej. Dlatego też tak ważne jest zaadresowanie tych aspektów w ramach prowadzenia OSINT-u.

W środowisku Sojuszu Północnoatlantyckiego istnieje rozróżnienie określeń, odpowiadających angielskiemu sformułowaniu „*intelligence*”, tj. „wywiad” i „rozpoznanie”. Pierwsze z nich stosowane jest dla działań na szczeblu taktycznym i taktyczno-operacyjnym, podczas gdy drugie oznacza działania na poziomie resortowym i rządowym[3]. Dla potrzeb niniejszej pracy, która obejmuje w ogromnej większości zakres działań cywilnych, niepodlegających nomenklaturze wojskowej, przyjęto zamienność powyższych terminów jako oznaczających działania, mające na celu zdobywanie informacji. W przypadku odniesień do wywiadu jako informacji przekazywanej przez służby, użyte zostało określenie „informacje wywiadowcze”, natomiast wywiad jako instytucja zdobywająca i analizująca informacje został przedstawiony jako „instytucja wywiadowcza”.

[2] Tamże, s. 5-6.
[3] J. Stróżyk, *Wybrane problemy międzynarodowej współpracy wywiadowczej. Czy NATO ma wywiad?*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa, 2020, s. 17.

ROZDZIAŁ 1

ZAŁOŻENIA BADAWCZE

1.1. Uzasadnienie wyboru tematu

Wywiad oparty na otwartych i ogólnodostępnych źródłach, określany skrótowo jako OSINT, swoją genezę czerpie z działalności takich służb jak brytyjskie BBC Monitoring Service czy amerykańska agencja Foreign Broadcast Monitoring Service (FBMS).

BBC Monitoring Service została utworzona w sierpniu 1939 roku, a więc krótko przed wybuchem II Wojny Światowej, podczas której zajmowała się głównie zbieraniem i jak najszybszą analizą informacji o wydarzeniach z całego świata. Podczas, gdy pierwsze podejścia do monitorowania przekazów radiowych, prowadzone przez BBC pod koniec lat 30. XX wieku miały na celu głównie dostosowanie własnych programów do treści nadawanych przez inne państwowe rozgłośnie, wybuch wojny zmienił kierunek tych działań i zorientował je na nasłuch propagandowych przekazów, prowadzonych przez takie kraje jak Niemcy czy Włochy. John Beresford, odpowiedzialny za procesy planistyczne w Ministerstwie Informacji, wskazywał we wrześniu 1938 roku, że radiowe przekazy propagandowe tych krajów powinny być ciągle monitorowane, gdyż jak wynikało z wniosków wyciągniętych z przebiegu I Wojny Światowej, te same mechanizmy były używane również tym razem i wskazywały na możliwość doprowadzenia do kolejnej wojny[4]. Jak wynika z dalszego przebiegu zdarzeń w Europie, podejrzenia te były jak najbardziej podstawne.

Pierwszych pracowników BBC Monitoring Service rekrutowano między innymi spośród aktualnej kadry BBC, legitymującej się znajomością języków obcych. W celu otrzymania angażu, musieli wysłuchać nagrań zarówno o doskonałej, jak i o słabej jakości i napisać na jego podstawie podsumowanie w określonym czasie. Języki, w których specjalizowały się nowoprzyjęte osoby, to: niemiecki, włoski, hiszpański, portugalski, francuski i arabski[5].

[4] L. M. Johnson, *Establishing Broadcast Monitoring as Open Source Intelligence: The BBC Monitoring Service during the Second World War*, King's College London, 2013, s. 79.
[5] Tamże.

Wartym zaznaczenia jest fakt zaangażowania przez BBC Monitoring (już w 1942 roku) wielu osób do wspólnego wysiłku na rzecz lepszego zobrazowania obszarów Europy pogrążonych w wojnie – w szczególności wybrzeży kontynentu od Norwegii, aż po Pireneje. Po apelu przekazanym przez BBC, słuchacze brytyjscy przekazali około 80 tysięcy fotografii z różnych zakątków Europy, tworząc pierwszą tego typu bibliotekę obrazową[6]. Umożliwiło to pracownikom RAF stworzenie trójwymiarowych modeli terenu oraz późniejsze dobranie odpowiednich miejsc desantu[7].

Po zakończeniu II Wojny Światowej BBC Monitoring dalej rozwijało się i prowadziło swoje działania nasłuchowe, mające na celu monitorowanie i analizę przekazów (już nie tylko radiowych, ale także telewizyjnych, prasowych, internetowych i pochodzących z agencji prasowych), między innymi podczas Zimnej Wojny i upadku Związku Radzieckiego, a także w trakcie wojny w Jugosławii i konfliktów na Bliskim Wschodzie.

Z kolei amerykańska agencja FBMS została zatwierdzona do działania przez administrację prezydenta Franklina D. Roosevelta w lutym 1941 roku. Podczas II wojny światowej agencja FBMS zajmowała się monitorowaniem informacji nadawanych przez zagraniczne źródła, w celu oceny ewentualnych zagrożeń ze strony przeciwnika. Dzień przed atakiem na Pearl Harbor, 6 grudnia 1941 roku, FBMS wydała swój pierwszy raport analityczny, w którym ostrzegała przed wzrostem niebezpiecznych sygnałów dochodzących ze strony japońskiej. Dopiero wtedy znaczenie agencji zostało zauważone, a jej nazwa została zmieniona na Foreign Broadcast Intelligence Service[8]. Jednym z przykładów, świadczących o możliwościach płynących z analizy danych pochodzących z ogólnodostępnych źródeł informacji, jest korelacja pomiędzy cenami pomarańczy w Paryżu, a skutecznością ataków bombowych na mosty kolejowe[9]. To pokazuje, że nawet z pozoru nieistotne informacje, dostępne dla każdego obywatela, mogą stanowić element wywiadu.

Wraz z końcem II wojny światowej i rozformowaniem amerykańskiego Biura Służb Strategicznych (ang. *Office of Strategic Services* – OSS), FBIS została w roku 1946

[6] H. Dylan, *Defence Intelligence and the Cold War: Britain's Joint Intelligence Bureau 1945-1964*, OUP Oxford, 2014, s. 12.
[7] *D-Day 75: How was the biggest ever seaborne invasion launched?*, https://www.bbc.co.uk/teach/d-day-how-was-the-biggest-ever-seaborne-invasion-launched/zrrs7nb – dostęp online 22.03.2022 r.
[8] J. E. Roop, *Foreign Broadcast Information Service. History. Part I: 1941-1947,* CIA, Waszyngton 1969.
[9] C. J. Schneider, D. Schneider, *World War II*, Infobase Publishing, Nowy Jork, 2014, s. 144.

przemianowana na Foreign Broadcast Information Service i stała się częścią CIA. Od tego momentu zajmowała się także monitorowaniem agencji prasowych, a od 1967 roku rozszerzono to także o zagraniczne massmedia – radio, telewizję i prasę.

W 2006 roku Departament Obrony USA sformalizował definicję OSINT-u jako „wywiadu opartego na publicznie dostępnych informacjach, które są zbierane, wykorzystywane i rozpowszechniane w określonym wymiarze czasowym odpowiednim odbiorcom, w celu uzyskania konkretnych potrzeb wywiadowczych”[10]. W glosariuszu NATO definicja OSINT-u przedstawiona została jako „dane wywiadowcze pochodzące zarówno z publicznie dostępnych informacji, jak i innych jawnych informacji o ograniczonym rozpowszechnianiu lub dostępie”[11].

Po atakach z 9 września 2001 roku wyartykułowana została w raporcie[12] komisji powołanej na skutek tych wydarzeń, potrzeba utworzenia osobnej agencji – Open Source Agency w ramach struktur CIA. To poskutkowało utworzeniem w 2005 roku komórki Open Source Center (OSC), która za zadanie miała analizę informacji pochodzących z Internetu, baz danych, prasy, radia, telewizji, filmu, danych geoprzestrzennych, zdjęć oraz obrazów dostępnych komercyjnie, a także podnoszenie kwalifikacji kadry w celu lepszego wykorzystania tych informacji. FBIS została włączona w szeregi nowopowstałej komórki CIA, która od roku 2015 nosi nazwę Open Source Enterprise (OSE) i podlega pod Dyrektoriat Innowacji Cyfrowych.

W dzisiejszym świecie OSINT wykorzystywany jest nie tylko w ramach działań wojskowych, choć w tym obszarze, ze względu na swoją genezę jest on niezmiernie ważnym źródłem informacji. Można zaryzykować stwierdzenie, że obecnie przewaga informacyjna jest równie ważna na arenie międzynarodowej, jak przewaga w zakresie broni konwencjonalnej. Poza zastosowaniem wojskowym, OSINT odgrywa jednak także ogromną rolę w działaniach organów ścigania, detektywów, dziennikarzy (w tym dziennikarzy śledczych), organizacji rządowych i pozarządowych oraz w szeroko pojętym biznesie. W ramach tego ostatniego obszaru OSINT daje nie tylko możliwości analizy rynku i działań konkurencji, ale także sposoby na zabezpieczenie swojej infrastruktury w przypadku wykrycia możliwych wektorów ataków (jak chociażby

[10] *National Defense Authorization Act for Fiscal Year 2006*, https://www.govinfo.gov/content/pkg/PLAW-109publ163/html/PLAW-109publ163.htm – dostęp online 02.03.2022 r. Tłumaczenie własne.
[11] AAP-6 (2019) PL: Słownik terminów i definicji NATO.
[12] *The 9/11 Commission Report*, https://www.9-11commission.gov/report/911Report.pdf – dostęp online 02.03.2022 r.

monitorowanie obecności adresów email w bazach danych z wycieków lub weryfikacja fizycznej możliwości ataku na budynki firmy w określonej lokalizacji).

Temat niniejszej pracy został przez autora wybrany z następujących powodów:

- pomimo ogromnego znaczenia, jakiego nabrał przez lata funkcjonowania w strukturach wojskowych wywiad otwartoźródłowy, opracowania tego zagadnienia w odniesieniu do środowiska cywilnego pozostają bardzo nieliczne, a wiele technik opiera się na wiedzy bardzo wąskiej grupy ekspertów w tej dziedzinie;
- perspektywa lepszego wykorzystania technik i narzędzi wywiadu otwartoźródłowego dla celów cywilnych bazuje na możliwości wykorzystania bardzo rozległego potencjału jednoczesnej pracy wielu specjalistów z różnych dziedzin, co pokazuje potrzebę dla lepszego usystematyzowania wiedzy w tym zakresie, w celu jej dalszego przekazywania;
- autor jako trener cyberbezpieczeństwa w zakresie technik wywiadu otwartoźródłowego i bezpieczeństwa operacyjnego, a także uczestnik międzynarodowych kursów w tym zakresie, zauważa potrzebę rozwijania wiedzy dotyczącej tematu niniejszej dysertacji w ramach dostępnych źródeł wiedzy w języku polskim.

Zaproponowana tematyka niniejszej rozprawy doktorskiej, dotycząca znaczenia wywiadu opartego na otwartych źródłach w zapewnieniu bezpieczeństwa systemów teleinformatycznych i bezpieczeństwa osobowego, będzie pierwszą pracą naukową z tego obszaru bezpieczeństwa.

1.2. Przedmiot badań

Jako przedmiot badań w niniejszej pracy przyjęto zabezpieczenia systemów teleinformatycznych, podłączonych do Internetu oraz procedury bezpieczeństwa osobowego w zakresie ochrony informacji o aktualnym położeniu i statusie osób, a także w odniesieniu do informacji technologicznych z nimi związanych.

1.3. Cel badań

Głównym celem poznawczym badań uczyniono identyfikację i ustalenie możliwości uzyskania szczegółowych informacji, wynikających z przeprowadzanego wywiadu otwartoźródłowego, w odniesieniu do systemów teleinformatycznych oraz indywidualnych osób, a także zidentyfikowanie zagrożeń wynikających z tego typu działań oraz metod skutecznej obrony przed przedmiotowym rozpoznaniem.

W sensie pragmatycznym jako cel badań określono opracowanie **koncepcji identyfikacji i weryfikacji** dostępnego zbioru informacji zawierających szczegóły techniczne, osobowe oraz geolokalizacyjne, dotyczące systemów teleinformatycznych oraz osób, a także **określenie możliwości wprowadzenia zabezpieczeń** przed działaniem zidentyfikowanych technik.

1.4. Problemy badawcze

Odpowiednio do zidentyfikowanego przedmiotu oraz celu badań, główny problem badawczy określono pytaniem: ***w jakim zakresie dostępne narzędzia, służące do gromadzenia informacji w ramach wywiadu otwartoźródłowego oraz techniki ich analizy wpływają na bezpieczeństwo systemów teleinformatycznych oraz bezpieczeństwo osobowe, a także jakie są możliwości obrony przed zidentyfikowanymi technikami?***

W celu rozwiązania określonego wyżej problemu badawczego, konieczne jest rozwiązanie kwestii, opisanych w następujących problemach szczegółowych:

1. *Jakie narzędzia i techniki wywiadu otwartoźródłowego są dostępne dla użytkowników Internetu?*
2. *W jaki sposób należy poddawać analizie zebrane informacje, aby uniknąć ich błędnej interpretacji?*
3. *Jakie zagrożenia płyną z powszechnej możliwości stosowania wywiadu otwartoźródłowego oraz nieprawidłowej analizy danych pozyskanych w ramach przedmiotowych działań w Internecie?*
4. *Jakie są możliwości zabezpieczenia infrastruktury teleinformatycznej oraz zapewnienia bezpieczeństwa osobowego przed działaniami wynikającymi z prowadzonego wywiadu otwartoźródłowego?*

1.5. Hipotezy badawcze

Zidentyfikowany cel oraz problemy badawcze, określone na podstawie aktualnie posiadanego stanu wiedzy, wynikającej z badań wstępnych, a także z analizy literatury, dały możliwość sformułowania następującej głównej hipotezy roboczej oraz wstępnych hipotez roboczych.

W odniesieniu do głównego problemu badawczego o treści: ***w jakim zakresie dostępne narzędzia, służące do gromadzenia informacji w ramach wywiadu otwartoźródłowego oraz techniki ich analizy wpływają na bezpieczeństwo systemów teleinformatycznych oraz bezpieczeństwo osobowe, a także jakie są możliwości obrony przed zidentyfikowanymi technikami?***

Zakładam, że w związku z coraz szerszym wachlarzem narzędzi i usług, dostarczających szeroki zakres danych w Internecie oraz poprzez coraz powszechniejszy dostęp do Internetu i znajomości sposobów na wyszukiwanie w nim treści, a także ze względu na fakt, że praktycznie wszystkie aspekty życia osobistego i zawodowego mają swoje odzwierciedlenie w systemach operujących w chmurze, istnieje zwiększające się zagrożenie zarówno dla bezpieczeństwa systemów teleinformatycznych, które te dane przetwarzają, jak i bezpieczeństwa osobowego, które jest bezpośrednio związane z kwestią poufności i integralności przetwarzanych danych. Możliwości i umiejętności użytkowników Internetu dają im sposobność na sprawdzenie jakie dane mogą zdobyć bez narażania się na bezpośrednie niebezpieczeństwo związane z infiltracją źródeł danych.

W odniesieniu do szczegółowego problemu badawczego o treści: ***jakie narzędzia i techniki wywiadu otwartoźródłowego są dostępne dla użytkowników Internetu***?

Przypuszczam, że ewolucja wyszukiwarek internetowych, zarówno umożliwiających przeglądanie zindeksowanej części Internetu, jak i wyszukiwarek kontekstowych i branżowych, operujących w wąskim zakresie niezindeksowanych danych, umożliwia dotarcie do zakresu danych praktycznie w każdym obszarze informacyjnym. Narzędzia, umożliwiające dostęp do danych graficznych, jak mapy drogowe i satelitarne, obrazy z kamer i zdjęcia opatrzone informacją o ich geolokalizacji, dają możliwość weryfikacji zdarzeń praktycznie w każdym miejscu na Ziemi. Należy także sądzić, że liczba agregatorów danych i materiałów szkoleniowych, dotyczących wywiadu otwartoźródłowego umożliwia bardzo prosty dostęp do całego portfolio

narzędzi i wiedzy, które jeszcze do niedawna znane były jedynie osobom, zajmującym się profesjonalnie przedmiotowymi tematami.

W odniesieniu do szczegółowego problemu badawczego o treści: ***w jaki sposób należy poddawać analizie zebrane informacje, aby uniknąć ich błędnej interpretacji?***

Zakładam, że poprzez niedoskonałość psychiki ludzkiej oraz wielu aspektów wpływających na możliwość zupełnie bezstronnej i nieograniczonej oceny zbieranych w procesie wywiadu otwartoźródłowego danych, wiele procesów jest zaburzonych przez ograniczenia związane z próbą uzyskania pożądanych efektów, a także z uproszczeniami i uogólnieniami tworzonymi podświadomie przez umysł osoby zajmującej się analizą zebranych danych. Należy sądzić, że uświadomienie analityków w zakresie sposobów nieprawidłowej i spolaryzowanej analizy danych jest w stanie uchronić te osoby przed popełnieniem błędów w zakresie błędnego procesu wnioskowania i uzyskiwania mylnych wyników.

W odniesieniu do szczegółowego problemu badawczego o treści: ***jakie zagrożenia płyną z powszechnej możliwości stosowania wywiadu otwartoźródłowego oraz nieprawidłowej analizy danych pozyskanych w ramach przedmiotowych działań w Internecie?***

Wykorzystanie ogólnodostępnych narzędzi i technik wywiadu otwartoźródłowego daje każdej osobie możliwość dotarcia do szerokiego zakresu informacji, bez wstępnej weryfikacji czy profil danej osoby jest odpowiedni do uzyskania dostępu do określonego zbioru danych i ich późniejszego wykorzystania. Zakładam, że brak przygotowania w zakresie poprawnej analizy danych skutkuje wyciąganiem błędnych i często spolaryzowanych wstępnie wniosków, co przekłada się na późniejszą możliwość wysuwania nieprawidłowych oskarżeń i tworzenia fałszywego obrazu sytuacji (umyślnie bądź nieumyślnie). Efekt wytworzenia sensacyjnego wyniku analizy danych może spowodować spolaryzowanie większej ilości użytkowników Internetu, co z kolei może prowadzić do efektu kuli śnieżnej, polegającego na bazowaniu kolejnych osób na pierwotnie nieprawidłowo przetworzonych informacjach.

W odniesieniu do szczegółowego problemu badawczego o treści: ***jakie są możliwości zabezpieczenia infrastruktury teleinformatycznej oraz zapewnienia bezpieczeństwa osobowego przed działaniami wynikającymi z prowadzonego wywiadu otwartoźródłowego?***

Przypuszczam, że wprowadzenie zasad, wynikających z wcześniejszej dogłębnej analizy możliwych do uzyskania danych w ramach prowadzonego wywiadu otwartoźródłowego, jest w stanie zmniejszyć ekspozycję systemów teleinformatycznych na zagrożenia płynące ze zbyt otwartej polityki w zakresie udostępniania danych oraz niedoskonałości oprogramowania, podatnego na techniki pozyskiwania danych pozornie ukrytych i niedostępnych dla przeciętnego użytkownika, a możliwych do uzyskania za pomocą wiedzy specjalistycznej.

Należy także sądzić, że procedury bezpieczeństwa, wynikające z dokumentów normatywnych, takich jak rodzina norm ISO 27000, a także zasady wynikające z przeszkolenia w zakresie bezpieczeństwa prowadzenia działań (ang. *Operations Security* – OPSEC) oraz bezpieczeństwa osobowego (ang. *Personal Security* – PERSEC), dają możliwości przeciwdziałania technikom wywiadu otwartoźródłowego i pozyskiwania danych, mogących mieć wpływ na bezpieczeństwo osób. Istotnym jest zatem określenie efektywnych metod ochrony przed zidentyfikowanymi technikami, które będą możliwe do wdrożenia w systemach teleinformatycznych oraz procedurach bezpieczeństwa osobowego.

1.6. Procedura badawcza

W ramach weryfikacji zidentyfikowanego głównego problemu badawczego, a także problemów szczegółowych oraz weryfikacji hipotez roboczych, dokonano podziału całości przedmiotowego procesu na następujące etapy:

- przygotowanie badań,
- realizacja procesu badawczego,
- interpretacja uzyskanych wniosków,
- sporządzenie dysertacji.

Pierwszy etap procesu badawczego – przygotowania badań – składał się z działań mających na celu zidentyfikowanie i zapoznanie się z dostępnymi technikami wywiadu otwartoźródłowego oraz uporządkowania dostępnych materiałów i narzędzi dotyczących tej tematyki. Działania te obejmowały ponadto określenie zagrożeń wynikających z powszechnego dostępu do technik i narzędzi OSINT-owych, a także ogólnodostępnej

wiedzy w tym zakresie. W tym etapie przeprowadzono analizę dostępnej literatury, włączając w to źródła występujące jedynie w postaci zasobów internetowych, dokumenty normatywne dotyczące bezpieczeństwa informacji i bezpieczeństwa prowadzenia operacji. Zakończeniem pierwszego etapu było przygotowanie koncepcji rozprawy doktorskiej.

W drugim etapie prac wykonano badania właściwe, w ramach których wykorzystywano zarówno teoretyczne, jak i empiryczne metody badawcze w zakresie opisywanej problematyki. Celem tego etapu była weryfikacja lub falsyfikacja przyjętych hipotez.

Wykorzystywane metody teoretyczne skupiały się głównie na analizie dostępnej wiedzy w zakresie wywiadu otwartoźródłowego, w tym publikacji książkowych i internetowych, instrukcji bezpieczeństwa prowadzenia operacji, a także informacji z konferencji naukowych.

W ramach tego etapu ważnym elementem było wykonanie badań empirycznych, które polegały na przeprowadzeniu wywiadu eksperckiego z osobami, które na co dzień wykorzystują w swojej pracy techniki wywiadu otwartoźródłowego i mają w tym zakresie szeroką wiedzę na poziomie eksperckim. W tym celu przygotowano kwestionariusz wywiadu eksperckiego, w którym zawarto pytania związane z szerokim spektrum kwestii będących w zakresie tematycznym niniejszej pracy. Otrzymane odpowiedzi następnie posłużyły do weryfikacji postawionych hipotez.

Trzeci etap polegał na interpretacji uzyskanych wcześniej wniosków poprzez analizę informacji zebranych w ramach wcześniejszych etapów i zestawienie ich z otrzymanymi odpowiedziami w ramach przeprowadzonych wywiadów eksperckich.

Czwartym i ostatnim etapem prac było sporządzenie niniejszej dysertacji, zawierającej w sobie opis dostępnych technik i narzędzi służących do prowadzenia wywiadu otwartoźródłowego, wyzwań i zagrożeń związanych z kwestią analizy zebranych danych, a także zaproponowanie możliwych do wdrożenia zabezpieczeń, uniemożliwiających uzyskanie informacji, dotyczących systemów teleinformatycznych i osób, z wykorzystaniem technik OSINT-owych.

1.7. Metody, techniki i narzędzia badawcze

„Nauki o bezpieczeństwie charakteryzują się eklektyzmem metodologicznym – nie dopracowały się własnej metodologii, więc wykorzystują dorobek innych dyscyplin naukowych. Metodologia w naukach społecznych to zestaw dyrektyw badawczych, opartych na przyjętych założeniach teoretycznych, a także sposoby formułowania, uzasadniania i sprawdzania twierdzeń. Dotyczy to m.in. takich zabiegów poznawczych jak stawianie i rozwiązywanie problemów, formułowanie zjawisk, ich opis i wyjaśnianie, prowadzenie badań, uzasadnianie twierdzeń oraz sposobów wnioskowania. W ramach zabiegów poznawczych w naukach o bezpieczeństwie następuje sformułowanie pytań badawczych, hipotez i twierdzeń o uwarunkowaniach procesu bezpieczeństwa, wypracowanie narzędzi badawczych, budowa modeli i teorii oraz tworzenie systemu organizacji badań naukowych, ich prowadzenia i kontroli. Narzędzia badawcze odnoszą się do rozpoznania zagrożeń i przeciwdziałania im, ale także (wraz z wypracowanymi metodami) – do systematycznego kumulowania wiedzy i rozwiązań praktycznych dotyczących bezpieczeństwa. Nauki o bezpieczeństwie jako dyscyplina naukowa w ramach nauk społecznych wykorzystują metody badań empirycznych. W przypadku trudności w sformułowaniu hipotez badawczych, w ich miejsce formułuje się pytania badawcze. Doświadczalną weryfikację hipotez badawczych cechuje postępowanie empiryczne, wykorzystujące najczęściej logikę indukcji, dopuszczalne jest wykorzystywanie logiki formalnej, rzadko – teorii dedukcji. Najbardziej reprezentatywne cechy przedmiotu badań ustala się za pomocą metod i technik badawczych. Rozumowanie indukcyjne zakłada cel w postaci uogólnień. Etapem wstępnym indukcji jest przedstawienie hipotezy. Najogólniejszymi metodami przetwarzania materiału badawczego w naukach społecznych są analiza i synteza, jako składniki szczegółowych metod i technik badawczych. Metoda analizy prowadzi do opracowania metodyki, dla uzyskania informacji o składzie badanej próby.”[13]

Obszar badań, ujęty w niniejszej pracy, wykracza niejako poza obszar nauk społecznych, w ramach którego funkcjonują nauki o bezpieczeństwie, wskazując na transdyscyplinarny charakter. W etapie badań wstępnych zidentyfikowano problem naukowy, a następnie przystąpiono do określenia problemu badawczego oraz

[13] A. Chodyński, *Podnoszenie poziomu bezpieczeństwa. Metody i narzędzia. Wprowadzenie* – Bezpieczeństwo. Teoria i praktyka, 2019, nr 4, s. 13-14.

sformułowania hipotez roboczych. Określono także szczegółową procedurę badawczą dla określonego problemu.

W niniejszej dysertacji zdecydowano się na wybór metody jakościowej prowadzenia badań. „Badania jakościowe powinny przebiegać w świecie rzeczywistym, gdy badane zjawiska zachodzą «naturalnie». Badacz powinien przyjąć postawę otwartości na to, co się wydarzyło, bez kontrolowania i ograniczeń. Powinien unikać sztywnych wzorów, które utrudniają poszukiwanie «nowych ścieżek».”[14] Skala prowadzonych badań była zatem mniejsza, niż miałoby to miejsce w przypadku metody ilościowej, jednak ze względu na specyfikę badanej tematyki oraz dostępność ekspertów, zdecydowano się na metodę jakościową. „Metody jakościowe charakteryzuje indywidualne podejście do respondenta, niewielka skala badań, stosowanie niezestandaryzowanych narzędzi badawczych oraz szczególna rola badacza, który wchodzi z badanym w interakcję. Badacz dokonuje też subiektywnej interpretacji uzyskanych wyników. Najczęściej metodom jakościowym przypisuje się techniki badawcze oparte na wywiadzie (spontaniczny wywiad narracyjny, swobody wywiad nieukierunkowany, wywiad ekspercki, wywiad zogniskowany, wywiad pogłębiony, wywiad grupowy) oraz obserwacji (obserwacja uczestnicząca lub nieuczestnicząca, jawna lub ukryta).”[15]

1.8. Analiza krytyczna literatury z zakresu przedmiotu badań

Z uwagi na dynamicznie zwiększającą się obecność tematyki wywiadu otwartoźródłowego w opracowaniach cywilnych dopiero w kilku ostatnich latach, w szczególności za sprawą rosnącej popularności tego typu działań w odniesieniu do konfliktów zbrojnych i protestów społecznych, niewiele jest opracowań naukowych, obejmujących ten zakres wiedzy. Duża liczba opracowań dotyczy aspektów militarnych, gdzie tematyka wywiadu otwartoźródłowego jest obecna już od połowy XX. wieku, jednak dotyczy to głównie opracowań amerykańskich.

W ramach niniejszej dysertacji wykorzystywane były nie tylko opracowania naukowe, ale także wiele opracowań pozanaukowych, w których zawarte zostały istotne informacje, związane bezpośrednio lub pośrednio z przedmiotem niniejszej pracy.

[14] S. Górski, *O metodach badawczych w naukach społecznych*, Nauki o bezpieczeństwie. Wybrane problemy badań, Wydawnictwo CNBOP-PIB, 2017, s. 62-63.
[15] Tamże, s. 63.

W ramach opracowań pozanaukowych brane pod uwagę były także opracowania, pochodzące ze źródeł internetowych, a będące artykułami niepublikowanymi nigdzie poza wersją internetową. Za każdym razem jednak weryfikowana była przez autora wiarygodność źródła i autora konkretnego opracowania.

Jako podstawę wiedzy w zakresie metod i narzędzi badawczych, wykorzystywanych w ramach pracy wykorzystano opracowania: A. Chodyński, *Podnoszenie poziomu bezpieczeństwa. Metody i narzędzia. Wprowadzenie* w: Bezpieczeństwo. Teoria i praktyka, 2019 oraz A. Czupryński, B. Wiśniewski, J. Zboina (red.), *Nauki o bezpieczeństwie. Wybrane problemy badań*, Wydawnictwo CNBOP-PIB, 2017.

Rozdział drugi zawiera opracowanie w zakresie charakterystyki metod pozyskiwania informacji z otwartych źródeł. W celu opisania tematu wykorzystano źródła: Mark M. Lowenthal, Robert M. Clark, *The Five Disciplines of Intelligence Collection*, CQ Press, 2015, Kamila Matela, *Wybrane aspekty systemów wywiadu, obserwacji i rozpoznania (ISR)*, WIEDZA OBRONNA, 2021, Vol. 276 No. 3, R. K. Hudnall, *No Safe Haven: Homeland Insecurity*, Omega Press, El Paso, Texas, 2004, *NATO OSINT Handbook v.1.2*, 2002, *Intelligence Community Directive Number 301*, National Open Source Enterprise, 2006, *National Defense Authorization Act for Fiscal Year 2006*, 2006, *Berkeley Protocol on Digital Open Source Investigations*, HR/PUB/20/2 (advance version), 2022, K. Król, *Geoinformation in the invisible resources of the Internet*, Geomatics, Landmanagement and Landscape No. 3, 2019, J. Saleem, R. Islam and M. A. Kabir, *The Anonymity of the Dark Web: A Survey*, IEEE Access, vol. 10, 2022, Mollie L. Coffey, *Library application of Deep Web and Dark Web technologies*, School of Information Student Research Journal, 10(1), 2020 oraz H. Bean, *Is Open Source Intelligence an Ethical Issue?*, Research in Social Problems and Public Policy, Volume 19.

Korzystano w omawianym zakresie także z wiedzy, zawartej w opracowaniach nienaukowych: Federation of American Scientists, The Interagency OPSEC Support Staff, *Operations Security, Intelligence Threat Handbook*, 1996, *AAP-6 (2019) PL: Słownik terminów i definicji NATO*, 2019, Federation of American Scientists, *Intellligence Resource Program, Measurement and Signature Intelligence (MASINT)*, C. Formeller, *Introducing 15 cm HD: The Highest Clarity From Commercial Satellite*

Imagery, 2020, *Number of global social network users 2018-2027,* Statista, 2022, *Most popular social networks worldwide as of January 2022, ranked by number of monthly active users (in millions)*, Statista, 2022, *How Much of the Internet is the Dark Web in 2022?*, Techjury, 2022, *Dark Web Searching*, OSINT Combine, 2021, *Search Party Rules of Engagement*, Trace Labs, 2022, *July 2022 Web Server Survey*, NetCraft, 2022, *Internet Live Stats: Total number of Websites*, InternetLiveStats, 2022, *How many active sites are there?*, NetCraft, 2022, *File types indexable by Google*, Google, 2022, Johnny Long*, The Google Hacker's Guide. Understanding and Defending Against the Google Hacker*, 2004, *AltaVista Photo Finder, and how to keep your images "unfound"*, *AltaVista Photo Finder Has Artists Concerned*, 1999, *News Watch; A Quick Way to Search For Images on the Web*, New York Times, 2001, Eric Schmidt*, The Tinkerer's Apprentice*, 2015, *Ooh! Ahh! Google Images presents a nicer way to surf the visual web*, Google, 2010, *Similar Images graduates from Google Labs*, Google, 2009, *Relevance meets the real-time web*, Google, 2009, *Knocking down barriers to knowledge*, Google, 2011, *Report: State of the Web*, httparchive, 2022, *Retired Chicago Firefighter Wrongly Accused of Participating in Capitol Violence*, Firefighter Nation, 2021, *How Open-Source Intelligence Is Helping Clear The Fog Of War In Ukraine*, BuzzFeed News, 2022 oraz *Feeling the Burden. Ethical Challenges and Practices in Open Source Analysis and Journalism*, Stanley Center, 2022.

Korzystano także z opracowań własnych, zawierających się w przedmiotowej tematyce: *Jak wyszukiwarki radzą sobie z analizą zawartości obrazów*, sekurak.pl, 2021.

W rozdziale trzecim przedstawiono kwestie odpowiedniej analizy informacji pozyskanych z zasobów internetowych jako fundamentu wywiadu. Korzystano ze źródeł: Robert M. Clark, *Intelligence Analysis: A Target-Centric Approach*, CQ Press, 2019, *Joint Chiefs of Staff, Joint Publication 2-0: Joint Intelligence*, 2013, Mark M. Lowenthal, *Intelligence: From Secrets to Policy*, wyd. 6, CQ Press, 2014, *Factbook on Intelligence*, Central Intelligence Agency, P. Dela, *Charakterystyka zagrożeń bezpieczeństwa cyberprzestrzeni*, PWN, Warszawa, 2022, D. Rumsfeld, *Known and Unknown: A Memoir*, Penguin, 2011, L. Krizan, *Intelligence essentials for everyone*, Joint Military Intelligence College, Waszyngton, 1999, ATP 2-33.4. *Intelligence Analysis*, Styczeń 2020, P. Dela, *Elementy propagandy w życiu publicznym*, Studia Politologiczne 54, R. J. Heuer Jr., *Psychology of Intelligence Analysis*, Center For The Study Of Intelligence, Central Intelligence Agency, 1999, L. Witlin, *Of Note: Mirror-Imaging and Its Dangers*,

SAIS Review of International Affairs, Johns Hopkins University Press, Volume 28, Number 1, Winter-Spring 2008, A. Tversky; D. Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, Science, New Series, Vol. 185, No. 4157. (27.09.1974) oraz A. Tversky, D. Kahneman, *Availability: A heuristic for judging frequency and probability*, Cognitive Psychology, Volume 5, Issue 2, wrzesień 1973.

W związku z występowaniem zagadnień, dotyczących śledztw cywilnych głównie w źródłach internetowych, posiłkowano się także opracowaniami nienaukowymi, opublikowanych jedynie w Internecie: *Cognitive biases – acaps technical brief*, ACAPS, 2016, N. Dekens, *Vicarious trauma and OSINT – a practical guide* i Benjamin Brown, *Cognitive Bias and Critical Thinking in Open Source Intelligence (OSINT)*, Circle City Con 2014.

W rozdziale czwartym skupiono się na zagadnieniach związanych z możliwością wykorzystania wywiadu opartego na otwartych źródłach w zakresie zapewnienia bezpieczeństwa systemów teleinformatycznych oraz bezpieczeństwa osobowego i biznesowego. Skupiono się tutaj nie tylko na identyfikacji zagrożeń, ale także na możliwych scenariuszach wyprzedzającej obrony.

W ramach tego rozdziału korzystano z amerykańskich wojskowych dokumentów i poradników, dotyczących zapewnienia bezpieczeństwa żołnierzy: *OPSEC - A Guide For Family and Friends,* 1st Information Operations Command (Land) Vulnerability Assessment Detachment Army OPSEC Support Element, 2013, *Operations Security (OPSEC) Guidance for Family Members,* 2nd Marine Division oraz *DoD OPSEC for Families*, OSPA, 2010.

Wykorzystywano także informacje zawarte w źródłach internetowych, ze względu na ich wartość jako źródeł informacji technicznych o zaobserwowanych atakach: *Przez TeamViewer-a aż do systemu kontroli uzdatniania wody. Hacker zmienił parametry chemiczne wody*, sekurak.pl, 2021, *Nastolatek uzyskał zdalny dostęp do ponad 25 Tesli na całym świecie. Lokalizowanie samochodów, otwieranie drzwi, wystartowanie auta…*, sekurak.pl, 2022, *Poszukiwany za przekręt na 50 mln zł złapany, bo jego młoda partnerka chwaliła się życiem w mediach społecznościowych*, AntyRadio, 2020, *Fitness app Strava lights up staff at military bases*, BBC, 2018 oraz *A Lance Corporal's Phone Selfie Got His Marine Unit 'Killed' at 29 Palms*, Military.com, 2020.

Jako bazę informacji o kategoriach zagrożeń wykorzystano bazę *MITRE Common Weakness Enumeration* (CWE).

Rozdział piąty opisuje możliwe do wprowadzenia zalecenia bezpieczeństwa w zakresie przeciwdziałania wywiadowi otwartoźródłowemu. W tym zakresie wykorzystywano źródła: Leszek Wiszniewski, *Rola i znaczenie analizy informacji wywiadowczej w zapewnianiu bezpieczeństwa państwa*, Przegląd Bezpieczeństwa Wewnętrznego, 2020 nr 22 (12) oraz *Purple Dragon. The Origin and Development of the United States OPSEC Program,* NSA, 2007.

Analizując sposoby zwiększenia poziomu bezpieczeństwa biznesowego, korzystano z materiałów z konferencji SANS 2022 OSINT Summit: Christina Lekati – *Protecting High-Value Individuals: An OSINT Workflow*, a także z informacji zawartych w artykule K. Zetter *Palin E-Mail Hacker Says It Was Easy*, Wired, 2008.

W tym rozdziale także korzystano z opisów kategorii zagrożeń dla sprzętu i oprogramowania, zawartych w bazie MITRE CWE, a także wykorzystywano opisy techniczne narzędzi i technik OSINT-owych w zakresie zapewnienia bezpieczeństwa systemów teleinformatycznych: DiSIEM Project Deliverable D4.1 – Techniques and tools for OSINT-based threat analysis.

1.9. Ograniczenia badawcze

Należy zauważyć, że duża część aktualnej wiedzy w zakresie pokrywającym się z tematyką niniejszej dysertacji, niebędąca publikacjami dostępnymi wyłącznie dla osób niezwiązanych z wojskowymi lub administracyjnymi komórkami prowadzącymi prace wywiadowcze, a także publikacjami naukowymi, jest wiedzą mocno rozproszoną, dostępną głównie w Internecie. Z tego względu konieczna jest każdorazowa weryfikacja źródeł informacji w zakresie ich reputacji, rzetelności oraz poprawności zawartych w nich danych, gdyż w przeciwieństwie do opracowań redagowanych, źródła internetowe pozbawione są często procesu weryfikacji poprawności zawartych w nich informacji.

Działania w ramach wywiadu OSINT-owego są stosowane w trakcie prac jednostek policji oraz służb, stąd zdobycie wiedzy z tego obszaru jest tematem nietrywialnym, gdyż często niemożliwe jest uzyskanie odpowiedzi na pytania w zakresie

metodologii prowadzenia rozpoznania przez przedmiotowe instytucje ze względu na charakter ich działania.

W związku z faktem, że charakterystyka rozpoznania OSINT-owego w obecnych czasach związana jest często ze źródłami i narzędziami internetowymi, dynamika ich zmian jest bardzo duża, co wpływa na możliwość dezaktualizacji niektórych z opisywanych w niniejszej pracy narzędzi i źródeł informacji lub znaczną zmianę ich zakresu.

1.10. Wywiad ekspercki

W celu zweryfikowania hipotez, założonych w niniejszej dysertacji, przeprowadzono wywiad ekspercki. Jego zakres oraz szczegółowe pytania i wnioski z analizy odpowiedzi ekspertów, zawarte zostały w Załączniku.

ROZDZIAŁ 2

CHARAKTERYSTYKA METOD POZYSKIWANIA INFORMACJI Z OTWARTYCH ŹRÓDEŁ

Uwagi wstępne

Przenoszenie coraz większej liczby aspektów ludzkiego życia do Internetu zmieniło sposoby prowadzenia śledztw, wykorzystujących dane z otwartych, ogólnodostępnych źródeł. Dotychczasowe przeglądanie doniesień prasowych, transmisji telewizyjnych i radiowych, zastąpione zostało przez monitorowanie mediów społecznościowych i portali dostarczających danych na wybrany temat w Internecie. Udostępnianie przez wiele osób praktycznie każdego aspektu swojego prywatnego i zawodowego życia w sieci umożliwiło z jednej strony lepsze rozeznanie biznesowe w przypadku firm oraz otworzyło całkiem nowy rozdział w zakresie metod poszukiwania osób zaginionych i poszukiwanych przez organy ścigania. Z drugiej jednak strony większa otwartość, jeśli chodzi o udostępnianie publicznie wielu dotychczas bardzo osobistych i poufnych danych otworzyło przed przestępcami i nieuczciwymi partnerami biznesowymi nowe możliwości w zakresie uzyskania lepszego rozeznania w celu późniejszego przeprowadzenia ataku, mającego na celu zdobycie zazwyczaj pieniędzy lub informacji (która de facto jest biznesowo równoważnikiem zysku pieniężnego). Transmisja zdarzeń w działaniach militarnych daje aktualnie możliwość niezależnej weryfikacji danych wywiadowczych, zdobytych przy użyciu klasycznych technik, a także przedstawia ogromne zagrożenie dla osób, które poprzez geolokalizację mogą stać się celem ataków.

Z tego względu pilnym okazało się przemodelowanie także sposobów na zabezpieczenie przed działaniem technik OSINT-owych, głównie w zakresie odpowiedniej konfiguracji środowiska teleinformatycznego, ale także dotyczących sposobów postępowania przez pracowników firm, śledczych (w tym dziennikarzy) oraz zwykłych użytkowników Internetu, którzy chcą zachować swoją prywatność oraz wysoki poziom bezpieczeństwa.

1.11. Otwarte źródła

Wśród sposobów pozyskiwania informacji wymienia się najczęściej pięć rodzajów rozpoznania[16]. Są to:

- HUMINT (ang. *Human Intelligence*) – pozyskiwanie informacji ze źródeł osobowych, najczęściej poprzez typowe działania wywiadowcze, ale także poprzez rozpoznanie dyplomatyczne lub poprzez jawne kontakty z osobami, mogącymi udzielić istotnych informacji. Podczas, gdy poprzez rozwój technologiczny coraz więcej informacji o zasobach rozpoznawanej organizacji lub państwa jest możliwych do zdobycia poprzez inne techniki wywiadowcze, HUMINT stanowi nadal źródło informacji o planowanych działaniach. Słownik terminów i definicji NATO[17] określa HUMINT jako wiadomości rozpoznawcze/wywiadowcze opracowane na podstawie informacji zebranych przez operatorów (rozpoznania osobowego) i zasadniczo dostarczonych przez źródła osobowe.
- SIGINT (ang. *Signals Intelligence*) – termin ogólny określający rozpoznanie radiowe i elektroniczne (wywiad radiowy i elektroniczny), wtedy, gdy nie ma potrzeby rozróżniania tych dwóch typów rozpoznania (wywiadu) lub należy przedstawić ich wspólny odpowiednik[18]. Pozyskiwanie informacji realizowane jest z transmisji elektronicznych, z wykorzystaniem zarówno odbiorników naziemnych, satelitów, jak i jednostek pływających oraz powietrznych. W ramach SIGINT-u można wyróżnić: COMINT (ang. *Communications Intelligence* – przechwytywanie komunikacji głosu, obrazu, połączeń telefonicznych i faksowych, a także np. sygnałów przesyłanych alfabetem Morse'a) oraz ELINT (ang. *Electronic Intelligence* – przechwytywanie informacji emitowanych za pomocą fal elektromagnetycznych, np. sygnałów radarowych, w celu określenia położenia i sposobu działania analizowanej infrastruktury).

[16] M. M. Lowenthal, R. M. Clark, *The Five Disciplines of Intelligence Collection*, CQ Press, 2015, s. xi oraz Federation of American Scientists, The Interagency OPSEC Support Staff, *Operations Security, Intelligence Threat Handbook*, 1996, https://irp.fas.org/nsa/ioss/threat96/index.html – dostęp online 01.09.2022 r.
[17] AAP-6 (2019) PL: Słownik terminów i definicji NATO – https://wcnjik.wp.mil.pl/u/AAP-6_2019_PL.xlsx – dostęp 18.03.2022 r.
[18] Tamże.

- IMINT (ang. *Imagery Intelligence*) – rozpoznanie obrazowe, bazujące zarówno na materiałach fizycznych (jak filmy lub zdjęcia), jak i elektronicznych. Oprócz zdjęć i filmów (PHOTOINT), obejmuje także m.in. obrazy radarowe, w podczerwieni, laserowe i elektrooptyczne. W zakresie rozpoznania analizy informacji geoprzestrzennej, mającej na celu identyfikację i ocenę obiektów naziemnych, funkcjonuje określenie GEOINT (ang. *Geospatial Intelligence*).

- MASINT (ang. *Measurement and Signatures Intelligence*) – obejmujące analizy ilościowe i jakościowe danych pozyskanych z różnych źródeł, w celu określenia cech specyficznych dla źródeł przechwytywanego sygnału. W jego zakres[19] wchodzą takie obszary rozpoznania jak: ACINT (ang. *Accoustic Intelligence* – akustyczne), RINT/URINT[20] (ang. *Unintentional Radiation Intelligence* – przypadkowej emisji ujawniającej), IRINT (ang. *Infrared Intelligence* – podczerwieni), CBINT/CBNINT (ang. *Chemical and Biological Intelligence* – chemiczne i biologiczne), DEWINT (ang. *Directed Energy Weapons Intelligence* – broni wiązkowej), NUCINT (ang. *Nuclear Intelligence* – nuklearne), EMPINT (ang. *Electromagnetic Pulse Intelligence* – impulsu elektromagnetycznego), ELECTRO-OPTINT (ang. *Electro-optical Intelligence* – elektro-optyczne, w tym LASINT – ang. *Laser Intelligence* – laserowe), MATINT[21] (ang. *Materials Intelligence* – analiza materiałowa), a także *Spectroscopic Intelligence* (analiza spektroskopowa) i Effluent/Debris Collection (analiza zanieczyszczeń atmosferycznych).

- OSINT (ang. *Open-Source Intelligence*) – rozpoznanie jawnoźródłowe bazujące na informacjach ogólnodostępnych (np. poprzez obserwację lub żądanie dostępu do nich), zbieranych z wykorzystaniem legalnych źródeł, zarówno darmowych, jak i dostępnych za opłatą. Ważnym elementem

[19] Opracowano na podstawie: K. Matela, *Wybrane aspekty systemów wywiadu, obserwacji i rozpoznania (ISR)*, WIEDZA OBRONNA, 2021, Vol. 276 No. 3, s. 238-253 oraz Federation of American Scientists, *Intellligence Resource Program, Measurement and Signature Intelligence (MASINT)* – https://irp.fas.org/program/masint.htm – dostęp online 01.09.2022 r.

[20] W związku z różnym zapisem niektórych skrótów w analizowanej literaturze, zapisane zostały one w obu występujących formach.

[21] Skrót ten rzadko występuje w opracowaniach, zazwyczaj ten rodzaj rozpoznania określany jest swoją pełną nazwą. Wersja skrótowa za: R. K. Hudnall, *No Safe Haven: Homeland Insecurity*, Omega Press, El Paso, Texas, 2004, s. 87.

działań OSINT-owych jest także odpowiednia weryfikacja danych, które poprzez swoją naturę ogólnej dostępności mogą być elementami przypadkowo podanych błędnych informacji lub celowych działań dezinformacyjnych.

OSINT często nazywany jest także „białym wywiadem", jednak na potrzeby niniejszej dysertacji autor używa tłumaczenia wywiad „otwartoźródłowy" lub „jawnoźródłowy" w celu lepszego oddania idei tego typu działań. Ma to związek z faktem, że podstawą do prowadzenia działań OSINT-owych jest zebranie wiedzy ze źródeł, które muszą być jawne i publicznie dostępne. W kolejnych krokach zebrane informacje są przetwarzane w celu ich odpowiedniej analizy i wytworzenia raportów, zawierających dane, będące odpowiedziami na zadane na początku pytania.

1.11.1. Informacje z jawnych źródeł

Jak wskazuje NATO Open Source Intelligence Handbook[22], same bazowe dane, zbierane w tym procesie, określa się mianem *Open Source Data* (OSD). Mogą to być materiały w swojej podstawowej, nieprzetworzonej jeszcze przez analityków formie: gazet, książek, audycji radiowych i telewizyjnych, wystąpień, a także nagrania audio, fotografie w swojej klasycznej formie lub w postaci zdjęć satelitarnych czy nawet korespondencja pocztowa.

Kolejnym stadium są informacje zebrane na podstawie posiadanych źródeł, czyli OSIF (lub OSINF – skrót od angielskiego terminu *Open Source Information*). Są one wynikiem procesów przetwarzania i edycji zebranych danych, mających na celu ich sprawdzenie i przefiltrowanie oraz niekiedy zmianę formy, w jakiej występują. Przykładem tego typu informacji mogą być gazety, książki, przekazy medialne lub raporty.

Informacje, które zostały zebrane, wyodrębnione i rozpowszechnianie wybranym odbiorcom w celu uzyskania odpowiedzi na konkretne pytanie, są określane mianem OSINT (*Open Source Intelligence*). Wymagają one przeprowadzenia procesów umysłowych na zebranych z otwartych źródeł danych, co jest konieczne, aby można było nazwać ten proces wywiadowczym.

[22] *NATO OSINT Handbook v.1.2* – https://ia800502.us.archive.org/14/items/NATOOSINTHandbookV1.2/NATO%20OSINT%20Handbook%20v1.2%20-%20Jan%202002.pdf – dostęp online 18.03.2022 r.

NATO Open Source Intelligence Handbook wskazuje jeszcze jeden poziom przetwarzania informacji, a mianowicie OSINT-V, czyli zwalidowany OSINT (ang. *Validated OSINT*). Wysoki stopień pewności co do informacji wywiadowczych w ramach tego określenia uzyskiwany jest poprzez zapewnienie odpowiednio wyszkolonej osoby, opracowującej zdobyte informacje, ze wsparciem pochodzącym z niejawnych źródeł wywiadowczych lub poprzez wykorzystanie źródła informacji, co do którego istnieje wysoki poziom pewności co do dostarczanych przez nie informacji (np. obraz samolotu lądującego na lotnisku, który jest transmitowany na żywo przez stacje telewizyjne).

Jako kategorie otwartych źródeł, można wyróżnić:

- tradycyjne formy przekazu medialnego (radio, telewizja, prasa, film),
- komercyjne serwisy informacyjne,
- serwisy internetowe (w tym media społecznościowe, strony internetowe, blogi, agregatory i bazy danych, pliki i inne),
- ekspertów,
- szarą literaturę[23],
- komercyjne źródła zobrazowania,
- serwisy mapowe.

Nieco inne podejście do informacji pochodzących z otwartych źródeł przekazuje Dyrektywa Społeczności Wywiadowczej nr 301, ogłoszona przez Dyrektora Centralnego Wywiadu, która określa je jako „publicznie dostępne informacje, które każdy może zgodnie z prawem pozyskać na żądanie, poprzez zakup lub obserwację”[24]. Mamy tutaj zatem podejście dookreślające źródła pochodzenia informacji jako nienaruszające prawa, co wyklucza wszystkie formy działań hackerskich. To odróżnia OSINT od innych działań wywiadowczych, takich jak HUMINT czy SIGINT, które polegają raczej na informacjach niedostępnych publicznie, zdobywanych poprzez przełamanie bariery niedostępności informacji niejawnych. W tym samym dokumencie OSINT z kolei określany jest jako

[23] Jako szarą literaturę określa się wszelkie publikacje, tworzone poza standardowym systemem wydawniczym. Termin ten obejmuje m.in.: raporty z badań, publikacje pokonferencyjne, dokumenty tworzone przez organizacje rządowe, pozarządowe i biznesowe, blogi i inne publikacje internetowe.

[24] *Intelligence Community Directive Number 301*, National Open Source Enterprise, 2006, https://irp.fas.org/dni/icd/icd-301.pdf – dostęp online 22.03.2022 r. Tłumaczenie własne.

„wytworzony na podstawie publicznie dostępnych informacji, które są zbierane, wykorzystywane i rozpowszechniane w określonym wymiarze czasowym odpowiednim odbiorcom, w celu uzyskania konkretnych potrzeb wywiadowczych”[25], co jest powieleniem definicji zawartej w cytowanym na wstępie niniejszej pracy dokumencie „National Defense Authorization Act for Fiscal Year 2006”[26]. Tego rodzaju określenie wskazuje na dwa aspekty. Pierwszym z nich jest ustalenie odbiorców, dla których przeznaczone są wyniki pracy analitycznej na zebranych wcześniej danych. Drugim aspektem jest obrany na początku cel prowadzonej analizy, czyli konieczność znalezienia odpowiedzi na konkretne zagadnienie, które jest przyczynkiem do wykonania całego procesu rozpoznania otwartoźródłowego.

W ramach tradycyjnych form przekazu nadal możliwe jest pozyskiwanie wartościowych informacji, chociażby poprzez programy nadawane przez stacje telewizyjne i radiowe oraz książki, czasopisma oraz gazety codzienne. Nie wszystkie publikowane w wymienionych mediach treści są dostępne w Internecie, dlatego nie należy pomijać przywołanych klasycznych form przekazu informacji podczas analizy możliwości rozpoznania otwartoźródłowego.

Komercyjne serwisy internetowe, także powinny stanowić możliwe źródło danych. Jednym z przykładów mogą być tutaj firmy dostarczające mapy satelitarne o wysokiej rozdzielczości. Dla przykładu firma Maxar, która zapewnia zdjęcia satelitarne o rozdzielczości do 15 cm[27], udostępnia co prawda część danych za darmo, jednak nie są to materiały o wysokiej jakości i rzadko można je wykorzystać w pełni do przeprowadzenia rozpoznania otwartoźródłowego. Dopiero płatny serwis zapewnia odpowiednią jakość zobrazowania i w takiej formie stanowi on bardzo cenną bazę dla śledczych, weryfikujących między innymi sprawy prowadzonych walk na całym świecie, przestępstw organizowanych na szeroką skalę (często na poziomie rządów państw), zaginionych jednostek pływających i latających lub zorganizowanego przemytu. Istnieje jednak wątpliwość dotycząca zakresu korzystania z płatnych serwisów i narzędzi komercyjnych jako źródeł OSINT-owych, gdyż użycie płatnych serwisów nie pozwala

[25] *Intelligence Community Directive Number 301*, National Open Source Enterprise, 2006, https://irp.fas.org/dni/icd/icd-301.pdf – dostęp online 22.03.2022 r. Tłumaczenie własne.
[26] *National Defense Authorization Act for Fiscal Year 2006*, https://www.govinfo.gov/content/pkg/PLAW-109publ163/html/PLAW-109publ163.htm – dostęp online 02.03.2022 r.
[27] C. Formeller, *Introducing 15 cm HD: The Highest Clarity From Commercial Satellite Imagery*, https://blog.maxar.com/earth-intelligence/2020/introducing-15-cm-hd-the-highest-clarity-from-commercial-satellite-imagery – dostęp online 19.04.2022 r.

dowolnej osobie na powtórzenie śledztwa ze względu na ograniczenia finansowe (przy czym granica finansowa jest niejasna, gdyż dla jednej osoby wydatek rzędu 10 USD może być pomijalnie mały, a dla innej już nie). W związku z brakiem jednej, spójnej definicji, dotyczącej wykorzystania źródeł płatnych w rozpoznaniu OSINT-owym, temat ten został przez autora ujęty w jednym z pytań zadanych w ramach wywiadu eksperckiego, który jest częścią niniejszej dysertacji.

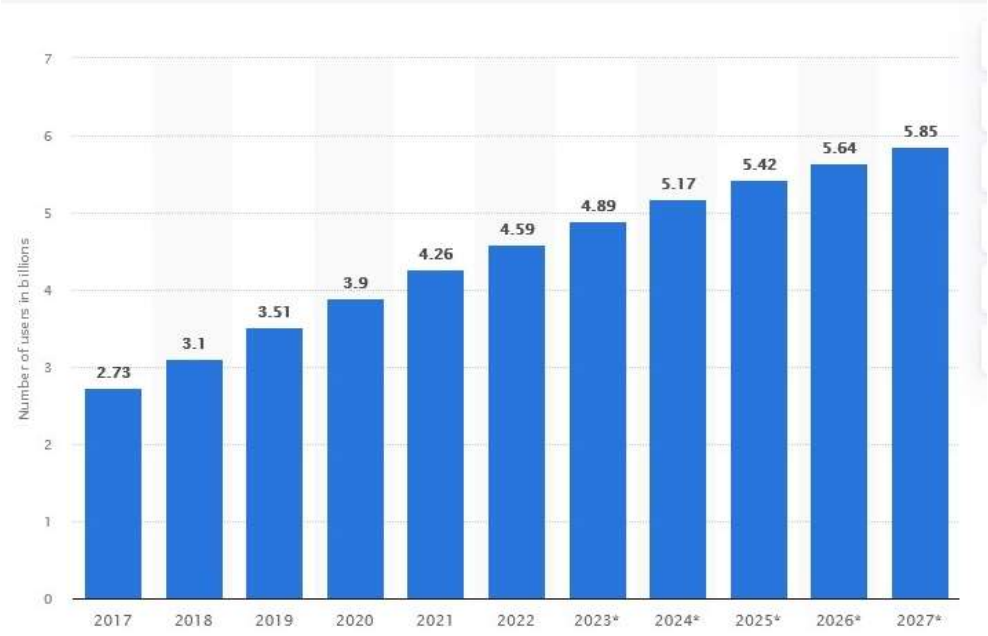
„Berkeley protocol on Digital Open Source Investigations"[28], wydany przez ONZ i Uniwersytet Kalifornijski w Berkeley poradnik[29], adresujący temat efektywnego wykorzystania informacji pochodzących z cyfrowych otwartych źródeł informacji w śledztwach, dotyczących naruszeń międzynarodowego prawa kryminalnego, humanitarnego oraz praw człowieka, w sekcji wskazującej możliwości pozyskiwania informacji wskazuje dla przykładu, że na potrzeby przedmiotowego dokumentu informacje pochodzące z otwartych źródeł obejmują także te, pozyskiwane z płatnych serwisów. Warunkiem jest jednak, że serwisy te są dostępne dla wszystkich, a nie ograniczone jedynie do określonych grup zawodowych, takich jak pracownicy organów ścigania czy licencjonowani prywatni detektywi. Jednocześnie w poradniku tym zaznaczono, że to, czy narzędzie jest płatne czy nie, powinno stanowić jedno z kryteriów w zakresie podejmowania decyzji co do wykorzystania danego narzędzia[30].

Podobnie do kwestii wykorzystania płatnych źródeł informacji można traktować kwestię wycieków danych, dostępnych w Internecie. Z jednej strony zawierają one informacje, które nie są jawne (jak na przykład hasła, poufna korespondencja czy prywatne materiały audiowizualne), jednak w momencie ich wycieku stają się informacjami dostępnymi publicznie, a więc są otwartym, dostępnym dla każdego źródłem danych.

[28] *Berkeley Protocol on Digital Open Source Investigations*, HR/PUB/20/2 (advance version), https://www.ohchr.org/sites/default/files/Documents/Publications/OHCHR_BerkeleyProtocol.pdf – dostęp online 19.04.2022 r.
[29] Dokument ten został wydany jako poradnik dla osób chcących wykorzystywać media społecznościowe jako dowody łamania praw człowieka. W ramach informacji, które mogą posłużyć jako dowody w śledztwach wymienia on zdjęcia, filmy i inne informacje publikowane w portalach społecznościowych, takich jak Facebook, Twitter czy YouTube.
[30] *Berkeley Protocol on Digital Open Source Investigations*, HR/PUB/20/2 (advance version), Annex V.

1.11.2. SOCMINT – wywiad w ramach mediów społecznościowych

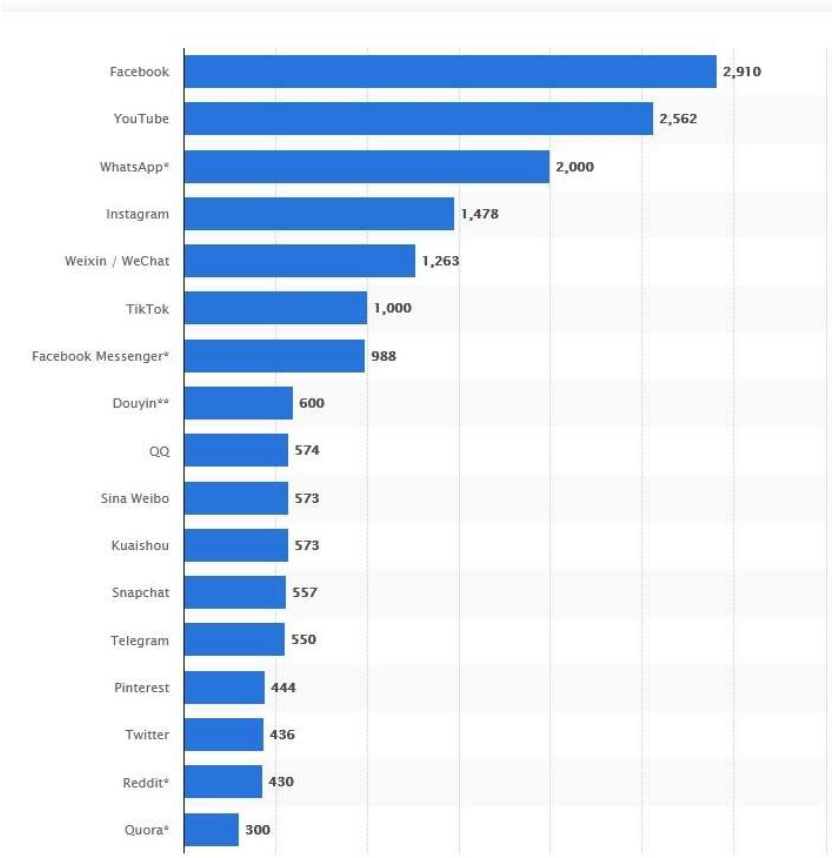
W ramach OSINT-u wyróżniana jest także jego gałąź – SOCMINT (ang. *Social Media Intelligence*), czyli rozpoznanie, bazujące na informacjach pozyskanych z mediów społecznościowych. W 2022 roku ogólna liczba osób korzystających z portali społecznościowych szacowana była na 4,59 miliarda, a do roku 2027 prognozowany jest wzrost tej liczby do 5,85 miliarda, co stanowić będzie dwukrotne zwiększenie liczby użytkowników w przeciągu 10 lat[31]. Dane te przedstawiono na Rysunku 1.



Rys. 1 – Liczba użytkowników mediów społecznościowych w latach 2017-2022 wraz z oszacowaniem dalszego wzrostu tej liczby do 2027 roku. Źródło: statista.com.

Najpopularniejszym portalem społecznościowym według zestawienia dostępnego na stronie statista.com[32] jest Facebook, a za nim plasują się takie serwisy jak: YouTube, WhatsApp, Instagram czy chiński WeChat, przy czym właściciele WhatsApp nie opublikowali w przeciągu ostatnich 12 miesięcy dokładnych danych, co sprawia, że liczba ta może w rzeczywistości różnić się od szacowanej (tak samo dla innych, oznaczonych na Rysunku 2 znakiem *).

[31] *Number of global social network users 2018-2027*, https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ – dostęp online 25.11.2022 r.
[32] *Most popular social networks worldwide as of January 2022, ranked by number of monthly active users (in millions)*, https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ – dostęp online 25.11.2022 r.



Rys. 2 – Zestawienie najpopularniejszych portali społecznościowych pod względem liczby użytkowników aktywnych w miesiącu (podane w milionach odsłon), stan na styczeń 2022. Źródło: statista.com.

W związku z ciągłym rozwojem tego typu wymiany informacji w Internecie, SOCMINT daje bardzo duże możliwości pozyskania aktualnych danych z wielu zakątków świata w różnej formie – informacji tekstowych, zdjęć, filmów, opinii, danych o osobach, firmach i organizacjach, a także umożliwia zawężenie poszukiwań poprzez analizę jedynie mediów o konkretnej tematyce lub z konkretnej lokalizacji. Można nawet uznać, że SOCMINT łączy ze sobą OSINT, IMINT i HUMINT, ponieważ z jednej strony bazuje na ogólnodostępnych danych, a z drugiej wykorzystuje techniki analizy danych obrazowych oraz pozyskiwanie informacji ze źródeł osobowych. Niekiedy granice wywiadu, prowadzonego w mediach społecznościowych mogą wykraczać nieco poza granice legalności pozyskiwania danych (a zatem jednej z podstaw OSINT-u), kiedy informacje uzyskiwane są na przykład z zamkniętych grup, do których osobom prowadzącym rozpoznanie udało się uzyskać dostęp, chociaż zagadnienie to może być

dyskusyjne, gdyż przekraczanie barier prywatności nie zawsze musi równać się łamaniu prawa.

Pozyskiwane z mediów społecznościowych dane muszą być jednak tak samo zweryfikowane, jak z każdego innego źródła, gdyż pomimo moderacji (która w wielu miejscach ogranicza się do weryfikacji czy wpisy nie łamią prawa i regulaminów serwisu), informacje publikowane przez użytkowników Internetu mogą być osobistym punktem widzenia, niepotwierdzonymi faktami lub nawet próbą manipulacji i dezinformacji.

SOCMINT, poprzez możliwość skierowania rozpoznania na serwisy o konkretnej tematyce i skierowane do wybranych grup odbiorców, jest często wykorzystywany do śledzenia cyberprzestępców, grup o charakterze terrorystycznym, przestępców seksualnych, dilerów narkotykowych, a także analizy trendów społecznych i potrzeb ich użytkowników.

1.12. Sieć indeksowana i nieindeksowana

W ramach sieci Internet możliwe jest wyróżnienie dwóch podstawowych jej obszarów: sieci indeksowanej i sieci nieindeksowanej. Podział ten odwołuje się do cechy danych dostępnych w Internecie, a mianowicie do możliwości skatalogowania ich przez wyszukiwarki[33] internetowe, takie jak Google, Bing, Yandex, Baidu i inne, lub braku takiej możliwości. Sieć indeksowana jest też często określana angielskim terminem *„surface web"* lub *„clear web"*, które odwołują się do bardziej obrazowego podziału Internetu, gdzie elementy indeksowane przez wyszukiwarki są tuż pod powierzchnią wody, natomiast elementy nieindeksowane stanowią szerszy obszar „głębi oceanu". Stąd też bierze się nazwa sieci nieindeksowanej w tym ujęciu: *„deep web"* („sieć głęboka") lub *„invisible web"* („sieć niewidoczna")[34], co odwołuje się do braku możliwości ich „zobaczenia" z wykorzystaniem zwykłej wyszukiwarki internetowej. Główną przeszkodę dla robotów indeksujących stanowi konieczność zalogowania się lub rozwiązania kodu CAPTCHA[35] w celu uzyskania dostępu lub dynamicznie generowana zawartość w odpowiedzi na wprowadzone przez użytkownika dane.

W przypadku rozważań na temat wyszukiwania danych w Internecie, najczęściej domyślnie stosuje się uproszczenie, przypisujące sieci indeksowanej wszystkie lub niemal wszystkie dostępne dla użytkownika zasoby. Rzeczywistość jednak jest zupełnie odmienna, gdyż według różnych źródeł objętość sieci indeksowanej to jedynie od około 1 do 4% [36] (a według niektórych opracowań nawet do 10%[37]) zasobów całego Internetu. Jednak według niektórych badaczy, *deep web* jest 500-krotnie lub nawet 5000-krotnie razy większa od sieci indeksowanej, a wyszukiwarki mają dostęp jedynie do 16% danych,

[33] Wyszukiwarki korzystają w tym zakresie z robotów indeksujących, które podążając za linkami na stronach tworzą mapę zawartości stron. Najczęściej spotkanymi innymi określeniami robotów indeksujących są: bot indeksujący, *web crawler*, *web spider* lub *spiderbot*.

[34] K. Król, *Geoinformation in the invisible resources of the Internet*, Geomatics, Landmanagement and Landscape No. 3, 2019, s. 53–66.

[35] CAPTCHA (skrót od ang. *Completely Automated Public Turing test to tell Computers and Humans Apart*) – technika stosowana na stronach internetowych w celu upewnienia się, że inicjatorem danej akcji (wysyłania lub pobierania danych) jest człowiek, często polegająca np. na odczytaniu zniekształconego kodu lub wskazaniu obrazów zawierających konkretne obiekty.

[36] *How Much of the Internet is the Dark Web in 2022?*, https://techjury.net/blog/how-much-of-the-internet-is-the-dark-web/#gref, dostęp online 22.06.2022 r.; J. Saleem, R. Islam and M. A. Kabir, *The Anonymity of the Dark Web: A Survey*, IEEE Access, vol. 10, 2022, s. 33628-33660.

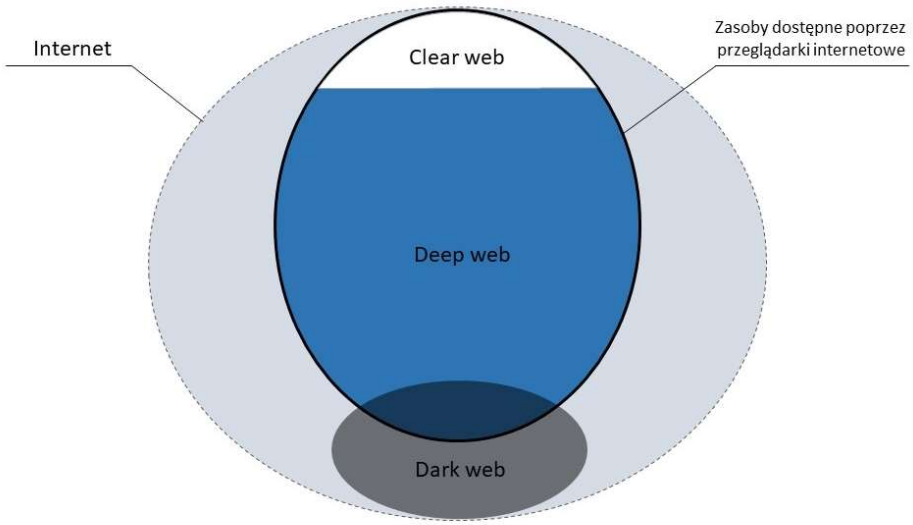
[37] M. L. Coffey, *Library application of Deep Web and Dark Web technologies*, School of Information Student Research Journal, 10(1), 2020.

zawartych w tej ostatniej[38]. Pozostałą, ogromną część zasobów stanowi sieć nieindeksowana *deep web*, czyli głównie:

- tematyczne bazy danych o ograniczonym dostępie,
- strony wymagające opłaty (np. gazety),
- fora dyskusyjne,
- zasoby dostępne z użyciem technologii stanowiącej barierę dla wyszukiwarek (np. sieci za VPN lub wykorzystujące niestandardowe protokoły / formaty danych),
- strony, do których linki nie są nigdzie opublikowane,
- zasoby *darknet*.

Określenie *„deep web”* jest niekiedy mylone z podobnym terminem – *„dark web”* lub *„dark net”* / *„darknet”* („ciemna sieć”), który jednak odwołuje się jedynie do niewielkiej jej części, dostępnej jedynie przy zastosowaniu odpowiedniego oprogramowania, umożliwiającego nawiązanie połączenia z darknetowymi serwerami. W niniejszej pracy przyjęto określenie *darknet* jako określenie sieci oraz *dark web* dla określenia serwisów w niej zawartych, chociaż często można spotkać się także z uproszczonym zamiennym stosowaniem podanych terminów jako określenia oznaczającego jedynie serwisy w tej części sieci.

[38] K. Król, *Geoinformation in the invisible resources of the Internet*, Geomatics, Landmanagement and Landscape No. 3, 2019, s. 53–66.



Rys. 3 – Przedstawienie podziału Internetu z perspektywy możliwości przeglądania poszczególnych jego części przy wykorzystaniu przeglądarek internetowych. Źródło: opracowanie własne.

Darknet jest siecią zawierającą w ogromnej większości serwisy dotyczące nielegalnych i ściganych prawem towarów oraz działań, chociaż daje także dużą możliwość zachowania anonimowości np. aktywistom lub dziennikarzom, operującym w obszarach o ścisłej cenzurze wypowiedzi.

W przypadku działań prowadzonych w darknecie, śledczy muszą dbać o bezpieczeństwo zarówno prowadzonych operacji, jak i osobiste, gdyż poruszanie się w tym obszarze sieci może ich narazić na ujawnienie własnej tożsamości lub, głównie ze względu na charakter operujących tam serwisów, na działanie złośliwego oprogramowania, mogącego zagrozić użytkownikom odwiedzającym tamtejsze strony lub portale. Wśród zalecanych[39] środków ostrożności można wyróżnić następujące sposoby uzyskiwania połączenia i przeglądania darknetu:

- skorzystanie z maszyny wirtualnej w przestrzeni chmurowej (np. Amazon Workspaces, Google Cloud, Microsoft Azure VDI, Paperspace) z zainstalowaną na niej przeglądarką (np. Tor Browser) oraz opcjonalnie

[39] *Dark Web Searching*, https://osintcombine.com/post/dark-web-searching – dostęp online 29.06.2022 r.

z zainstalowanym oprogramowaniem do połączeń przez VPN[40] na maszynie wirtualnej (jeśli dostawca usług na to pozwala);

- skorzystanie z lokalnie hostowanej maszyny wirtualnej (z wykorzystaniem np. środowiska VirtualBox), która może być zainstalowana przez użytkownika lub uruchomiona z wcześniej pobranego z Internetu obrazu (np. Trace Labs VM), skonfigurowanie na niej VPN-a oraz uruchomienie na niej przeglądarki (np. Tor Browser);
- skorzystanie z komputera przeznaczonego wyłącznie do działań śledczych, z systemem operacyjnym w wersji live[41] (np. dystrybucja Tails) i skonfigurowanym VPN-em.

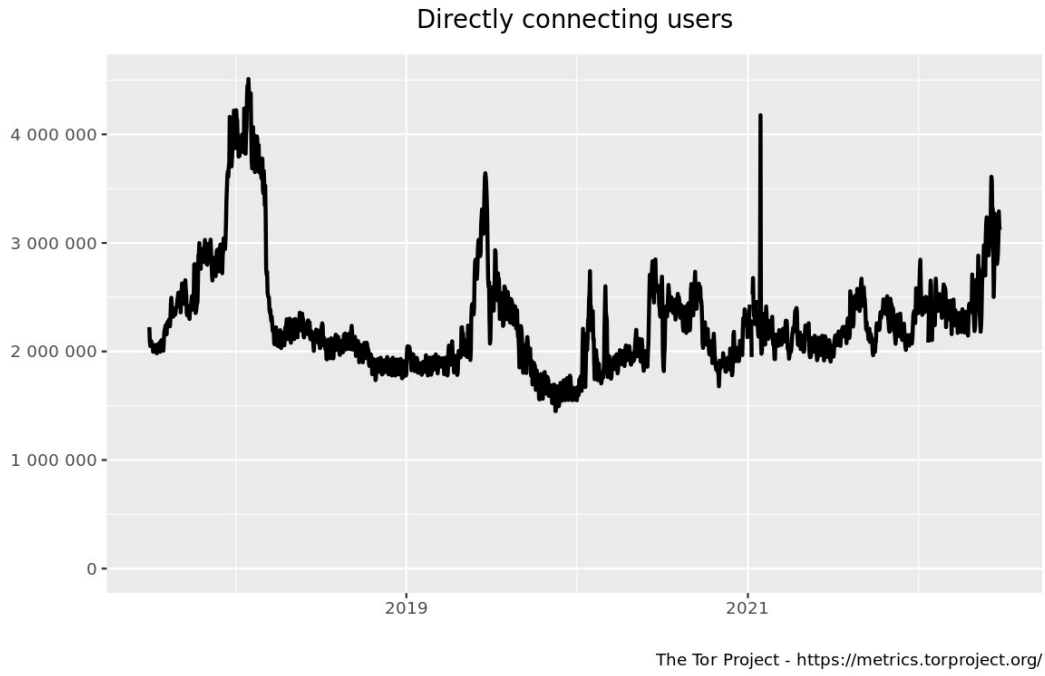
Najpopularniejszym sposobem korzystania z zasobów darknetowych jest aktualnie sieć Tor[42]. Ze względu na swoją popularność i szeroką dostępność dla użytkowników Internetu, wielokrotnie można spotkać się z uogólnieniem, że Tor jest synonimem darknetu, podczas gdy istnieją także inne odmiany tego obszaru sieci, jak np. I2P (Invisible Internet Project) czy Freenet. Z uwagi na specyfikę i rozległość darknetu, w dysertacji uwzględniono tylko sieć Tor.

Rysunek 4 ilustruje liczbę użytkowników podłączających się codziennie bezpośrednio do sieci Tor w okresie pięciu lat (od 30.06.2017 r. do 30.06.2022 r.).

[40] VPN (skrót od ang. Virtual Private Network) – wirtualna sieć prywatna – sposób tunelowania ruchu sieciowego w Internecie, pozwalający na zapewnienie jego bezpieczeństwa przez osoby podłączone do sieci pomiędzy węzłem wejściowym (najczęściej maszyną użytkownika) a węzłem wyjściowym (zlokalizowanym w sieci wewnętrznej np. w przypadku firmowych dostępów VPN lub na serwerze dostawcy usług VPN w przypadku dostawców komercyjnych). Zabezpiecza to ruch sieciowy przed podglądem lub modyfikacją np. na poziomie dostawcy usług internetowych. W działaniach OSINT-owych wykorzystywane jest często do zabezpieczenia przed wykryciem lokalizacji osoby poszukującej, gdyż adresem źródłowym, logowanym przez serwery docelowe będzie w takich przypadkach jedynie adres węzła VPN.

[41] Systemy operacyjne w wersji live są uruchamiane z napędu CD/DVD lub pendrive'a podłączonego do komputera i nie przechowują danych ani ustawień użytkownika, tzn. po wyłączeniu komputera wszystkie niezapisane na innych nośnikach dane są tracone. Dodatkowo, systemy takie jak Tails lub Whonix posiadają dodatkowe funkcjonalności, umożliwiające ukrycie prawdziwej tożsamości użytkownika w Internecie.

[42] Tor – otwarta i darmowa wirtualna sieć komputerowa, umożliwiająca komunikację z wysokim poziomem anonimowości, co jest szczególnie ważne dla osób dbających o prywatność, ale w rejonach świata objętych cenzurą. Do swojego działania wykorzystuje mechanizmy trasowania cebulowego. Jej nazwa jest właśnie skrótem od określenia „The Onion Router", które nawiązuje do sposobu zabezpieczania danych przesyłanych w ramach tej sieci, poprzez ich wielokrotne szyfrowanie i przesyłanie przez kolejne węzły.



Rys. 4 – Liczba użytkowników podłączonych do sieci Tor – okres od 30.06.2017 r. do 30.06.2022 r. (źródło: Tor Metrics – https://metrics.torproject.org – dostęp online 30.06.2022 r.)

Z punktu widzenia rozpoznania OSINT-owego, zasoby w sieci *deep web*, włączając w to także *darknet*, są istotnym elementem poszukiwań, chociaż co do tych ostatnich można mieć wątpliwości w zakresie legalności ich pozyskiwania (np. w przypadku danych z wycieków lub wykradzionych dokumentów).

Ze względu na przenikanie się sieci *surface web* i *dark web*, istnieje możliwość deanonimizacji technologii i osób stojących za serwisami darknetowymi poprzez poszukiwanie informacji, które przypadkowo lub przez nieostrożność obsługujących je osób przedostały się do sieci ogólnodostępnej. Mogą to być adresy e-mail, adresy IP serwerów, loginy, a nawet nazwiska osób operujących w darknecie. Dla śledztwa OSINT-owego tego typu znalezisko może stanowić podstawę do rozszerzenia lub ukierunkowania poszukiwań i w efekcie deanonimizację celu.

Wyszukiwanie danych w sieci nieindeksowanej wymaga od śledczych także dodatkowych umiejętności odnajdywania powiązań i znajomości możliwych do przeszukania źródeł ze względu właśnie na brak możliwości skorzystania ze standardowych wyszukiwarek. Niemniej jednak istnieją dwa sposoby na wyszukiwanie informacji w sieci Tor: korzystanie z wyszukiwarek typowo nastawionych na

przeszukiwanie stron w domenie .onion[43] oraz korzystanie z serwisów typu tor2web, przy czym ten ostatni sposób może narazić użytkownika na ujawnienie swojej tożsamości. Do wyszukiwarek w sieci Tor można zaliczyć takie serwisy jak m.in.: Ahmia.fi[44] (dostępna także z sieci *surface web* i obsługująca również sieć I2P), Torch[45], a także serwis The Hidden Wiki[46], który nie jest wyszukiwarką *sensu stricto*, ale posiada bardzo rozbudowany spis przydatnych adresów w sieci Tor.

Należy zwrócić uwagę, że nie wszystkie osoby pragnące uzyskać dostęp do zasobów sieci *deep web* są w stanie go uzyskać, gdyż niektóre serwisy wymagają np. adresu poczty e-mail w domenie akademickiej lub wykazania powiązania z organami ścigania czy też inną grupą, dla której konkretne dane są przeznaczone.

[43] Adresy w domenie .onion są specyficzne dla sieci Tor, a dostęp do nich umożliwiają jedynie przeglądarki takie jak Tor Browser. W momencie pisania niniejszej pracy obowiązuje specyfikacja v3 dla adresów .onion, która określa m.in. jego długość, tzn. wszystkie nazwy domen składają się z 56 znaków i sufiksu .onion. Wcześniejsze, krótsze adresy v2 (16-znakowe) od 15 lipca 2021 roku nie są już wspierane.
[44] Adres w sieci Tor: http://juhanurmihxlp77nkq76byazcldy2hlmovfu2epvl5ankdibsot4csyd.onion/
[45] Adres w sieci Tor: http://torchdeedp3i2jigzjdmfpn5ttjhthh5wbmda2rr3jvqjg5p77c54dqd.onion/
[46] Adres w sieci Tor: http://zqktlwiuavvvqqt4ybvgvi7tyo4hjl5xgfuvpdf6otjiycgwqbym2qad.onion/

1.13. Rekonesans pasywny i aktywny

Etap poszukiwania informacji w ramach rozpoznania otwartoźródłowego można podzielić w ogólności na dwie kategorie, w zależności od poziomu interakcji z obiektem (celem) przedmiotowych działań:

- rekonesans pasywny, który zakłada brak bezpośredniej interakcji z obiektem w celu ukrycia przed nim faktu prowadzenia rozpoznania;
- rekonesans aktywny, który dopuszcza wejście w interakcję z obiektem, jedocześnie wprowadzając niebezpieczeństwo wykrycia.

Rodzaj prowadzonego rekonesansu w ramach działań OSINT-owych powinien być jasno określony na samym początku, w zależności od charakteru i jawności prowadzonego rozpoznania – jest to jeden z elementów bezpieczeństwa operacyjnego (OPSEC[47]), które stanowi jeden z kluczowych aspektów poprawnego prowadzenia działań OSINT-owych. Zbyt późne określenie konieczności zastosowania wyłącznie rekonesansu pasywnego może mieć wpływ na brak powodzenia całej operacji.

Należy wskazać także, iż działania pasywne są obarczone dużo większym błędem ze względu na konieczność wykorzystania jedynie zebranych wcześniej danych o technologii (np. ze skanów infrastruktury sieciowej, wykonanych przez serwisy takie jak Shodan[48] lub Censys[49]) oraz osobach (np. dane dostępne w dokumentach i rejestrach publicznych lub pochodzące z wycieków), które mogą nie być kompletne, adekwatne do celu prowadzonego rozpoznania i aktualne na moment jego przeprowadzania. Dla informacji zbieranych w sposób pasywny wysoki jest także poziom tzw. szumu informacyjnego[50], co również niekorzystnie wpływa na przebieg późniejszych działań analitycznych, w ramach których trudniejsze jest oddzielenie informacji istotnych od nieistotnych. Aktywne rozpoznanie daje dużo szybsze i dokładniejsze efekty, jednak niesie za sobą prawdopodobieństwo ujawnienia nie tylko informacji o fakcie bycia analizowanym, ale także ujawnienia szczegółów dotyczących osoby, organizacji lub infrastruktury śledzącej.

[47] Temat szerzej poruszany w rozdziale **Błąd! Nie można odnaleźć źródła odwołania.**.
[48] Shodan (https://www.shodan.io) – serwis internetowy umożliwiający wyszukiwanie informacji na temat infrastruktury teleinformatycznej podłączonej do Internetu na podstawie zadanych filtrów: adresacji IP, lokalizacji, rodzaju oprogramowania itp.
[49] Censys (https://search.censys.io) – wyszukiwarka urządzeń podłączonych do Internetu, podobnej klasy jak Shodan.
[50] Szum informacyjny – nadmiar informacji utrudniający wyodrębnienie informacji prawdziwych i istotnych (za: Słownik języka polskiego PWN).

Należy także mieć na uwadze, iż w niektórych śledztwach OSINT-owych prowadzenie rozpoznania pasywnego jest zakazane. Dzieje się tak m.in. podczas poszukiwania osób zaginionych przez kanadyjską organizację non-profit – Trace Labs. W ramach Search Party CTF – cyklicznego, ogólnoświatowego wydarzenia, w ramach którego uczestnicy, zorganizowani w zespoły (maksymalnie 4-osobowe), prowadzą OSINT-owe rozpoznanie, mające na celu zebranie jak największej ilości informacji o osobach zaginionych, które następnie są przekazywane organom ścigania. Pomimo dość niepozornej nazwy, przedmiotem śledztw są osoby faktycznie zaginione, których dane nie są publikowane poza zamkniętym forum uczestników danej edycji wydarzenia. Zgodnie z regułami[51], uczestnikom nie wolno prowadzić rozpoznania aktywnego, tzn. komunikować się z osobami poszukiwanymi ani ich rodzinami, tagować ich na zdjęciach lub wydarzeniach w mediach społecznościowych, lajkować ich wpisów ani wchodzić w żadną inną interakcję. W przypadku wykrycia interakcji, osoby takie są natychmiast dyskwalifikowane z udziału w wydarzeniu. Tego rodzaju środki bezpieczeństwa wprowadzone są po to, aby nie utrudniać faktycznie toczących się śledztw ani nie powodować dodatkowej traumy u rodziny osób zaginionych.

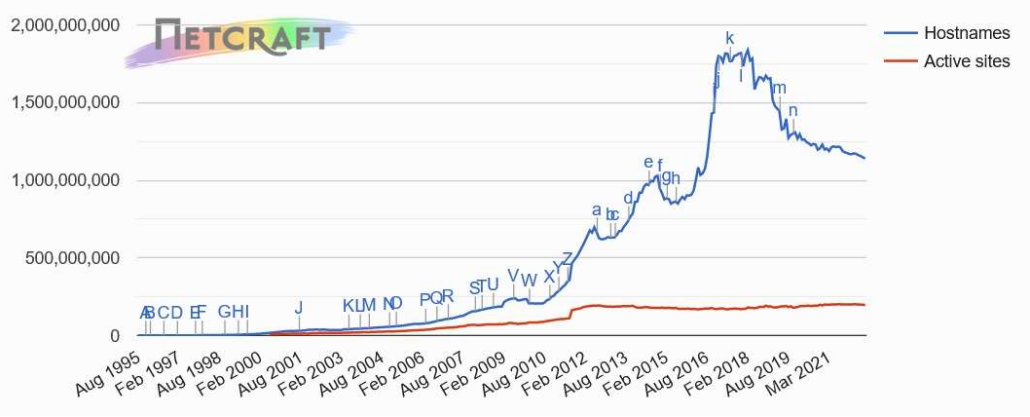
[51] Trace Labs, *Search Party Rules of Engagement* – https://www.tracelabs.org/about/search-party-rules – dostęp online 09.09.2022 r.

1.13.1. Rekonesans pasywny

Prowadzenie pasywnego OSINT-u obejmuje działania, możliwe do wykonania z wykorzystaniem odpowiednich narzędzi, z których wybrano i przedstawiono reprezentatywne przykłady. Wybór konkretnych narzędzi został dokonany na podstawie ich popularności w Internecie oraz własnego doświadczenia autora niniejszej pracy. Narzędzia zostały pogrupowane w kategorie, w zależności od ich przeznaczenia.

1.13.1.1. Wyszukiwanie w witrynach internetowych na podstawie zapytania tekstowego

Liczba aktywnych stron internetowych w lipcu 2022 roku szacowana była na około 1,1 miliarda (a dokładnie 1 139 467 659 stron[52]), powiązanych z 271 728 559 domenami i 12 341 172 hostami podłączonymi do sieci. Istnieją także statystyki określające liczbę stron na niemal 2 miliardy[53]. Należy jednak pamiętać, że zastawienia te zawierają także domeny nieaktywne, tzn. jedynie zarezerwowane przez firmy hostingowe lub na potrzeby reklamowe, których zawartość jest generowana automatycznie. Na poniższym wykresie widoczne jest, że liczba stron aktywnych (zidentyfikowanych przez algorytmy[54] jako stworzone przez człowieka) wynosi jedynie ok. 17% (dokładnie jest to 197 046 670) ogólnej liczby zidentyfikowanych stron w Internecie.



Rys. 5 – Wykres pokazujący liczbę zarejestrowanych stron internetowych oraz liczbę aktywnych stron od 1995 roku. Źródło: July 2022 Web Server Survey – https://news.netcraft.com/archives/2022/07/28/july-2022-web-server-survey.html – dostęp online 18.08.2022 r.

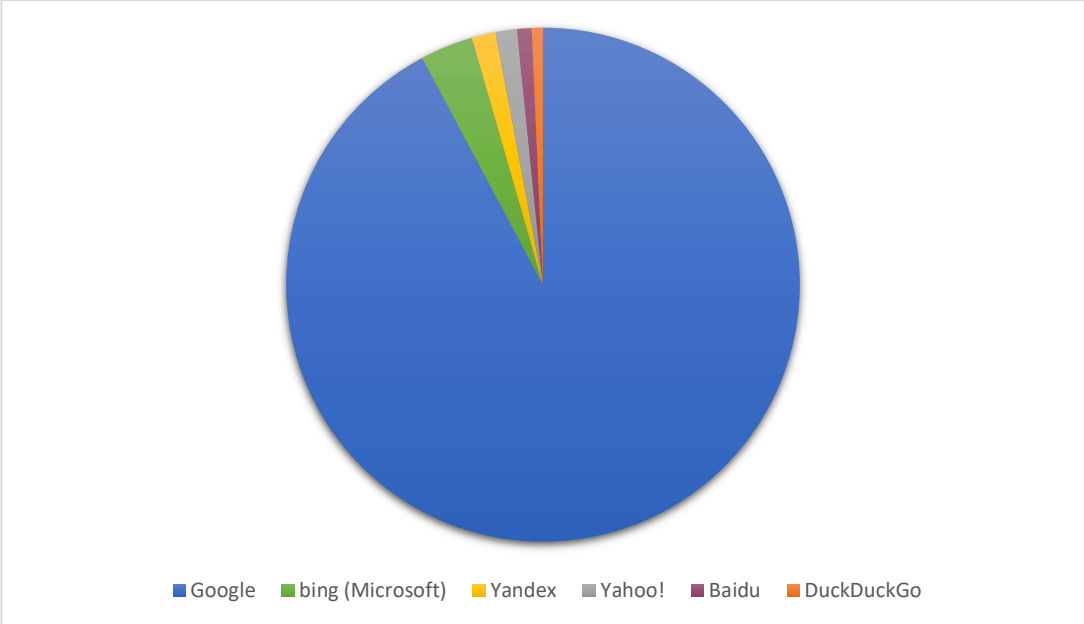
[52] *July 2022 Web Server Survey* – https://news.netcraft.com/archives/2022/07/28/july-2022-web-server-survey.html – dostęp online 18.08.2022 r.
[53] *Internet Live Stats: Total number of Websites* – https://www.internetlivestats.com/total-number-of-websites/ – dostęp online 18.08.2022 r.
[54] *How many active sites are there?* – https://www.netcraft.com/active-sites/ – dostęp online 18.08.2022 r.

Stan popularności poszczególnych wyszukiwarek internetowych pokazuje ogromną przewagę Google nad wyszukiwarkami innych firm. Podział rynku wyszukiwarek w lipcu 2022 roku przedstawiał się następująco:

Wyszukiwarka	**Udział w rynku**
Google	91,43%
Bing (Microsoft)	3,30%
Yandex	1,49%
Yahoo!	1,33%
Baidu	0,91%
DuckDuckGo	0,70%

Tabela 1 – Udział poszczególnych wyszukiwarek w rynku w lipcu 2022 roku. Źródło: Search Engine Market Share Worldwide - July 2022 – https://gs.statcounter.com/search-engine-market-share – dostęp 18.08.2022 r.

Aby lepiej unaocznić dysproporcję pomiędzy, powyższe dane przedstawiono w formie wykresu kołowego na Rysunku 6:



Rys. 6 – Udział poszczególnych wyszukiwarek w rynku w lipcu 2022 roku. Źródło: Search Engine Market Share Worldwide - July 2022 – https://gs.statcounter.com/search-engine-market-share – dostęp online 18.08.2022 r.

Z powodu swojej popularności Google jest głównym źródłem poszukiwań informacji w sieci indeksowanej, jednak ze względu na różnorodność sposobów indeksowania informacji przez różne wyszukiwarki, nie jest wskazane opieranie się jedynie na tej wyszukiwarce. Możliwe są przypadki, w których informacje, które nie będą w bazach Google, będą prezentowane np. przez wyszukiwarkę Bing.

Sposobem na ukierunkowanie wyszukiwania tekstowego w celu otrzymania mniejszej i bardziej zawężonej liczby jego wyników, co pozwala na łatwiejsze wyszukiwanie pożądanych informacji, jest wykorzystywanie operatorów wyszukiwania. Zarówno Google, jak i inne wyszukiwarki tekstowe, a także portale pozwalające na rekonesans konkretnych zagadnień (np. portal Shodan, który bardziej szczegółowo opisany został w podrozdziale 1.13.1.3), pozwalają na używanie określonych operatorów lub ich kombinacji. Operatory wyszukiwania określają zakres, w jakim ukierunkowane jest wyszukiwanie, a ich argumenty określają jakie elementy tego zakresu mają być brane pod uwagę. Aby lepiej przedstawić to zagadnienie, poniżej znajduje się zestawienie przykładowych operatorów wyszukiwania w Google dla wyszukiwarek tekstowych wraz z ich objaśnieniem.

Operator	**Działanie**	**Przykład zastosowania**
" "	Wyszukiwanie określonej frazy – wyniki muszą zawierać konkretny ciąg znaków podany w cudzysłowie. Przydatny w przypadku wyszukiwania tekstu zawierającego spacje, który bez użycia cudzysłowu byłby traktowany jako wyszukiwanie wszystkich podanych słów, jednak w losowej kolejności i niekoniecznie użytych obok siebie.	"Nauki o bezpieczeństwie"
filetype: lub ext:	Pozwala na wyszukiwanie jedynie wyników, będących plikami w określonym formacie, przy czym rozszerzenie pliku musi być podane bezpośrednio po dwukropku, bez spacji. Dla wyszukiwarki Google nie ma znaczenia czy zastosujemy formę „filetype:” czy „ext:”, jednak np. Bing rozpoznaje tylko formę „filetype:”, a w wyszukiwarce Yandex rolę tę pełni operator „mime:”. Wyszukiwarki mają	"Raport dzienny" filetype:pdf

	także ograniczoną liczbę typów plików[55], które można wyszukiwać z użyciem tego operatora.	
intitle:	Powoduje wyszukanie stron lub dokumentów, zawierających w tytule słowo lub frazę (więcej słów w cudzysłowie) podane po operatorze. W przypadku chęci wyszukania większej ilości słów w tytule, konieczne jest powtórzenie operatora przed kolejnym wyrazem.	intitle:”raport z badań”
allintitle:	Działa podobnie jak „intitle:”, jednak wyszukuje strony, które w tytule posiadają wszystkie wymienione po tym operatorze wyrazy lub frazy.	allintitle: raport końcowy
inurl:	Powoduje wyszukanie stron lub dokumentów, zawierających w adresie słowo lub frazę podane po operatorze.	inurl:admin
allinurl:	Działa podobnie jak „inurl:”, jednak wyszukuje strony, które w adresie posiadają wszystkie wymienione po tym operatorze wyrazy lub frazy, przy czym ignorowana jest interpunkcja, co powoduje, że wyrazy złączone znakiem - lub / będą wyszukiwane jako osobne wyrazy.	allinurl: login admin
intext:	Powoduje wyszukanie stron lub plików, zawierających w treści słowo lub frazę podaną po operatorze.	intext:panel
allintext:	Działa podobnie jak „intext:”, jednak wyszukuje strony, które w tytule posiadają wszystkie wymienione po tym operatorze wyrazy lub frazy.	allintext: panel administracyjny

[55] *File types indexable by Google* – https://support.google.com/webmasters/answer/35287?hl=en – dostęp online 24.08.2022 r.

site:	Powoduje wyszukanie wyników w zakresie podanej po operatorze domeny.	"wniosek paszportowy" site:gov.pl
-	Znak „-” wyklucza słowo, które po nim następuje (bez spacji pomiędzy znakiem a słowem) z wyszukiwania. Może być stosowane także z innymi operatorami, co pozwala na wykluczanie np. całych domen lub typów plików z wyników.	"Raport dzienny" -filetype:pdf
*	Znak „*” zastępuje dowolny ciąg znaków we frazie wyszukiwania. Może być używany np. w celu znalezienia osób, których pierwsze imię i nazwisko jest nam znane, ale drugiego imienia nie znamy, a wiemy, że te osoby posługują się nim w Internecie.	"Jan * Kowal"
AND lub &	Operator iloczynu logicznego – wszystkie frazy objęte operatorem muszą występować w wynikach wyszukiwania. Możliwe jest zarówno zastosowanie operatora słownego „AND”, jak i znaku „&”.	samolot AND Łoś AND uzbrojenie
OR lub |	Operator sumy logicznej – co najmniej jedna fraza objęta operatorem musi występować w wynikach wyszukiwania (ale mogą także występować w większej ilości lub nawet wszystkie). Możliwe jest zarówno zastosowanie operatora słownego „OR”, jak i znaku „|”.	samochód OR auto OR pojazd
( )	Nawiasy ograniczają zakres działania operatorów logicznych, przez co możliwe jest budowanie bardziej złożonych logicznie zapytań.	Fiat AND (125p OR 126p)
cache:	Pokazuje wynik ostatniego zapisu zawartości strony przez Google, której	cache:microsoft.com

	adres musi być podany po operatorze, bez spacji. Umożliwia to np. weryfikację jakie treści prezentowała dana strona kilka godzin wcześniej (czas ten zależy od momentu indeksowania danej strony przez Google).	
related:	Umożliwia sprawdzenie jakie strony według Google są związane ze stroną, której adres jest podany po operatorze. Umożliwia to weryfikację jak mechanizmy Google kategoryzują daną stronę.	related:cnn.com
..	Operator .. działa w połączeniu z dwiema liczbami – jedną poprzedzającą go i jedną następującą po nim (zapisane bez spacji). Powoduje on wyszukanie wartości od pierwszej liczby do ostatniej. Google nie pozwala jednak na wyszukiwanie zbyt dużych zakresów liczb, dodatkowo często sprawdzając czy zapytania są formułowane przez człowieka (poprzez mechanizmy CAPTCHA). Tego typu zabezpieczenie ma przeciwdziałać m.in. zbyt łatwemu wyszukiwaniu numerów identyfikacyjnych, takich jak chociażby amerykańskie numery ubezpieczenia społecznego (obecnie będące, pod względem ich wykorzystywania, odpowiednikiem naszego numeru PESEL).	"raport roczny" 2020..2022

Tabela 2 – Zestawienie najpopularniejszych operatorów wyszukiwania w Google

Jeszcze lepsze wyniki wyszukiwania otrzymać można dzięki kombinacji operatorów i ich parametrów. Tego typu zestawienia, za pomocą których możliwe jest znalezienie słabych punktów serwisów i urządzeń w Internecie, są często określane jako Google Dorks lub Google hacking, chociaż oryginalnie określenie „Google Dork" oznaczało osobę, która nieudolnie zabezpieczała stronę dostępną w Internecie. Później twórca tego określenia (jak i pierwszego znanego na szeroką skalę zestawu zapytań, pomagającego znajdować podatności serwisów www), Johnny Long, zmienił nieco znaczenie tego określenia na osobę, która wykorzystuje znajomość tego typu zapytań do wyszukiwania poufnych informacji w Internecie[56]. Od roku 2002, w którym Johnny Long opublikował pierwsze zestawienie Google Dorks, baza tych zapytań systematycznie się rozrastała i obecnie dostępna jest na stronie Google Hacking Database (GHDB)[57]. Przykładowym zapytaniem z rodziny „Google Dorks", które pozwala na wyszukanie w Internecie serwerów z kopią zapasową bazy danych umieszczoną w możliwym do wylistowania folderze „backup" i posiadającą w nazwie pliku lub katalogu słowo „sql", jest zapytanie:

intitle:"Index of" sql intext:"/backup"

Określenie *„dorks"* jest też używane w odniesieniu do zestawów operatorów innych wyszukiwarek (nie tylko Google) i jest dzisiaj terminem oznaczającym po prostu konkretne kombinacje parametrów wyszukiwania, pozwalające na wyszukanie interesujących wyników w Internecie.

Wspomnieć należy także o możliwości wyszukiwania treści zarchiwizowanych przez różne serwisy internetowe, np. Internet Archive[58], Ta internetowa biblioteka non-profit jest zbiorem zapisanych wersji stron internetowych, oprogramowania oraz dokumentów, publikowanych w Internecie. Poprzez wykorzystanie własnych botów do indeksowania sieci, a także z pomocą osób ręczne wskazujących elementy do archiwizacji, możliwe było stworzenie archiwum ponad 771 miliardów stron internetowych[59]. Poniżej przedstawiono wykres liczby archiwizacji jednej z najstarszych stron polskiego Internetu – onet.pl – na osi czasu.

[56] J. Long, *The Google Hacker's Guide. Understanding and Defending Against the Google Hacker* – https://pdf.textfiles.com/security/googlehackers.pdf – dostęp online 29.08.2022 r.
[57] Dostępna pod adresem: https://www.exploit-db.com/google-hacking-database
[58] https://archive.org
[59] Dane na dzień 22.11.2022 r.

Rys. 7 – Wykres liczby archiwizacji strony onet.pl na osi czasu. Źródło: archive.org.

Innym sposobem na uzyskanie dostępu do zarchiwizowanej wersji danej strony internetowej, jest wykorzystanie operatora „cache" w wyszukiwarce Google lub skorzystanie z innego narzędzia, które oferuje możliwość archiwizacji stron, np. archive.today[60] lub archiwów kontekstowych, jak np. UNSECO Web Archive[61].

1.13.1.2. Wyszukiwanie obrazów i wyszukiwanie na podstawie obrazu

Wczesne wyszukiwarki pozwalały znajdować jedynie wyniki w formie linków do stron internetowych oraz informacji tekstowej o ich treści. Dopiero 13 października 1998 roku, Altavista, jedna z wiodących w ówczesnym czasie wyszukiwarek, wprowadziła na swojej stronie usługę „Photo finder", która na podstawie wprowadzonych słów zwracała wyniki w formie miniatur obrazów, wyszukanych w swoich zindeksowanych zbiorach wraz z odnośnikami do nich[62]. Początkowo na stronach z wynikami nie widniały żadne informacje dotyczące praw autorskich odnośnie do prezentowanych obrazów, jednak środowiska fotografów i artystów zaprotestowały przeciw tego typu udostępnianiu ich prac, co poskutkowało dodaniem na stronie wyszukiwania informacji o konieczności skontaktowania się z właścicielem danego obrazu w celu upewnienia się co do możliwości jego wykorzystania. Takie rozwiązanie nie spotkało się jednak z akceptacją środowisk twórczych[63], co jednak nie przeszkodziło Altaviście kontynuować działanie wyszukiwarki obrazów, która zaoferowała im możliwość wykluczenia swoich obrazów

[60] https://archive.ph
[61] https://webarchive.unesco.org
[62] *AltaVista Photo Finder, and how to keep your images "unfound"* – http://www.photodude.com/av.htm – dostęp online 24.08.2022 r.
[63] *AltaVista Photo Finder Has Artists Concerned* – https://web.archive.org/web/19990427131502/http://www.searchenginewatch.com/sereport/9811-photofinder.html – dostęp online 24.08.2022 r.

z wyszukiwania poprzez odmowę indeksowania ich przez robota Altavisty na poziomie pliku robots.txt[64].

Wyszukiwanie informacji na podstawie zapytań tekstowych ograniczone jest jednak przez umiejętność odpowiedniego opisania przez człowieka tego, co chciałby uzyskać w odpowiedzi. Z związku z faktem, że słowa są tylko reprezentacją obrazów, które zapamiętujemy i uwieczniamy w postaci zdjęć, znacznie wydajniejszym wydaje się wyszukiwanie na podstawie konkretnego obrazu.

W połowie 2001 roku Google uruchomiło możliwość wyszukiwania obrazów, które w tamtym czasie nadal bazowało na zapytaniu tekstowym, jednak możliwość bezpośredniego otrzymania wyników graficznych z puli 150 milionów plików[65] w odpowiedzi na zadaną frazę wyszukiwania była dużym rozszerzeniem możliwości, której oczekiwało bardzo wielu użytkowników[66]. Do 2005 roku ilość zindeksowanych przez Google obrazów przekroczyła 1 miliard plików[67].

Wyszukiwarki obrazów stopniowo powiększały swoje możliwości, chociażby poprzez wykorzystanie mechanizmów CBIR[68] do lepszego wyszukiwania obrazów na podstawie ich zawartości (kolorów, kształtów, tekstur itp.), a nie opisujących je metadanych, co dawało znacznie większe możliwości znalezienia dokładnie tego pliku, który był poszukiwany przez osobę formułującą zapytanie.

Możliwość wyszukiwania obrazów na podstawie podanego pliku z grafiką pojawiła się w 2008 roku wraz z początkiem działania serwisu TinEye, który do dzisiaj jest jednym z najbardziej rozpoznawalnych serwisów oferujących wyszukiwanie za

---

[64] Plik robots.txt – plik, którego zawartość definiuje możliwość indeksowania wszystkich lub wybranych elementów strony internetowej przez boty indeksujące. Plik ten powinien być zawsze umieszczony w głównym katalogu zasobów www danej domeny. Więcej informacji na temat tego pliku zawarto w rozdziale 1.20.

[65] *News Watch; A Quick Way to Search For Images on the Web* – https://www.nytimes.com/2001/07/12/technology/news-watch-a-quick-way-to-search-for-images-on-the-web.html – dostęp online 24.08.2022 r.

[66] Według ówczesnego dyrektora generalnego Google, Erica Schmidta, pomysł na wyszukiwarkę obrazów zrodził się w odpowiedzi na występ Jennifer Lopez na rozdaniu nagród Grammy Awards w 2000 roku, gdzie wystąpiła ona w tak charakterystycznej, zielonej sukni, że wygenerowało to bardzo dużą ilość zapytań w Google od osób, które chciały na własne oczy zobaczyć tę odważną kreację (źródło: Eric Schmidt, *The Tinkerer's Apprentice* – https://www.project-syndicate.org/magazine/google-european-commission-and-disruptive-technological-change-by-eric-schmidt-2015-01 – dostęp online 24.08.2022 r.).

[67] *Ooh! Ahh! Google Images presents a nicer way to surf the visual web* – https://googleblog.blogspot.com/2010/07/ooh-ahh-google-images-presents-nicer.html – dostęp online 24.08.2022 r.

[68] CBIR (skrót od ang. *Content-based image retrieval*) – technika wyszukiwania obrazów w oparciu o ich graficzną zawartość.

pomocą obrazu z bazą 55,9 miliarda plików[69]. 27 października 2009 roku Google zaoferowało, po półrocznych testach, użytkownikom swojej wyszukiwarki korzystanie z opcji „Znajdź podobne obrazy", co umożliwiło wyszukiwanie grafik podobnych wizualnie według mechanizmów wyszukiwarki[70]. Z kolei 7 grudnia 2009 roku Google udostępniło użytkownikom telefonów komórkowych aplikację Google Goggles[71], która umożliwiała szukanie na podstawie zdjęcia, wykonanego aparatem fotograficznym telefonu, a 11 czerwca 2011 roku, niemal dokładnie 10 lat po rozpoczęciu działania wyszukiwarki obrazów Google, firma ta dała możliwość wyszukiwania obrazem także użytkownikom komputerów[72], którzy od teraz mogli wyszukiwać obrazem poprzez wskazanie wzorcowego pliku na dysku lub podanie adres URL obrazu w Internecie.

Biorąc pod uwagę fakt, że wyszukiwarki używają różnych metod analizy obrazów, różnią się przez to skutecznością w wyszukiwaniu grafik, należących do różnych kategorii. Podczas, gdy Google wykorzystuje nie tylko obraz, na podstawie którego tworzy model matematyczny i w powiązaniu z metadanymi opisującymi obraz poszukuje podobnych grafik w swojej bazie, jeden z jego głównych konkurentów – Yandex – wykorzystuje dodatkowo mechanizmy sztucznej inteligencji, co powoduje, że wyniki są często trafniejsze, niż te, wyszukiwane przez Google.

Obecnie wyszukiwarki rozszerzają swoje możliwości o kolejne funkcjonalności, takie jak na przykład rozpoznawanie tekstu w obrazie. W tym aspekcie Google jednak pozostaje w tyle za konkurencją, a opcję tę oferują Bing i Yandex, przy czym Bing potrafi jedynie rozpoznawać alfabet łaciński, a Yandex posiada funkcjonalność rozpoznawania tekstu zapisanego w różnych alfabetach, co czyni go jednym z najużyteczniejszych narzędzi w tym obszarze.

W zakresie porównania możliwości trzech najpopularniejszych wyszukiwarek – Google, Bing i Yandex – pod względem trafności wyszukiwania za pomocą obrazu przez nie, autor niniejszej dysertacji przeprowadził własne badanie w 2021 roku, których

[69] Stan na dzień 26.08.2022 r. – źródło: https://tineye.com/faq#count
[70] *Similar Images graduates from Google Labs*, https://googleblog.blogspot.com/2009/10/similar-images-graduates-from-google.html – dostęp online 22.08.2022 r.
[71] *Relevance meets the real-time web*, https://googleblog.blogspot.com/2009/12/relevance-meets-real-time-web.html – dostęp online 22.08.2022 r.
[72] *Knocking down barriers to knowledge*, https://googleblog.blogspot.com/2011/06/knocking-down-barriers-to-knowledge.html – dostęp online 22.08.2022 r.

wyniki opublikowane zostały na portalu sekurak.pl[73] oraz w wersji anglojęzycznej na stronie Securitum Research[74]. Badanie to skupiło się na sprawdzeniu umiejętności rozpoznania: miasta, pojazdu, owoców, napisów na tablicy informacyjnej w nie-łacińskim alfabecie, znaków graficznych logo oraz twarzy znanej osoby.

Przeprowadzone wyszukiwanie z użyciem zdjęcia, przedstawiającego widok obszaru miejskiego wykazało, że Google i Yandex rozpoznały miasto, a Bing miał z tym problem. Na korzyść Binga przemawia jednak funkcjonalność wyboru elementu obrazu, który ma posłużyć jako baza wyszukiwania. Taką samą opcję oferuje Yandex. Chcąc wyszukać fragment obrazu w Google, musimy niestety ręcznie go wyedytować lub skorzystać z dodatku do przeglądarki, dzięki któremu możliwe będzie zaznaczenie interesującego nas obszaru wyświetlonego w przeglądarce obrazu i wysłanie go od razu do jednej lub kilku obsługiwanych wyszukiwarek. Taką możliwość oferuje chociażby dodatek Search by Image[75]. Wykorzystani funkcjonalności wskazywania jedynie fragmentu obrazu do analizy działa na korzyść wyszukiwarki Bing, która dzięki temu zabiegowi potrafi lepiej rozpoznać charakterystyczne obiekty na zdjęciach niż w przypadku obrazu zawierającego szersze ujęcie.

W przypadku badania funkcjonalności rozpoznawania tekstu na zdjęciach, Google także okazał się gorszy od swojej konkurencji i wskazał jedynie podobne obrazy oraz informację o miejscu, z którym może być powiązane zdjęcie. Bing, pomimo posiadania opcji rozpoznawania tekstu na obrazie, nie był w stanie podać ani informacji o tym jaki napis znajduje się na grafice użytej w badaniu (zdjęcie przedstawiało tablicę z napisem po tajsku), ani dodatkowych informacji, które mogłyby wskazywać na miejsce wykonania zdjęcia. Tutaj jedynymi podanymi wynikami były podobne graficznie obrazy. W tym zakresie najlepiej poradził sobie Yandex, który nie tylko rozpoznał tekst na zdjęciu i zaproponował możliwość jego przetłumaczenia, ale także wskazał podobne wizualnie grafiki i strony, zawierające podobne obrazy, co może naprowadzić osoby wyszukujące na dodatkowe źródła informacji, związanych z daną grafiką.

[73] K. Wosiński, *Jak wyszukiwarki radzą sobie z analizą zawartości obrazów*, https://sekurak.pl/jak-wyszukiwarki-radza-sobie-z-analiza-zawartosci-obrazow-osint-hints/ – dostęp online 22.08.2022 r.
[74] K. Wosiński, *Comparison of reverse image searching in popular search engines*, https://research.securitum.com/comparison-of-reverse-image-searching-in-popular-search-engines-osint-hints/ – dostęp online 22.08.2022 r.
[75] https://github.com/dessant/search-by-image

Badanie zdolności wyszukiwarek do rozpoznawania samochodu na zdjęciu wykazało, że Google poprzez przypisanie słów kluczowych do wyszukiwanego obrazu (co jest standardowym działaniem tej wyszukiwarki i w zależności od sytuacji może polepszyć lub pogorszyć trafność wyszukiwania), rozszerzył zakres poszukiwań i wskazał wiele innych modeli samochodów, chociaż wśród wyników znalazły się także modele zgodne z wyszukiwanym. Bing w tej kategorii nie potrafił rozpoznać konkretnego modelu samochodu, a jedynie wskazał dość podobne wizualnie pojazdy. Yandex natomiast nie tylko poprawnie zidentyfikował model samochodu na zdjęciu, ale także standardowo wskazał podobne obrazy i linki do stron z informacjami o poszukiwanym modelu auta.

W przypadku analizy zdjęcia, zawierającego egzotyczne owoce, Google znów rozszerzył wyszukiwanie poprzez ogólne słowa kluczowe, jednak wśród wizualnie podobnych obrazów przeważały te, zawierające inne rodzaje owoców. Jedynie poprzez wskazane strony, zawierające podobne obrazy można było dotrzeć do informacji co to za owoce. W przypadku Binga sytuacja przedstawiała się lepiej, gdyż zarówno wskazane strony internetowe, jak i obrazy podobne wizualnie przedstawiały poprawnie tej sam rodzaj owoców co w obrazie, który posłużył jako bazowy. Podobnie Yandex wskazał wyniki z poprawnymi informacjami.

Kolejnym etapem badania była analiza japońskiego znaku graficznego, będącego logo jednej z firm kurierskich, nieznanych jednak szerzej poza Japonią. Zbadano tutaj logo w dwóch postaciach – oryginalnej grafiki, pobranej ze strony firmy oraz fragmentu zdjęcia, przedstawiającego logo pod niewielkim kątem, a także zdjęcie z logo Igrzysk Olimpijskich w Pekinie i logo sieci znanych polskich cukierni. O ile dla pierwszego przypadku wyszukiwarki nie miały problemu, o tyle dla zdjęcia zrobionego pod kątem oraz zdjęcia logo Igrzysk, Google wskazał podobne grafiki, które można było wykorzystać do dalszych poszukiwań, Bing zupełnie nie potrafił znaleźć podobnej grafiki, a Yandex wskazał poprawnie strony z poszukiwanym znakiem oraz obrazy podobne graficznie. Dla logo znanych polskich cukierni żadna z wyszukiwarek nie potrafiła znaleźć poprawnych wyników.

Ostatnim elementem badania była weryfikacja umiejętności rozpoznawania twarzy, w którym za materiał badawczy posłużyły zdjęcia znanych skoczków narciarskich, zmodyfikowane zdjęcie znanego biznesmena oraz zdjęcie dziewczyny, pobrane z portalu ze zdjęciami stockowymi. Najciekawszym przypadkiem okazały się

zdjęcia sportowców, których nie rozpoznały ani Google, ani Yandex. Bing natomiast wskazał konkretnie imię i nazwisko obu skoczków (ich zdjęcia były analizowane osobno) pomimo, że na pewno tego zdjęcia nie mógł wcześniej „widzieć", gdyż było to zdjęcie, które nie zostało nigdy opublikowane w Internecie. Oznacza to, że sposób analizy twarzy przez wyszukiwarkę Bing stoi na naprawdę wysokim poziomie. Dla zmodyfikowanego (w zakresie konwersji na odcienie szarości i znaczne podniesienie kontrastu, a także obrót) zdjęcia biznesmena, Google także nie był w stanie wskazać osoby na zdjęciu, natomiast Bing i Yandex poprawnie wskazały osobę, która znajdowała się na zdjęciu. Stockowe zdjęcie dziewczyny przez Bing i Yandex poprawnie zidentyfikowane, co spowodowało wskazanie profili w mediach społecznościowych, wykorzystujących ten obraz. Google w tym zakresie niestety wskazał zupełnie niepowiązane wyniki.

Porównanie wyników badań przedstawionych powyżej pozwala stwierdzić, że w każdym z jego etapów wyszukiwarka Yandex radzi sobie co najmniej prawidłowo, a w niektórych jest dużo lepsza od konkurencji. Google i Bing uzupełniają się pod względem wyszukiwania obrazami w różnych kategoriach, jednak ten drugi wyraźnie lepiej radzi sobie z rozpoznawaniem twarzy. Wyniki w formie tabeli z rezultatami poszczególnych sprawdzeń przedstawiono na Rys. 8.

	Google	Bing	Yandex
Miejsca	+	-	+
Tekst	-	+	++
Samochody	+	-	++
Owoce	-	+	+
Logo	+	-	+
Twarze	-	++	+

Rys. 8 – Wyniki badania wyszukiwarek Google, Bing i Yandex w zakresie wyszukiwania obrazem. Źródło: https://sekurak.pl/jak-wyszukiwarki-radza-sobie-z-analiza-zawartosci-obrazow-osint-hints – dostęp online 31.08.2022 r.

Rozpoznawanie twarzy jest technologią, która jest coraz bardziej pożądana, szczególnie przez organy ścigania w celu identyfikacji sprawców przestępstw, a także przez służby i środowiska akademickie. Narzędziem, które wyróżnia się w tym obszarze jest PimEyes – wyszukiwarka twarzy, której twórcami jest dwóch absolwentów Politechniki Wrocławskiej – Łukasz Kowalczyk i Denis Tatina. Z jej pomocą możliwe jest wyszukanie wszystkich obrazów zawierających twarz należącą najprawdopodobniej do tej samej osoby, którą przedstawia zdjęcie źródłowe. Aby jednak otrzymać informację z linkiem do miejsca w Internecie, w którym dane zdjęcie zostało znalezione, konieczne jest wykupienie subskrypcji. Algorytmy PimEyes wykorzystują sieci neuronowe do tworzenia wzorców twarzy w celu dopasowania ich do wzorców obecnych już w bazie danych, a także do dalszego „uczenia się" w celu jeszcze lepszego dopasowywania obrazów. Są one tak dobre, że potrafią znaleźć zdjęcia osób wykonane przed laty lub nawet dopasować zdjęcie w maseczce na twarzy lub w okularach słonecznych, przez co to narzędzie stało się obiektem oskarżeń o pomoc nie tylko osobom działającym w dobrej sprawie, ale także przestępcom śledzącym swoje ofiary. Co prawda przed rozpoczęciem wyszukiwania należy potwierdzić, że zdjęcie, za pomocą którego wyszukujemy przedstawia naszą osobę i taka sama zasada opisana jest w regulaminie, lecz aktualni właściciele[76] liczą raczej na etyczne postępowanie swoich użytkowników i nie wymuszają go w żaden dodatkowy sposób.

W ramach wyszukiwania obrazów, można także wyróżnić kategorię analizy avatarów kont osób i organizacji w Internecie, na podstawie powiązanego adresu e-mail, co zostało szerzej opisane w podrozdziale 1.13.1.7.

1.13.1.3. Wyszukiwanie IoT

Rekonesans urządzeń podłączonych do Internetu, niezależnie od poziomu ich zaawansowania technologicznego ani funkcji, którą pełnią, możliwy jest dzięki wyszukiwarkom, które skanują publicznie dostępną sieć w poszukiwaniu otwartych portów i działających na nich usług. Dzięki temu późniejsze korzystanie z tego typu narzędzi jest skanowaniem typowo pasywnym, gdyż analizowane są wcześniej zebrane

[76] W 2020 roku firma przeniosła się na Seszele, a w 2021 roku kupił ją Giorgi Gobronidze – gruziński profesor, który zarejestrował firmę w Dubaju, a siedzibę umieścił w Tbilisi. Gobronidze poznał twórców w 2017 roku, kiedy był na wymianie akademickiej w Polsce i od tego czasu śledził ich poczynania będąc pod wrażeniem algorytmów przez nich opracowanych.

dane, a nie informacje pozyskane w czasie rzeczywistym. Mankamentem takiego rozwiązania jest oczywiście brak aktualności uzyskanych wyników wyszukiwania, jednak zazwyczaj zmiany w konfiguracji i dostępności urządzeń w Internecie nie są na tyle dynamiczne, aby stanowiło to poważny argument przeciwko możliwości pasywnego rozpoznania infrastruktury.

Przykładem wyszukiwarki urządzeń w Internecie jest Shodan[77] – narzędzie, które umożliwia każdemu użytkownikowi Internetu wyszukanie urządzeń w oparciu o wprowadzone słowa kluczowe. W przypadku chęci skorzystania z zaawansowanych filtrów wyszukiwania (np. dotyczących podatności), możliwości podejrzenia lokalizacji znalezionych urządzeń na mapie czy też zwiększenia limitu wyszukiwanych lub monitorowanych adresów IP, konieczne jest wniesienie jednorazowej opłaty za podstawowe konto lub wykupienie jednej z oferowanych subskrypcji (w zależności od potrzeb ilościowych). Wyszukiwane frazy są poszukiwane w bannerach, zwracanych przez oprogramowanie działające na skanowanych portach, więc odpowiednie skonstruowanie tych fraz jest w stanie wyłuskać konkretne urządzenia z konkretnym rodzajem oprogramowania, a także np. wskazać jak wygląda interfejs graficzny, dostępny na danym porcie. Shodan skanuje wszystkie rodzaje urządzeń, które są dostępne w Internecie pod publicznymi adresami IP, co skutkuje tym, że możliwe jest wyszukanie informacji nie tylko o typowych serwerach www, ale także o kamerach z wystawionym do Internetu interfejsem www, komputerach z udostępnionym zdalnym pulpitem czy nawet sterownikach przemysłowych, kontrolujących zarówno przydomowe instalacje, jak i ogromne systemy np. w elektrowniach czy oczyszczalniach.

Bardzo podobny zestaw funkcjonalności oferuje chiński portal ZoomEye[78]. Jego dodatkowymi atutami są możliwość podejrzenia podatności powiązanych z wyszukanymi usługami, a także przeszukiwanie nie tylko bannerów zwracanych przez urządzenia, ale także tekstu stron internetowych. Ze względu na tę ostatnią właściwość możliwe jest np. wyszukiwanie w treści stron internetowych linków do konkretnych avatarów (zdjęć profilowych), hostowanych w serwisie gravatar.com, co pozwala na wyszukiwanie, gdzie osoby powiązane z wyszukiwanymi avatarami pozostawiały swoje ślady, np. wpisy na forum (szerzej o powiązaniu adresów e-mail ze zdjęciami profilowym opisano w podrozdziale 1.13.1.7). Tego typu rekonesans nie jest możliwy z wykorzystaniem

[77] https://www.shodan.io
[78] https://www.zoomeye.org

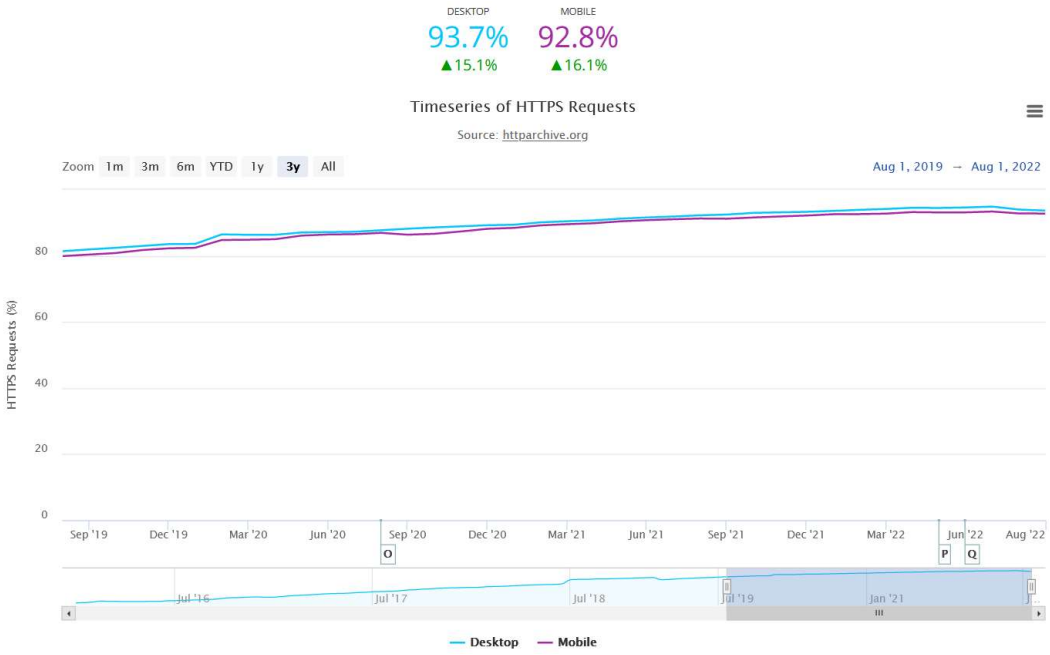
zwykłych wyszukiwarek (jak Google czy Bing), gdyż wyszukują one informacje jedynie w tekstowej reprezentacji stron, a nie w ich kodzie źródłowym.

Innym, często porównywanym z Shodanem portalem, pozwalającym na wyszukiwanie informacji o urządzeniach dostępnych z poziomu Internetu jest Censys[79]. Jego opcje wyszukiwania umożliwiają korzystanie z tagów bez konieczności wykupywania odpowiedniego pakietu, jak to ma miejsce w przypadku Shodana, a jedynie wymagają założenia konta użytkownika. Poprzez system podpowiedzi korzystanie z wyszukiwarki hostów jest uproszczone, a ilość wyników w porównaniu do serwisu Shodan potrafi być podobna lub znacznie się różnić, zarówno w górę, jak i w dół, co spowodowane jest innymi algorytmami, jakimi posługuje się wyszukiwarka (tak samo jak ma to miejsce w przypadku wyszukiwarek typu Google czy Bing). Oprócz hostów, Censys udostępnia także możliwość wyszukiwania certyfikatów, która szerzej została opisana w kolejnym podrozdziale.

1.13.1.4. Wyszukiwanie informacji na podstawie certyfikatu

Certyfikaty stały się obecnie standardem, jeśli chodzi o zabezpieczenie komunikacji pomiędzy przeglądarką użytkownika Internetu a serwerem hostującym stronę internetową. Według raportu HTTP Archive: State of the Web[80], w ciągu ostatnich 3 lat liczba analizowanych stron, które dostępne są przez połączenie HTTPS (czyli z wykorzystaniem zabezpieczonego połączenia, do którego wymagany certyfikat jest wystawiony dla witryny) wzrosła o 15,1% w przypadku wersji desktopowych i o 16,1% w przypadku wersji mobilnych, i wynosiła na dzień 1 sierpnia 2022 roku odpowiednio 93,7% i 92,8%. Dane o adresach stron do badania pobrane zostały z listy adresów udostępnianych przez Google w ramach danych zbieranych od użytkowników przeglądarki Chrome. Wykres zmian w procentowej ilości stron obsługujących protokół HTTPS przedstawiono na Rys. 9.

[79] https://search.censys.io
[80] *Report: State of the Web*, https://httparchive.org/reports/state-of-the-web – dostęp online 31.08.2022 r.



Rys. 9 – Wykres procentowej ilości stron obsługujących protokół HTTPS.
Źródło: State of the Web, https://httparchive.org/reports/state-of-the-web – dostęp online 31.08.2022 r.

Biorąc powyższe pod uwagę, można przyjąć, że certyfikaty wystawione dla serwisów operujących w Internecie mogą być źródłem wiedzy o przeważającej ilości jego zasobów. Stąd też bardzo ważny zasób informacyjny stanowią serwisy, pozwalające na wyszukiwanie informacji o aktualnych oraz historycznych certyfikatach stron internetowych. Informacje o wydanych certyfikatach są publicznie dostępne ze względu na standard Certificate Transparency (CT), stosowany powszechnie od 2018 roku, kiedy to opublikowana została jego wersja 2.0. Poprzez system publicznych logów dla wydawanych certyfikatów, możliwość szybkiego znalezienia certyfikatów wystawionych błędnie lub w złośliwych celach, została znacznie ułatwiona. Pozwala to także na zdobywanie informacji o infrastrukturze, ponieważ certyfikaty wystawione dla serwisów (np. dla subdomen używanych do celów testowych), o których informacja nie została upubliczniona, będą dostępne do wyszukania w logach. Innym rodzajem danych, dostępnych poprzez certyfikaty są informacje o osobach, uczestniczących w procesie wydawania certyfikatów z ramienia danej organizacji (np. poprzez uwzględnienie ich adresu e-mail lub nazwiska) i pozwalają na zbudowanie dodatkowej wiedzy o niej.

Wyszukiwarką, pozwalającą na znalezienie certyfikatów, a co za tym idzie informacji, dotyczących danej firmy lub domeny, jest m.in. wspomniany w poprzednim

podrozdziale Censys, który nie tylko korzysta z publicznych logów CT, ale także zbiera informacje o certyfikatach, które otrzymuje podczas nawiązywania połączenia ze skanowanymi przez siebie hostami w Internecie.

Inną wyszukiwarką certyfikatów na podstawie domeny, nazwy firmy lub unikalnego znacznika certyfikatu (tzw. *fingerprint*) jest portal crt.sh[81], który podobnie jak Censys oferuje bardzo przejrzysty interfejs i umożliwia analizę zarówno aktualnych, jak i wygasłych certyfikatów oraz szybki podgląd listy subdomen, dla których generowane były certyfikaty.

1.13.1.5. Wyszukiwanie informacji o organizacji na podstawie analizowanych plików, adresów URL i domen

Źródłem informacji o organizacji, jej infrastrukturze oraz technologii może być serwis VirusTotal, którego podstawową funkcją jest analizowanie plików i adresów URL pod kątem złośliwego oprogramowania. Mało znaną jednak jego funkcją jest także możliwość wyszukiwania w zgromadzonych zasobach informacji o organizacjach. Jest to możliwe, gdyż wszystkie próbki przesyłane do serwisu są pozostawiane w jego bazie, czego wiele osób korzystających z VirusTotal nie jest świadomych. Skutkuje to zatem w wielu przypadkach publikowaniem danych, które nie powinny zostać upublicznione. Poprzez wyszukiwanie informacji na podstawie podanej nazwy lub domeny w VirusTotal, możliwe jest uzyskanie danych zawartych w plikach, wytworzonych lub dotyczących danej organizacji, certyfikatów wystawionych dla jej domen, a także danych o serwerach hostujących strony internetowe w analizowanej domenie oraz powiązań pomiędzy wymienionymi informacjami. Do pobrania pliku, który wcześniej został wgrany w celu jego analizy wymagane jest posiadanie płatnego konta w VirusTotal, co ogranicza w pewien sposób możliwość dotarcia do zawartości wgranych do analizy plików, jednak nie czyni tego zadania niemożliwym. W przypadku omyłkowego wgrania pliku lub gdy wgrany plik zawiera poufne dane, możliwy jest kontakt z obsługą serwisu w celu jego usunięcia.

[81] https://crt.sh

1.13.1.6. Wyszukiwarki kont w serwisach internetowych

Jako osobną kategorię można przyjąć wyszukiwanie informacji o kontach użytkowników w serwisach internetowych, w tym także w serwisach mediów społecznościowych. Funkcjonalność taką udostępnia wiele narzędzi OSINT-owych, z których poniżej opisanych zostało kilka przykładowych, oferujących różne zestawy funkcjonalności.

Narzędzie Whatsmyname[82] powstało jako skrypt napisany w języku Python i umożliwiało z poziomu linii komend wyszukanie w których serwisach internetowych konto użytkownika o danym loginie będzie znalezione. Idea działania narzędzia opiera się na weryfikacji odpowiedzi serwera dla kont istniejących i nieistniejących. Jeśli możliwe są do wychwycenia różnice w stronach, wyświetlanych przy próbie uzyskania dostępu do strony z profilem istniejącego i nieistniejącego użytkownika, możliwe jest także wskazanie tylko tych serwisów, na których dany użytkownik ma konto. Z czasem lista serwisów, która jest tworzona przez osoby z wielu krajów (w tym przez autora niniejszej pracy oraz jeszcze jedną osobę z Polski, dzięki której lista możliwych do analizy polskich serwisów została znacznie poszerzona), stawała się coraz większa. Aktualnie narzędzie posiada jakże wersję przeglądarkową, dostępną pod adresem https://whatsmyname.app, która umożliwia wyszukiwanie w aktualnej bazie serwisów z poziomu strony internetowej, bez konieczności używania skryptów pythonowych. Warto jednak podkreślić, że samo znalezienie kont, należących do osoby o podanym loginie nie oznacza ich powiązania z osobą, która jest przedmiotem rekonesansu, gdyż może się okazać, że wiele osób używa tej samej nazwy użytkownika. Im ta nazwa jest bardziej oryginalna, tym prawdopodobieństwo uzyskania trafniejszych wyników jest większe, jednak za każdym razem konieczna jest ręczna weryfikacja uzyskanych wyników. Ułatwieniem dla tego typu działań może być inne narzędzie – WMN_screenshooter[83], które po wskazaniu pliku z listą serwisów, na której bazuje Whatsmyname, jest w stanie wykonać wrzuty ekranu dla wszystkich zidentyfikowanych stron z kontami należącymi do użytkownika o podanej nazwie.

Bardzo podobnym narzędziem, pozwalającym wyszukanie kont w mediach społecznościowych na podstawie podanej nazwy użytkownika jest Sherlock[84], który także

[82] https://github.com/WebBreacher/WhatsMyName
[83] https://github.com/swedishmike/WMN_screenshooter
[84] https://github.com/sherlock-project/sherlock

jest narzędziem napisanym w języku Python. Interesującą funkcją tego narzędzia jest możliwość wyszukiwania poprzez sieć Tor lub przez serwer pośredniczący (*proxy*), co zmniejsza niebezpieczeństwo identyfikacji osoby wyszukującej.

Jeśli danymi wejściowymi do poszukiwań osoby nie jest jednak nazwa użytkownika, a jej imię i nazwisko, możliwe jest skorzystanie z narzędzia NAMINT[85], opracowanego przez autora niniejszej pracy, które pozwala na stworzenie listy możliwych kombinacji zapisu imienia i nazwiska (bądź pierwszych liter jednego z nich), które następnie mogą posłużyć m.in. jako baza do wyszukiwania w Google (w wyszukiwarce tekstowej, grafiki, mapach, dokumentach), wyszukiwarkach Bing i Yandex, a także weryfikację istnienia kont osób o wprowadzonych danych w wybranych mediach społecznościowych (m.in. Facebook, Twitter, TikTok). Utworzone z podanego imienia i nazwiska prawdopodobne kombinacje tworzące loginy, stanowią także bazę do wyszukania w wyszukiwarkach adresów email, a także profili o danych nazwach użytkownika w mediach społecznościowych. Dodatkowo, na podstawie wygenerowanych loginów narzędzie wyświetla zdjęcia profilowe, powiązane z adresami e-mail o podanych loginach i w podanych domenach, pod warunkiem, że są powiązane z kontami w serwisie Gravatar. Mnogość kombinacji, stworzonych na podstawie podanego imienia (lub imion) i nazwiska daje możliwość szerokiego wyszukania osoby o podanych danych.

Jeśli jako dane wejściowe posłuży adres e-mail, powiązany z kontem Google, użytecznym do identyfikacji dodatkowych informacji o takiej osobie będzie wyszukiwarka firmy Epieos[86]. Wynikiem jej działania są informacje, takie jak: imię i nazwisko, zdjęcie profilowe, identyfikator konta, a także datę ostatniej aktualizacji konta czy odnośnik do map Google, które mogą zawierać recenzje, pozostawione przez danego użytkownika w różnych lokalizacjach. na podstawie recenzji możliwe jest określenie skąd prawdopodobnie dana osoba pochodzi – dużo wpisów w jednej okolicy może wskazywać na częste przebywanie tam, jednak należy do tego podchodzić z dużą ostrożnością, ponieważ wpisy mogą być sfabrykowane lub dana osoba może zostawiać je tylko w danym miejscu, pomimo, że wcale tam na co dzień nie przebywa.

[85] https://seintpl.github.io/NAMINT
[86] https://epieos.com

1.13.1.7. Wyszukiwanie zdjęć profilowych (avatarów)

Powyżej opisane narzędzie firmy Epieos umożliwia uzyskanie zdjęcia profilowego konta Google na podstawie powiązanego adresu e-mail, przy czym nie musi to być adres w domenie gmail.com. Nie jest to jednak jedyna metoda, gdyż usługi Google, takie jak Gmail, Dokumenty i Formularze Google czy czat Hangouts, umożliwiają podgląd zdjęć profilowych osób, które zostaną dodane jako adresaci e-maila oraz które są wpisywane na listę osób do udostępnienia dokumentu lub formularza. Jest to metoda pasywna, gdyż nie jest konieczne zatwierdzenie wysyłania wiadomości lub udostępniania dokumentu do tego, aby uzyskać od Google podpowiedź ze zdjęciem profilowym osoby, której adres wpisujemy w odpowiednie pole.

Ten sam mechanizm działa dla kont pocztowych w Microsoftowych kontach w domenie outlook.com. Poprzez dodanie nowego adresata maila z konta Outlook lub osoby do wydarzenia w Kalendarzu, udostępniane jest zdjęcie profilowe osoby, do której należy wpisany adres e-mail (pod warunkiem oczywiście, że dla ich konta Outlook zostało takie zdjęcie ustawione). W przeciwieństwie do adresów powiązanych z kontami Google, ważna jest tu ilość i umiejscowienie kropek w loginie adresu e-mail. Dla Google nie ma to znaczenia, zarówno jeśli chodzi o dostarczanie poczty, jak i wyszukiwanie profili.

Podana w rozdziale 1.13.1.6, dotyczącym wyszukiwania kont w serwisach internetowych, możliwość powiązania zdjęć profilowych użytkowników z ich adresem e-mail poprzez serwis Gravatar[87], wykorzystuje bardzo prostą metodę tworzenia bezpośrednich odnośników do tych zdjęć. Serwis Gravatar udostępnia funkcjonalność stworzenia konta, zawierającego zdjęcie profilowe, powiązane z adresem e-mail właściciela poprzez stworzenie stałego odnośnika, gdzie odwołanie do konkretnego zdjęcia profilowego bazuje na skrócie MD5 właśnie z adresu e-mail. Zdjęcia te są szeroko wykorzystywane w Internecie, np. na forach lub w komentarzach, ponieważ podanie adresu e-mail przez osobę dokonującą wpisu umożliwia wyświetlenie także jej zdjęcia profilowego obok wypowiedzi. Tak więc znając więc adres e-mail danej osoby można bez problemu sprawdzić czy i jakie zdjęcie profilowe ma ustawione w serwisie Gravatar.

[87] https://gravatar.com

1.13.1.8. Inne narzędzia

Jednym ze sposobów na wykonywanie rozpoznania z jednoczesną archiwizacją danych i przedstawianiem ich w formie graficznej, jest oprogramowanie Maltego[88], dostępne w wielu konfiguracjach – także w wersji darmowej (Community Edition – CE), która pozwala na ograniczone, ale nadal dość szerokie skorzystanie z posiadanych opcji rekonesansu. Ograniczenia dla wersji CE obejmują m.in. brak możliwości wykorzystywania jej w celach komercyjnych oraz możliwość zwracania jedynie do 12 wyników na transformację. Maltego działa właśnie w oparciu o tzw. transformacje (ang. *transforms*), czyli dedykowane skrypty, które poprzez określone wyszukiwania na podstawie zadanych danych wejściowych, wprowadzają na budowany graf zależności prowadzące do znalezionych danych. Przykładem użycia transformacji jest wyszukanie adresów email w zadanej domenie na podstawie wyników z WHOIS, czyli bazy danych serwerów DNS, zawierającej m.in. informacje o właścicielach domen i dane kontaktowe osób odpowiedzialnych za ich obsługę. Dzięki wielu dostępnym do zainstalowania transformacjom, przygotowanym zarówno przez zespół odpowiedzialny za oprogramowanie Maltego, ale także przez inne firmy i osoby, wykorzystującym nie tylko możliwości wyszukiwarek Google czy Bing, ale także takich narzędzi jak Shodan lub VirusTotal oraz opcję wyszukiwania w portalach społecznościowych (np. Twitter lub LinkedIn), możliwości analizy danych i ich powiązań z wykorzystaniem Maltego są ogromne. Istnieje także możliwość własnego przygotowania transformacji i wykorzystania ich w oprogramowaniu, więc w przypadku szczególnych wymagań śledczych, może być to duże ułatwienie w prowadzenia rozpoznania OSINT-owego. Pomimo faktu, że same wyszukiwania wykonywane z Maltego są pasywne, tzn. nie powodują pozostawiania śladów w rozpoznawanej infrastrukturze, to jednak zasada działania wbudowanych transformacji opiera się na wysyłaniu zapytań do serwera transformacji Maltego. Użytkownicy, którzy ze względu na wysoką poufność prowadzonych działań śledczych nie chcą, aby jakiekolwiek dane były logowane przez producenta Maltego, muszą korzystać z lokalnej kopii serwera transformacji. Istnieje także możliwość jeszcze większego ukrycia prowadzonych działań, poprzez włączenie w ustawieniach prywatności opcji „niewidoczny” (ang. *stealth*), co spowoduje

[88] https://www.maltego.com

zablokowanie pobierania danych z Internetu, tzn. brak automatycznego pobierania m.in. ikon stron internetowych i obrazów dla elementów grafu powiązań.

Bardzo duże możliwości rekonesansu pasywnego[89] daje także oprogramowanie Recon-ng[90], które jest darmowym frameworkiem[91], służącym do wykonywania OSINT-owego wyszukiwania danych o organizacjach, osobach czy elementach (w tym podatnościach) systemów teleinformatycznych poprzez wykorzystanie wielu modułów, które mogą być instalowane przez użytkownika z istniejącego repozytorium narzędzi (tzw. marketplace) lub tworzone indywidualnie. Jako dane wejściowe, możliwe jest podanie np. nazwy organizacji, danych osobowych lub domeny. Poprzez wykorzystanie modułów, z których każdy korzysta z innego źródła informacji, możliwe jest uzupełnianie swojej bazy danych dotyczącej danego rozpoznania. Moduły podzielone są na grupy, ze względu na rodzaj danych wejściowych oraz wyjściowych, co ułatwia znacznie wyszukiwanie odpowiedniego modułu do pożądanego wyszukiwania. Bazując na już zdobytych informacjach (np. liście adresów IP zdobytych na podstawie nazw domenowych), możliwe jest ich przekazanie jako danych wejściowych do wykorzystania przez kolejny moduł. Tym sposobem każde nowo zdobyte informacje rozszerzają możliwości dalszego wyszukiwania informacji o celu. Dodatkowo, możliwe jest generowanie raportów z przeprowadzonego rekonesansu.

W przypadku rekonesansu na podstawie nazwy domenowej, możliwe jest wykorzystanie narzędzia theHarvester[92], które zbiera informacje dostępne z wykorzystaniem rekonesansu OSINT-owego i umożliwia zbadanie ekspozycji danych powiązanych z daną domeną w Internecie. Daje to pogląd na widoczne adresy IP, e-mail, URL, a także subdomeny, nazwiska i subdomeny. Wykorzystuje do tego różne inne narzędzia i wyszukiwarki, co powoduje duży zakres możliwych sprawdzeń. Większość z wykorzystywanych narzędzi działa pasywnie, jednak istnieje także możliwość użycia technik aktywnych, jak DNS brute-force czy wykonywanie zrzutów ekranów dla zidentyfikowanych subdomen.

[89] Większość modułów Recon-ng działa pasywnie, jednak niektóre moduły skanują sieci i hosty aktywnie, o czym zawsze należy pamiętać podczas wykonywania rekonesansu.
[90] https://github.com/lanmaster53/recon-ng – dostęp online 22.08.2022 r.
[91] Framework – platforma programistyczna, będąca szkieletem, na którym może zostać zbudowana aplikacja w oparciu o strukturę i sposób działania, zdefiniowane w samym frameworku.
[92] https://github.com/laramies/theHarvester

Narzędziem, które także może zostać wykorzystane do pasywnego rekonesansu jest Wireshark[93]. To znane oprogramowanie, służące do analizy ruchu sieciowego może dać dużo informacji w przypadku możliwości podłączenia się do sieci firmowej lub podsłuchiwania ruchu sieciowego, generowanego przez komputery pracowników, podłączonych np. do otwartej sieci Wi-Fi. Wspomniany jednak wcześniej wzrost połączeń szyfrowanych, a także coraz szersze wykorzystywanie szyfrowania zapytań DNS może jednak stanowić dużą przeszkodę w praktycznym zdobywaniu wiedzy na podstawie przechwyconego ruchu sieciowego.

1.13.2. Rekonesans aktywny

W ramach rekonesansu aktywnego, głównymi technikami są metody skanowania sieci (nie tylko Internetu, ale także sieci wewnętrznych) w celu uzyskania jak największej ilości informacji o danym celu. Należy jednak pamiętać, że wchodzenie w interakcje z badaną infrastrukturą może zostać uznane za cyber-atak, dlatego zawsze należy mieć pewność co do legalności wykonywanych działań (np. poprzez podpisanie umowy z klientem, dla którego wykonywane jest rozpoznanie podatności jego infrastruktury, zawierającej dokładny zakres możliwych działań). Przykładami sposobów rekonesansu aktywnego mogą być:

- skanowanie portów / usług,
- enumeracja subdomen,
- enumeracja SMTP
- nietechniczne rodzaje aktywnego rozpoznania.

Sposoby te opisano w kolejnych podrozdziałach.

[93] https://www.wireshark.org

1.13.2.1. Skanowanie portów / usług

Często wykonywane w ramach testów bezpieczeństwa w celu odkrycia jakie porty i pracujące na nich usługi są dostępne dla każdego użytkownika sieci. Poprzez analizę odpowiedzi serwerów (np. na podstawie kodów odpowiedzi lub bannerów tekstowych[94]), możliwe jest rozpoznanie szczegółów dotyczących technologii, w jakiej zbudowana jest dana infrastruktura. Jednym z najczęściej używanych i najbardziej wszechstronnych narzędzi do tego typu działań jest Nmap[95], który oprócz wielu sposobów skanowania sieci, umożliwia także opcje ukrywające proces skanowania poprzez używanie odpowiednich pakietów i rozłożenie skanowania w czasie, jednak nie daje to pewności, że skanowanie nie zostanie wykryte przez systemy zabezpieczeń w badanej infrastrukturze. Poprzez tę technikę możliwe jest do ustalenia np. jakie systemy operacyjne są używane przez daną organizację (tzw. *OS fingerprinting* – w dosłownym tłumaczeniu: zbieranie odcisków palców systemu operacyjnego) lub nawet jakie podatności są w nich obecne.

1.13.2.2. Enumeracja subdomen

Ten rodzaj rekonesansu ma na celu wydobycie informacji z serwerów DNS, które odpowiedzialne są za zmianę nazw domenowych na adresację IP. Możliwe jest zatem zdobycie informacji zarówno o subdomenach w ramach danego adresu, które mogą posłużyć do rozszerzenia zakresu rekonesansu, jak i o zakresach adresów IP, wykorzystywanych w danej infrastrukturze lub organizacji. W przypadku wykrycia niepoprawnej konfiguracji serwera DNS, możliwe jest np. wykorzystanie podatności na atak typu *Zone transfer*, dający możliwość przechwycenia wszystkich informacji z serwera DNS, jednak tego typu działanie wykracza już poza zakres legalnego rekonesansu.

Analizę informacji związanych z domenami oraz rekordami DNS umożliwia bardzo wiele narzędzi, także darmowych, dostępnych do zainstalowania na komputerze lub w formie serwisu internetowego. W związku z faktem, że opisanie wszystkich tych aplikacji byłoby niemożliwe ze względu na ich mnogość, a także na fakt, że wiele z nich

[94] Technikę rozpoznawania usług na podstawie ich odpowiedzi tekstowych określa się mianem „Banner grabbing". Słowo „banner" oznacza w tym przypadku właśnie informację, która jest wyświetlana użytkownikowi przy próbie połączenia się do danego serwera.
[95] https://nmap.org

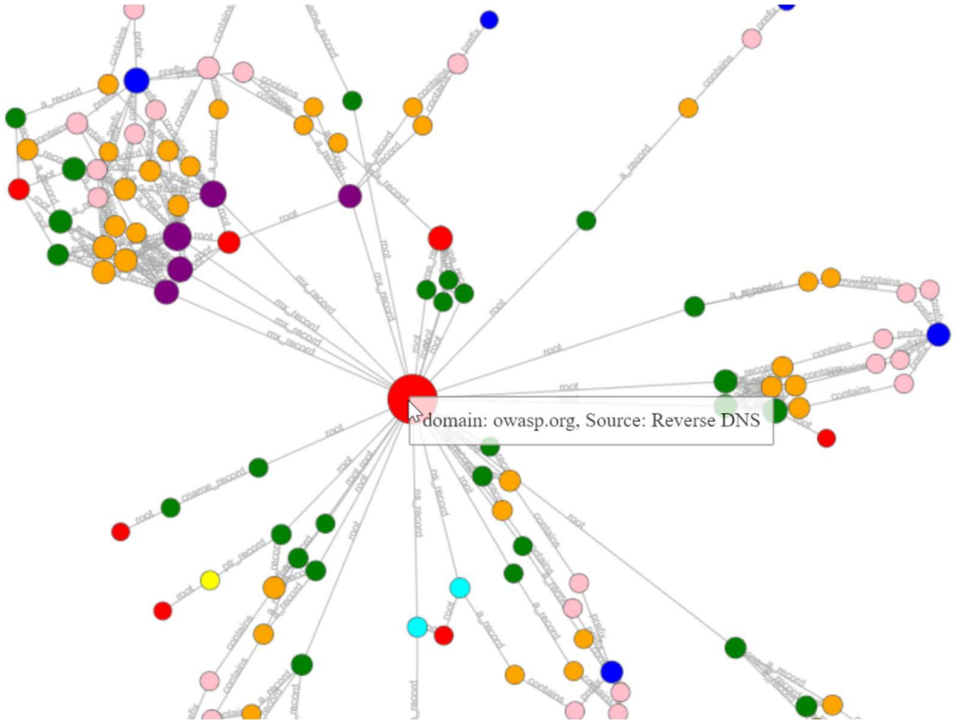
oferuje bardzo podobny zestaw funkcjonalności, w niniejszej pracy opisane zostanie tylko kilka reprezentatywnych zdaniem autora przykładów.

Serwisem, który oferuje prosty interfejs do wyszukiwania informacji związanych z analizowaną domeną internetową jest dnsdumpster.com - narzędzie stworzone przez zespół HackerTarget.com. Jak wskazują sami twórcy, narzędzie to ma pomagać zarówno atakującym pentesterom i poszukiwaczom *bug bounty*[96], jak i zespołom zorientowanym za ochronę sieci. OSINT prowadzony w odniesieniu do zasobów sieciowych ma im pomóc lepiej zarządzać swoimi sieciami. Siłą tego serwisu jest wiele źródeł pozyskiwania informacji - rekordy odczytywane z serwerów DNS, dane z wyszukiwarek, certyfikatów i repozytoriów danych.

W zakresie oprogramowania, które możliwe jest do uruchomienia na własnym komputerze, wymienić można takie narzędzia jak Sublist3r, Recon-ng (opisywany szerzej w ramach podpunktu 1.13.1.8) czy amass. Sublist3r[97] to napisany w języku Python skrypt, który umożliwia OSINT-ową enumerację subdomen. Wykorzystuje on do tego informacje pozyskane z takich wyszukiwarek jak Google, Yahoo, Bing, Baidu czy Ask, a także korzysta z takich źródeł danych jak Netcraft, Virustotal, ThreatCrowd, DNSdumpster czy ReverseDNS. Do narzędzia dołączony jest także moduł subbrute, który umożliwia aktywne poszukiwanie subdomen metodą brute-force przy wykorzystaniu słowników typowych nazw.

Z kolei narzędzie amass[98] zostało stworzone przez OWASP (Open Web Application Security Project[99]) do mapowania sieci oraz pozyskiwania informacji o jej zasobach przy wykorzystaniu ogólnodostępnych informacji, a także aktywnych form rekonesansu. Jego siłą są rozbudowane możliwości korzystania z wielu źródeł danych: API serwisów, certyfikatów, serwerów DNS, danych o routingu, zbiorów i archiwów danych, a także baz WHOIS. Pozyskane dane mogą być później bazą do ich dalszej analizy, m.in. przez ich zwizualizowanie w formie grafu zależności (pokazanego na Rys. 10) lub wyeksportowanie do narzędzia Maltego.

[96] bug bounty – programy poszukiwania błędów, publikowane przez firmy, oferujące korzyści w zamian za znalezienie uchybień w ich produktach (jak serwisy internetowe i aplikacje mobilne), głównie w zakresie problemów w zakresie zabezpieczeń i błędów funkcjonalnych.
[97] https://github.com/aboul3la/Sublist3r
[98] https://github.com/OWASP/Amass
[99] https://owasp.org



Rys. 10 – Przykładowy graf zależności pomiędzy poznanymi danymi o domenie.
Źródło: https://github.com/OWASP/Amass/blob/master/doc/tutorial.md – dostęp online 08.11.2022 r.

1.13.2.3. Enumeracja SMTP

Poprzez wykorzystanie specyfiki protokołu SMTP[100] i takich jego poleceń jak: VRFY (potwierdzenie, że podany adres e-mail istnieje), EXPN (informacja o liście mailingowej lub faktycznym odbiorcy e-maila w przypadku zapytania o alias) czy RCPT TO (lista odbiorców e-maila, które może także wskazać czy dany adres istnieje), możliwe jest uzyskanie informacji o kontach poczty elektronicznej w danej organizacji. Możliwe jest zatem sprawdzenie czy istnieją np. konta o nazwach „root" lub „admin", a także na jakie konta administracyjne są kierowane wiadomości wysyłane na adresy techniczne (np. „postmaster").

[100] SMTP (ang. *Simple Mail Transfer Protocol*) – protokół wykorzystywany do komunikacji za pomocą wiadomości poczty elektronicznej.

1.13.2.4. Nietechniczne rodzaje aktywnego rozpoznania

W ramach rozpoznania aktywnego, oprócz metod technicznego skanowania infrastruktury, wyróżnić można także metody nietechniczne, polegające na bezpośredniej interakcji z osobami, które są obiektem śledztwa lub są z nim związane. Wśród przykładów wykorzystania tego typu rekonesansu można zarówno wyróżnić dołączanie do grona znajomych w mediach społecznościowych, dołączanie do grup tematycznych lub lokalnych, wysyłanie wiadomości do badanych osób lub interakcja z nimi poprzez reakcję na ich wpisy, jak i techniki o większym stopniu ukrycia, np. kontakt na portalach typu LinkedIn jako rekruter, co zazwyczaj nie wzbudza podejrzeń ze względu na dużą ilość prób podobnych kontaktów, inicjowanych przez faktycznych rekruterów.

Niestety aktywne próby nawiązania kontaktu z analizowanymi osobami lub przedstawicielami rozpoznawanych organizacji, jest obarczone dużym ryzykiem wykrycia i może negatywnie odbić się na końcowym wyniku prowadzonego rozpoznania.

1.14. Ograniczenia etyczne

Podczas prowadzenia działań OSINT-owych, jednym z kluczowych aspektów jest legalność wykorzystywanych technik i narzędzi, jednak oprócz legalności regulowanej prawnie, należy mieć na uwadze także etykę działań.

Rozpoznanie, bazujące na otwartych źródłach, poprzez swoją powszechną dostępność dla każdego użytkownika Internetu, niesie także za sobą niebezpieczeństwa wynikające z różnorodności zahamowani etycznych poszczególnych osób. Brak odgórnego nadzoru może skutkować ujawnieniem nadmiarowych danych lub wyciągnięciem nieprawidłowych wniosków, co z kolei może prowadzić nawet do tego, że niesłusznie oskarżona osoba zacznie obawiać się o swoje życie, jak to miało miejsce w przypadku strażaka niesłusznie oskarżonego o śmiertelne zranienie policjanta podczas ataku na Kapitol 6 stycznia 2021 roku[101].

Niekiedy pomimo korzystania wyłącznie z otwartych źródeł, przekraczana jest granica poczucia bezpieczeństwa i komfortu przez osoby, co do których prowadzone jest rozpoznanie. Jak to zostało wspomniane na wstępie rozdziału 1.13 przy okazji omawiania zasad poszukiwania osób zaginionych, niekiedy wyznacza się także etyczne granice prowadzonego śledztwa, aby nie powodować dodatkowych emocji u osób związanych z poszukiwaną osobą.

Zdarzenia, wywołujące duże emocje, jak chociażby atak bombowy podczas maratonu w Bostonie w 2013 roku, powodują też duże poruszenie pośród społeczności internetowej, co z jednej strony ma pozytywne skutki, gdyż daje możliwość nakłonienia bardzo dużej ilości osób do podzielenia się swoimi zdjęciami, filmami oraz spostrzeżeniami na temat danego zdarzenia z organami ścigania i służbami. Z drugiej jednak strony, chęć szybkiego doprowadzenia do oskarżenia i osądzenia osób odpowiedzialnych, czyli ogólnie rzecz ujmując „wymierzenie sprawiedliwości", powoduje pomijanie etapu analizy zebranych dowodów, szybkie publikowanie poszlak jako wniosków oraz poddanie się efektowi podczepienia[102], co prowadzi z kolei do niesłusznych osądów. Po wspomnianym ataku w Bostonie wielu internautów wskazywało

[101] *Retired Chicago Firefighter Wrongly Accused of Participating in Capitol Violence*, https://www.firefighternation.com/leadership/retired-chicago-firefighter-wrongly-accused-of-participating-in-capitol-violence/ – dostęp online 01.09.2022 r.

[102] Efekt podczepienia lub zasad podczepienia – termin z zakresu psychologii poznawczej, który określa tendencję do wykonywania czegoś lub myślenia w określony sposób tylko dlatego, że inni tak robią lub myślą. Nazywany jest często także „owczym pędem".

niesłusznie osoby jako sprawców ataku tylko dlatego, że miały ze sobą plecak (w którym podobno ukryty był ładunek wybuchowy), patrzyły w innym kierunku niż reszta widzów, miały ciemniejszy kolor skóry lub były widziane jak uciekały (chociaż w momencie ataku dużo osób także uciekało, więc nie stanowiło to żadnej podstawy do oskarżeń). To wskazuje jak niebezpieczne może być umożliwienie rzeszy użytkowników Internetu prowadzenia śledztwa OSINT-owego bez nadzoru.

Duży dostęp do danych, na podstawie których możliwe jest rozpoznanie oraz coraz większa popularność śledztw internetowych sprawia, że zagrożone stają się operacje wojskowe, których powodzenie nierzadko zależy od ukrycia prowadzonych działań. Przykładem działania, którego autor nie miał złych intencji, a jednak przyczynił się do zagrożenia życia ludzkiego w rejonie walk była publikacja informacji o postoju ukraińskich żołnierzy pod centrum handlowym na przedmieściach Kijowa przez tamtejszego tiktokera w marcu 2022 roku. W ataku rakietowym, który nastąpił niedługo po publikacji tej informacji w mediach społecznościowych, zginęło 8 osób. Służba Bezpieczeństwa Ukrainy opublikowała nagranie[103], z przeprosinami tiktokera, który przyznaje się on do błędu i zaleca innym powściągliwość w publikowaniu tego typu treści. Wiele portali i profili w mediach społecznościowych (zarówno wojskowych, jak i cywilnych) po rozpoczęciu wojny w Ukrainie pod koniec lutego 2022 roku rozpoczęło kampanie informacyjne, mające na celu uświadomienie osobom publikującym treści w Internecie, że publikacja zdjęć, filmów czy chociażby informacji tekstowych o ruchach lub zasobach wojsk, może przyczynić się do znaczącego obniżenia bezpieczeństwa i skuteczności prowadzonych operacji, a także mieć wpływ na bezpieczeństwo ludności cywilnej i samych żołnierzy.

Powszechność dostępu do otwartych źródeł informacji nie jest jednak jedynie negatywnym zjawiskiem w zakresie informacji o działaniach wojskowych. To dzięki prywatnym osobom, zaangażowanym w śledzenie ogólnodostępnych informacji, możliwe było uświadomienie osobom na całym świecie, jak wygląda sytuacja w rejonach walk oraz jakie mogą być kolejne zamiary wojsk rosyjskich. Co więcej, dane wskazujące na plany inwazji na Ukrainę zostały zauważone zanim ona się jeszcze de facto zaczęła. Tego typu rozpoznanie było wcześniej zarezerwowane raczej dla instytucji wywiadowczych, a nie dla przeciętnych obywateli. Jednak to obserwacje map Google i korków, tworzących się na drodze łączącej rosyjskie miasto Biełgorod i ukraiński Charków, doprowadziły

[103] https://www.facebook.com/SecurSerUkraine/videos/3153483434931349/ – dostęp online 14.09.2022 r.

badaczy z Middlebury Institute of International Studies w Kalifornii do wniosku, że wcześniej zauważone w tej okolicy rosyjskie pojazdy militarne zatarasowały drogę samochodom prywatnych osób, których telefony informowały Google o dłuższym czasie przejazdu w tym rejonie[104].

W raporcie „Feeling the Burden. Ethical Challenges and Practices in Open Source Analysis and Journalism"[105], zawarto między innymi wyniki wywiadów z 20 analitykami i 8 dziennikarzami, zapytanych o momenty, kiedy mieli wątpliwości dotyczące tego, czy powinni publikować wyniki swojego rozpoznania, bazującego na otwartych źródłach. Ich odpowiedzi można podzielić na cztery główne kategorie problemów etycznych, z którymi musieli się zmierzyć. Dotyczyły one:

- dyplomacji, kryzysów i konfliktów,
- niezamierzonych konsekwencji swoich działań,
- prywatności,
- interakcji pomiędzy analitykami i dziennikarzami.

W pierwszej kategorii, dotyczącej dyplomacji, kryzysów i konfliktów, analitycy i dziennikarze wskazywali momenty, w których mieli wątpliwości czy poprzez publikację wyników swojej pracy nie utrudnią działań dyplomatycznych lub nie wpłyną na eskalację konfliktów. Jako jeden z przykładów podano sytuację, kiedy po irańskim ataku na amerykańską bazę wojskową, okazało się w ramach dziennikarskiego śledztwa, że atak był bardziej udany, niż oficjalnie podawano. Osoba, która dotarła do informacji o skali ataku wstrzymała się jednak z publikacją tych danych, aby nie spowodować publicznego napięcia i nacisku na wykonanie akcji odwetowej.

W kategorii dotyczącej niezamierzonych konsekwencji swoich działań, wskazane zostały wątpliwości, dotyczące możliwości przypadkowej pomocy, jaką opublikowane informacje mogłyby przynieść przestępcom lub wyrządzenia szkód działaniom wykonywanym w słusznej sprawie. Za jeden z przykładów posłużyła tutaj historia analizy zdjęć satelitarnych, na których widoczni byli terroryści działający w południowej Afryce.

[104] *How Open-Source Intelligence Is Helping Clear The Fog Of War In Ukraine* – https://www.buzzfeednews.com/article/peteraldhous/osint-ukraine-war-satellite-images-plane-tracking-social – dostęp online 14.09.2022 r.
[105] *Feeling the Burden. Ethical Challenges and Practices in Open Source Analysis and Journalism* – https://stanleycenter.org/wp-content/uploads/2022/01/NWRPT-FeelingtheBurden_122-v2.pdf – dostęp online 14.09.2022 r.

W ramach tego samego rozpoznania przypadkowo odkryte zostały ruchy wojsk amerykańskich i sojuszniczych. Informacje te, po konsultacjach, nie zostały opublikowane, co zapobiegło ujawnieniu operacji wojskowej.

W kategorii dotyczącej prywatności głównym aspektem, na który zwrócono uwagę, jest zapewnienie bezpieczeństwa źródłom informacji, podmiotom śledztw, a także osobom postronnym. Podczas publikowania zdjęć lub filmów, możliwe jest w wielu przypadkach zidentyfikowanie z jakiego dokładnie miejsca dane ujęcie zostało wykonane. Stwarza to niebezpieczeństwo zidentyfikowania autora materiału (poprzez dokładne wskazanie budynku, piętra czy nawet mieszkania, z którego wykonano film lub zdjęcie), który później może ponieść konsekwencje z rąk osób i organizacji, na których niekorzyść działa dany dowód. Dlatego zawsze przed wykorzystaniem i opublikowaniem pozyskanych informacji należy zadać sobie pytanie „czy możemy wykorzystać ten dowód?".

Także w zakresie komunikacji pomiędzy analitykami, pracującymi na otwartych źródłach, a osobami wykorzystującymi tego typu informacje w swojej pracy (np. dziennikarzami) powinien istnieć taki sposób komunikacji, aby wychwytywać informacje, mogące być powodem późniejszych błędnych wniosków lub działań mających negatywny wpływ na opisywane osoby lub organizacje. W innym przypadku może powstać mylna informacja (ang. *misinformation*), która nie jest sama w sobie stworzona w celu wprowadzenia odbiorców w błąd (jak dezinformacja), jednak powoduje wyciąganie błędnych wniosków.

W działaniach nie podlegających pod reguły, jakie obowiązują w strukturach wojskowych, a więc głównie w cywilnych śledztwach dziennikarskich, częstym wynikiem konieczności szybkiego publikowania informacji, pod presją bycia wyprzedzonym przez inne źródła informacji, brakuje czasu na zastanowienie się nad etycznymi aspektami publikowanych informacji. Przez to ujawniane są także dane, dotyczące sposobu gromadzenia informacji o celu rozpoznania, co w ramach działań wywiadowczych jest utrzymywane w jak największej tajemnicy, aby nie pozwolić rozpoznawanym obiektom na uodpornienie się na wykorzystywane techniki rozpoznania. W celu uniknięcia tego typu negatywnych skutków działań rozpoznania otwartoźródłowego, przywołany wcześniej raport wskazuje na konieczność opracowania odpowiednich szkoleń i wytycznych w zakresie etyki działań oraz zapewnienie

możliwości kontaktu pomiędzy analitykami, pracującymi na otwartych źródłach a organizacjami rządowymi, w celu możliwości weryfikacji czy ujawnienie pozyskanych informacji nie będzie stanowiło zagrożenia ujawnienia zbyt dużej ilości wskazówek, dotyczących technik i źródeł.

Wskazana otwartość i przejrzystość działań OSINT-owych może także być podstawą do lepszej edukacji dla innych osób, borykających się z podobnymi problemami etycznymi. Jak wskazuje Hamilton Bean w artykule „*Is Open Source Intelligence an Ethical Issue?*”[106], pomimo, że pozyskiwane metodami OSINT-owymi dane pochodzą z jawnych i ogólnodostępnych źródeł, ich wykorzystanie powoduje cały czas dylematy etyczne.

Innym rodzajem dylematów etycznych, napotykanych podczas rozpoznania OSINT-owego jest wykorzystanie informacji pochodzących z wycieków danych, przy czym ten dylemat należy także rozpatrywać w aspekcie prawnym. Informacje, które stały się publiczne poprzez ich wykradzenie lub przypadkowe ujawnienie, spowodowane głównie błędami ludzkimi, stanowią bardzo dobrą bazę dla poszukiwań informacji dotyczących osób, jednak ich wykorzystanie w oficjalnych zleceniach, za które śledczy otrzymują wynagrodzenie, może być uznane jako czerpanie zysków z przestępstwa, dlatego też zawsze należy ostrożnie podchodzić do tego typu źródeł.

[106] H. Bean, *Is Open Source Intelligence an Ethical Issue?*, Research in Social Problems and Public Policy, Volume 19, s. 385-402.

Wnioski

Rozpoznanie, bazujące na otwartych źródłach, poprzez rozwój Internetu staje się możliwe do przeprowadzenia dla coraz szerszego grona osób. Dostępność zarówno narzędzi, jak i informacji o technikach do tego wykorzystywanych jest powszechna i w połączeniu z coraz większą ilością źródeł informacji, w postaci portali informacyjnych, serwisów społecznościowych, a także źródeł danych przestrzennych, powoduje wzrost popularności i wykorzystanie działań OSINT-owych.

W ramach wielu definicji informacji pozyskanych z rozpoznania otwartoźródłowego można wskazać powtarzające się określenia, dotyczące pochodzenia danych, które muszą być publicznie dostępne, a także sposób ich pozyskiwania, który musi być zgodny z prawem, przy czym dostępność nie musi oznaczać braku konieczności uiszczenia opłaty za możliwość pozyskania informacji.

Tradycyjne formy przekazu, na których w pierwszej połowie XX wieku bazował OSINT, aktualnie zastępowane są często przez źródła internetowe, co jednak nie zmienia często źródła ich pochodzenia, a jedynie formę przekazu i sposób odtarcia do odbiorcy.

Publiczna dostępność informacji w Internecie utożsamiana jest niekiedy z ich dostępnością w wyszukiwarkach internetowych, a często nawet w jednej, najpopularniejszej wyszukiwarce, jaką jest Google. Tego typu podejście pomija niestety ogromną większość informacji, które dostępne są w sieci nieindeksowanej, niedostępnej dla wyszukiwarek internetowych. Dopiero świadomość istnienia sieci *deep web* (w tym także *dark web*), a także sposobów dotarcia do zawartych w nich informacji, daje możliwość pozyskania danych przechowywanych w branżowych i tematycznych bazach danych, serwisach o ograniczonym dostępie oraz zasobach sieci *darknet*.

Rekonesans, który jest prowadzony w Internecie, musi być przeprowadzony z określonym wcześniej poziomem bezpieczeństwa, co implikuje decyzję o wykorzystaniu technik pasywnych (bez nawiązywania kontaktu bezpośredniego z rozpoznawanym obiektem, co jednak powoduje niższą dokładność, niż w przypadku rekonesansu aktywnego) oraz aktywnych (dających bardziej aktualne i dokładniejsze wyniki, ale także narażających osoby wykonujące rozpoznanie na wykrycie). Mnogość narzędzi dostępnych w Internecie oraz łatwość ich użycia sprawia, że działania OSINT-owe są dostępne dla każdego użytkownika światowej sieci, który będzie chciał tego typu rozpoznanie wykonać, przy czym należy zaznaczyć, że doświadczenie

w prowadzeniu analiz OSINT-owych umożliwia dużo wydajniejsze wykorzystanie narzędzi, a więc, co za tym idzie, możliwość otrzymania dokładniejszych i szerszych wyników. Dostępność serwisów, przekazujących dane przestrzenne, jak zdjęcia satelitarne, a także dane w postaci audiowizualnej, powoduje możliwość niemal natychmiastowej weryfikacji i samodzielnego osądzenia informacji, przekazywanych przez środki masowego przekazu oraz pojedyncze osoby np. w mediach społecznościowych.

Poprzez dużą dostępność opisywanych narzędzi oraz technik, a także coraz powszechniejszy dostęp do Internetu, bardzo ważną kwestią jest edukacja w zakresie etyki działań OSINT-owych, gdyż ich wykorzystanie w nieprzemyślany sposób może narazić osoby, firmy, a także operacje wojskowe na negatywne konsekwencje.

ROZDZIAŁ 3

ANALIZA INFORMACJI POZYSKANYCH Z ZASOBÓW INTERNETOWYCH JAKO FUNDAMENT WYWIADU

Uwagi wstępne

Przedstawione w poprzednim rozdziale narzędzia oraz sposoby zbierania informacji z jawnych źródeł w Internecie nie stanowią całości procesu wywiadu otwartoźródłowego. Zebrane informacje muszą zostać przetworzone oraz poddane analizie, w wyniku której dopiero uzyskiwany jest rezultat w postaci poprawnego rozpoznania.

Niestety nie zawsze praca wywiadowcza przynosi pożądane rezultaty. Niekiedy wynikiem błędnych analiz jest klęska operacji lub doprowadzenie do tragicznych w skutkach zdarzeń. Przykładami katastrofalnych skutków błędnych analiz są między innymi zdarzenia z ubiegłego tysiąclecia, które wskazuje Robert M. Clark w swojej książce „Intelligence Analysis: A Target-Centric Approach”[107]:

- Operacja Barbarossa z 1941 roku, podczas której Stalin postanowił sam przeprowadzić analizę, co okazało się katastrofalne w skutkach, gdyż błędnie założył on, że niemieccy dezerterzy którzy ostrzegali o mającym się wkrótce odbyć ataku z zaskoczenia, są jedynie prowokatorami i kazał ich rozstrzelać. Atak sił niemieckich zupełnie zaskoczył sowieckich generałów, co spowodowało znaczące straty po stronie ZSRR.
- W 1942 roku w Singapurze, wojska brytyjskie, australijskie i indyjskie poniosły klęskę i musiały się wycofać po tym jak mniejsza liczebnie armia japońska zmusiła ich do poddania się. Analizy brytyjskie nie doceniły możliwości japońskich myśliwców typu Zero oraz nie przewidziały możliwości użycia przez Japończyków czołgów podczas walk w dżungli, co dało armii japońskiej dużą przewagę nad Brytyjczykami.

[107] R. M. Clark, *Intelligence Analysis: A Target-Centric Approach*, CQ Press, 2019, s. 8.

- W 1973 roku izraelski wywiad sądził, że Egipt nie zaatakuje, nie odbudowawszy wcześniej swoich sił powietrznych i bez zawarcia sojuszu z Syrią. Dodatkowo, izraelscy dowódcy byli przekonani o sile swojego wojska, która miała być tak duża, że sama w sobie mogła odwieść potencjalnych agresorów od ataku. Przez to izraelscy dowódcy dali się zwieść egipskiej operacji, mającej na celu zmylenie przeciwnika i zignorowali prawidłowo zanalizowane informacje wywiadowcze, które wskazywały na możliwość ataku przez Egipt. Kiedy ten zaatakował podczas żydowskiego święta Yom Kippur 6 października 1973 roku, wojska Izraela poniosły bardzo duże straty odpierając atak.

Jako przyczyny porażek, związanych z błędnym procesem analitycznym, Robert M. Clark wskazuje trzy aspekty:

- Brak współpracy pomiędzy osobami pozyskującymi dane a analitykami.
- Błędy ludzkie, mające wpływ na jakość wytworzonych analiz – głównie są to wewnętrzne nastawienie oraz błędy poznawcze, które powodują brak obiektywizmu podczas przetwarzania danych.
- Brak odpowiedniego skorzystania z opracowanych danych przez ich odbiorców – podczas przekazywania analiz wywiadowczych, kluczowym jest nie tylko samo przekazanie, ale także upewnienie się, że odbiorca rozumie otrzymane informacje i będzie potrafił na ich bazie podejmować decyzje.

Na podstawie popełnionych błędów znacznie łatwiej można nauczyć się wagi, jaką należy przykładać do poprawnego prowadzenia tego etapu cyklu wywiadu niż w przypadku poprawnie podjętych decyzji, jednak każda niepoprawna analiza lub odrzucenie jej na etapie podejmowania kluczowych decyzji, może przynieść tragiczne skutki.

Biorąc pod uwagę powyżej opisane przypadki, można stwierdzić, że etap analizy zebranych danych i wyciągania wniosków powinien być poprzedzony edukacją w zakresie niebezpieczeństw, wynikających w dużej mierze z błędów ludzkich oraz że jego odpowiednie przeprowadzenie ma kluczowe znaczenie dla całego procesu wywiadu otwartoźródłowego.

1.15. Cykl wywiadowczy

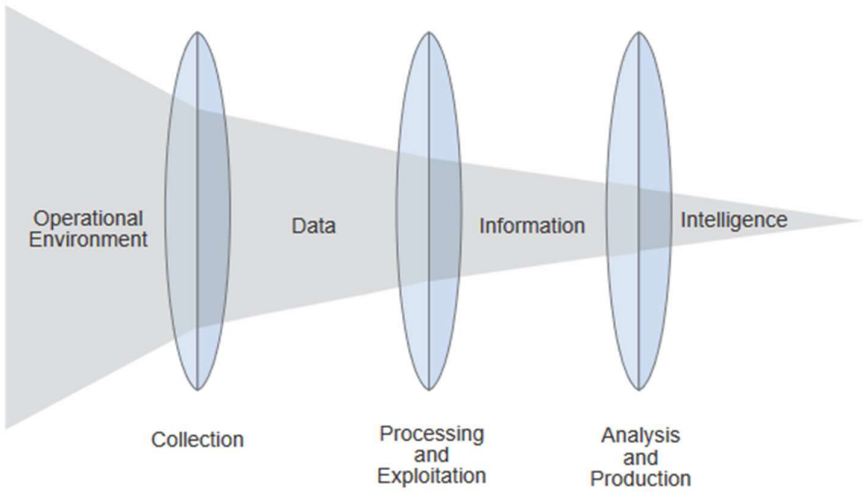
Etap analizy danych jest częścią cyklu wywiadowczego, który składa się najczęściej z pięciu etapów, jednak w zależności od konkretnego zastosowania i potrzeb, może być ich od trzech do siedmiu. Dodatkowe kroki mogą obejmować: analizę wymagań przed etapem planowania (tego typu podejście wskazuje FBI[108]), dodatkowy krok oceny i informacji zwrotnej (zgodnie z Joint Publication 2-0: Joint Intelligence[109]) lub osobne, dodatkowe dwa kroki po etapie przekazania – spożytkowanie i przekazanie informacji zwrotnej od odbiorców produktów wywiadowczych[110].

Podstawę cyklu wywiadowczego stanowi jednak proces w formie pięcioetapowej, co przedstawione zostało na Rys. 11.

Rys. 11 – Cykl wywiadowczy.
Źródło: opracowanie własne na podstawie „Central Intelligence Agency: Factbook on Intelligence"[111]

[108] *Intelligence Cycle Graphic*, https://www.fbi.gov/image-repository/intelligence-cycle-graphic.jpg/view – dostęp online 03.10.2022 r.
[109] Joint Chiefs of Staff, Joint Publication 2-0: *Joint Intelligence*, 2013, s. I-6, https://fas.org/irp/doddir/dod/jp2_0.pdf – dostęp online 03.10.2022 r.
[110] M. M. Lowenthal, *Intelligence: From Secrets to Policy*, wyd. 6, CQ Press, 2014, s. 71.
[111] Central Intelligence Agency, *Factbook on Intelligence*, https://irp.fas.org/cia/product/facttell/index.html – dostęp online 03.10.2022 r.

W trakcie prac w ramach cyklu wywiadowczego, zbierane surowe dane są przetwarzane i stają się informacjami, które stanowią większą wartość dla możliwości podjęcia decyzji na ich podstawie. Dopiero jednak porównanie uzyskanych informacji z innymi, a także inne formy analizy, bazujące na przykład na dotychczasowej wiedzy oraz doświadczeniu, tworzą wiedzę wywiadowczą, pozwalającą przewidzieć przebieg wydarzeń oraz wskazuje możliwe ich ścieżki[112]. Wskazaną transformację ilustruje Rys. 12.



Rys. 12 – Dane, informacje oraz wiedza wywiadowcza. Źródło: Joint Publication 2-0: Joint Intelligence

Poszczególne etapy cyklu wywiadowczego charakteryzują się następującymi zakresami działań:

- Planowanie / podejmowanie decyzji – stworzenie wymagań i celów dla działań rozpoznania, co rozpoczyna cykl wywiadowczy, ale także stanowi jego zakończenie, którego wynikiem są informacje wywiadowcze i nowe wymagania.
- Zbieranie danych – pozyskiwanie danych w formie „surowej", na podstawie jawnych źródeł, takich jak m.in. prasa, radio, telewizja czy Internet, zarówno w formie pisanej, jak i obrazowej, w ramach rodzajów rozpoznania wskazanych w rozdziale 1.11.

[112] Joint Chiefs of Staff, Joint Publication 2-0: *Joint Intelligence*, 2013, s. I-1, https://fas.org/irp/doddir/dod/jp2_0.pdf – dostęp online 03.10.2022 r.

- Przetwarzanie – obejmuje zmianę formy pozyskanych wcześniej danych poprzez odszyfrowywanie, tłumaczenie, konwersję kodowania, uzyskanie danych z nośników magnetycznych i światłoczułych, a także redukcję danych.
- Analiza i wytwarzanie – oznacza zamianę wcześniej pozyskanych i przygotowanych informacji w informację wywiadowczą poprzez ich łączenie, weryfikację prawdziwości, pewności i przydatności oraz analizę, a także wytworzenie przygotowanie całościowego podsumowania.
- Przekazywanie – dystrybucja gotowych raportów wywiadowczych do ich odbiorców, w formie zależnej od ich potrzeb, co może zakończyć cykl wywiadowczy lub prowadzić do przejścia ponownie do etapu podejmowania kolejnych decyzji na podstawie przedstawionych w tym kroku wyników.

Jakość otrzymywanych informacji wywiadowczych, wypracowanych w ramach powyższego cyklu będzie zależna od poziomu jakości każdego z jego kroków. Możliwe jest jednocześnie zakłócanie procesu rozpoznania, co jest tym bardziej prawdopodobne, im bardziej dynamiczne będzie pole rywalizacji, im więcej zjawisk będzie zachodziło w otoczeniu i im mniej czasu będzie na zdobywanie i opracowywanie danych[113].

[113] P. Dela, *Zagrożenia działań w cyberprzestrzeni*, PWN, Warszawa, 2022, s. 134.

1.16. Analiza zebranych informacji

Nawet ogromny i bardzo dokładny zestaw zebranych danych nie będzie użyteczny bez ich odpowiedniej analizy. To właśnie proces analizowania informacji, które zostały zebrane i przygotowane w ramach poprzednich etapów cyklu wywiadowczego ma odpowiedzieć na zadane wcześniej pytania, stworzyć tzw. mozaikę wywiadowczą, wskazać możliwe dalsze kroki oraz, jeśli zajdzie taka konieczność, stworzyć nowe wymagania i pytania, na które należy znaleźć odpowiedź.

W tym etapie konieczne jest utrzymywanie przez osoby przeprowadzające analizę odpowiedniego kursu, wskazanego przez postawione wcześniej wymagania i pytania. Wszelkie nadmierne rozszerzenia zakresu analizy, zbytnie odchodzenie od głównego pytania, a także próby samodzielnego wypełnienia braków w danych będą miały negatywny skutek na jakość wytworzonych informacji wywiadowczych. W przypadku luki w informacjach konieczne jest podkreślenie tego faktu, aby odbiorca raportu miał świadomość co jest tzw. „wiadome niewiadome". Pomimo pozornej zawiłości zakresów posiadanej wiedzy, dobrze ilustruje jej podział wypowiedź Sekretarza Stanu USA, Donalda Rumsfelda z 2002 roku:

„*Zawsze interesują mnie raporty, które informują, że czegoś nie wiemy, ponieważ jak wiemy istnieją wiadome wiadome – są to rzeczy, o których wiemy, że o nich wiemy. Wiemy także, że istnieją wiadome niewiadome – oznacza to, że wiemy o rzeczach, o których nie wiemy. Lecz istnieją także niewiadome niewiadome – te, o których nie wiemy, że o nich nie wiemy. I jeśli spojrzeć na historię naszego kraju, jak i innych wolnych krajów, to ta ostatnia kategoria zazwyczaj okazuje się być tą, przysparzającą najwięcej kłopotów.*"[114]

Wskazywanie możliwego przebiegu zdarzeń w przyszłości nie może być mylone z przewidywaniem przyszłości, ale powinno być przedstawieniem oceny sytuacji z określonym poziomem pewności, bazującym na zestawieniu faktów i przypuszczeń[115]. W celu określenia jak daleko analityk może posunąć się w określaniu prawdopodobieństwa swoich przypuszczeń, możliwe jest zastosowanie tzw.

[114] Cytat z wypowiedzi Donalda Rumsfelda podczas konferencji prasowej w Pentagonie w dniu 12.02.2002 r. Tłumaczenie własne na podstawie tekstu z książki: D. Rumsfeld, *Known and Unknown: A Memoir*, Penguin, 2011, s. 8.
[115] L. Krizan, *Intelligence essentials for everyone*, Joint Military Intelligence College, Waszyngton, 1999, s. 29.

wywiadowczego łańcucha pokarmowego (ang. *Intelligence Food Chain*), który składa się z czterech stopni[116]:

- Fakty – zweryfikowane informacje powiązane z danym rozpoznaniem;
- Znaleziska – wiedza ekspercka, bazująca na informacjach, wskazująca np. trendy lub wzorce;
- Przewidywania – osądy oparte na faktach i możliwe do obronienia przy wykorzystaniu jasnych i solidnych argumentów;
- Wróżenie – niewystarczająco wyjaśnione i bronione osądy.

Wszystkie powyższe terminy w języku angielskim zaczynają się od litery „F" (*Facts, Findings, Forecasts, Fortunetelling*), stąd przy pomocy wyznaczenia granicy „4 F minus jeden" możliwe jest zaznaczenie, jak powinny być prowadzone analizy, by nie zafałszować ich wyniku.

W celu znalezienia jak najpoprawniejszych odpowiedzi na zadane na początku pytania, analitycy muszą posłużyć się odpowiednim rozumowaniem, czyli nabytymi umiejętnościami wyjaśniania zdarzeń, w oparciu o posiadane informacje i świadomość zdarzeń, które wystąpiły już wcześniej. W ramach analiz wywiadowczych, wyróżnić można cztery podstawowe typy rozumowań[117]:

- Rozumowanie dedukcyjne – typ logicznego rozumowania, wykorzystujący zbiór posiadanych danych lub faktów, w celu przedstawienia wniosku, dający zawsze prawdziwe wyniki, gdyż jedynie reorganizuje istniejące już dane, bez tworzenia nowych (jeśli to jest prawdą, to tamto też jest prawdą). Tylko ten rodzaj rozumowania ze wszystkich przedstawionych tutaj daje wynik zawsze prawdziwy, jednak rzadko podczas działań rozpoznawczych występuje możliwość zastosowania tej reguły.
- Rozumowanie indukcyjne – wyszukiwanie w posiadanych danych lub faktach wzorców i trendów oraz wnioskowanie, że zależności zostaną utrzymane. Nie daje ono pewności co do spodziewanych wyników, a jedynie pewne przypuszczenie (Bazującym na tej zależności, to jest prawdopodobnie prawdą).

[116] Tamże, s. 29.
[117] Opracowano na podstawie: ATP 2-33.4. *Intelligence Analysis*, Styczeń 2020 – https://irp.fas.org/doddir/army/atp2-33-4.pdf – dostęp online 07.10.2022 r.

- Rozumowanie abdukcyjne – podobnie jak rozumowanie indukcyjne bazuje na informacjach prawdopodobnych i tworzy na ich podstawie także prawdopodobne wyjaśnienia i wnioski (Ponieważ to jest prawdopodobnie prawdą, to tamto także może być prawdą).

- Rozumowanie przez analogię – metoda bazująca na znajdowaniu podobieństw i powiązań pomiędzy danymi obiektami, umożliwiająca tworzenie na tej podstawie wniosków – im więcej podobieństw występuje pomiędzy obiektami, tym silniejszy jest wniosek.

To, jakie korzyści w ramach analizy przyniesie korzystanie z podanych typów rozumowania zależy w dużej mierze od samego analityka i jego wyszkolenia oraz doświadczenia. Wśród personelu wojskowego, wykonującego prace analityczne, wskazane jest posiadanie trzech podstawowych umiejętności w zakresie umiejętności myślowych: szeregowania informacji (w celu ich łatwiejszej analizy, syntezy oraz lepszego zrozumienia), rozpoznawania wzorców (w celu odnalezienia powiązań i oddzielenia faktów nieistotnych od tych ważnych) oraz właśnie rozumowania w celu odnalezienia wyjaśnień i nadania znaczenia wydarzeniom i działaniom[118]. Niestety we wszystkich tych umiejętnościach kryje się niebezpieczeństwo zastosowania znanych wzorców, osobistych przekonań oraz skrótów myślowych, co może prowadzić do ograniczeń w pracy analityków lub do formułowania nieprawdziwych wniosków. Temat ten jako ważny aspekt wywiadu, szerzej omawiany jest w kolejnym podrozdziale.

Za dobrą praktykę uznać można także tworzenie zestawień posiadanych danych oraz ich wizualizację na różne sposoby, gdyż nawet najlepsze dane w formie surowej nie będą tak dobrze opisywać sytuacji, jak dane w formie np. tabelarycznej lub graficznej. Dopiero porównanie danych, ich przypisanie do odpowiednich grup i zakresów lub też wskazanie trendu będzie dawało pogląd ich odbiorcom na ich znaczenie.

[118] ATP 2-33.4. *Intelligence Analysis*, Styczeń 2020, p.B-2 – https://irp.fas.org/doddir/army/atp2-33-4.pdf – dostęp online 07.10.2022 r.

1.17. Błędy poznawcze w ocenie danych

Analizy, wykonywane nawet przez najlepiej wykwalifikowane osoby, są w każdym momencie narażone na błędy ludzkie, wynikające z niedoskonałości oraz sposobów funkcjonowania ludzkiego rozumowania. Analitycy są ponadto nie tylko narażeni na błędy wynikające ze zbyt małej lub zbyt dużej ilości zebranych wcześniej informacji, ale także na te, wynikające z umyślnego zakłamania pozyskanych danych, jak na przykład akcje mające na celu zmylenie przeciwnika. Współcześnie, elementem polityki państw i korporacji jest walka informacyjna, w ramach której można wyróżnić m.in. prowadzenie kampanii informacyjnych z elementami propagandy i dezinformacji (ang. *denial and deception*), co jest nieodzownym elementem bezpieczeństwa informacyjnego[119]. Dlatego też tak ważne jest zidentyfikowanie czynników, które mogą zachwiać pewność analizy ogromnej ilości danych, dostępnych w otwartych źródłach.

To, na co analitycy mają wpływ, to zdolność do weryfikowania swoich wzorców myślenia i próby jak największego zdystansowania się emocjonalnie, kulturowo i światopoglądowo od analizowanego tematu, aby uniknąć błędów poznawczych.

1.17.1. Wzorce w ludzkim rozumowaniu

Sukces procesu analizy danych wywiadowczych zależy w ogromnym stopniu od umiejętności samego analityka, jego wykształceniu w zakresie prowadzenia analiz, ale także doświadczeniu i znajomości słabości ludzkich sposobów rozumowania i wyciągania wniosków. Richards J. Heuer Jr., długoletni pracownik Dyrektoratu Wywiadu CIA, a także wykładowca i nauczyciel wielu analityków pracujących w tej instytucji, w swojej książce „Psychology of Intelligence Analysis”[120] wskazuje, że to, co i w jakiś sposób ludzie postrzegają, uzależnione jest silnie od ich dotychczasowych doświadczeń, edukacji, wartości kulturowych i wymagań społecznych, a także tego, co odbierają ich zmysły. Jednym z podstawowych problemów podczas postrzegania informacji jest to, że ludzie zazwyczaj dostrzegają to, czego się spodziewają dostrzec.

[119] P. Dela, *Elementy propagandy w życiu publicznym*, Studia Politologiczne 54, s. 68-95.
[120] R. J. Heuer Jr., *Psychology of Intelligence Analysis*, Center For The Study Of Intelligence, Central Intelligence Agency, 1999.

Percepcja według Heuera jest procesem nie pasywnym, rejestrującym faktyczne zdarzenia, lecz procesem aktywnym, w którym ludzie konstruują ich własną wersję rzeczywistości na podstawie doznań ze wszystkich swoich zmysłów[121].

Z jego pracy można wyodrębnić trzy podstawowe zagadnienia, z którymi musi mierzyć się każdy analityk:

- Ludzki mózg ma problemy z radzeniem sobie jednocześnie z niepewnością wbudowaną, tzn. z naturalną mgłą, otaczającą złożone i nieokreślone zagadnienia wywiadowcze, oraz z niepewnością tworzoną, tzn. ze sztucznie wytwarzanym przez ludzi w ramach działań zaprzeczania i oszukiwania zamgleniem sytuacji.

- Nawet zwiększona świadomość istnienia zarówno błędów poznawczych, jak i innych „bezwiednych" błędów, jak np. tendencja do lepszego rozpoznawania informacji potwierdzających dotychczasowe wnioski niż informacji im zaprzeczających, sama w sobie niewiele jest w stanie pomóc analitykom w radzeniu sobie z niepewnością.

- Narzędzia oraz techniki wspomagające pracę analityków w zakresie osiągnięcia wyższego poziomu krytycznego myślenia, mogą istotnie poprawić analizę złożonych zagadnień, w których posiadane informacje są niepełne, niejednoznaczne, a także często umyślnie zniekształcone. W tym celu kluczowym będzie wykorzystanie technik strukturyzacji informacji, konkurujących przypuszczeń oraz rozpoznawanie alternatywnych interpretacji.

Jak dalej zauważa Heuer, ludzie postrzeganie zdaje się być łatwe w tworzeniu, lecz odporne na zmiany, a nowe informacje są dołączane do już istniejących wzorców. To może tłumaczyć przypadki, kiedy nowi analitycy dostrzegali szczegóły przeoczone przez tych, którzy pracowali nad danym zagadnieniem już nawet 10 lat. Także długotrwałe analizowanie „rozmytych" informacji ma wpływ na późniejsze dostrzeżenie poprawnych odpowiedzi, gdyż jak pokazuje przywoływany przez Heuera eksperyment, osoby, którym przez dłuższy czas pokazywano rozmazane zdjęcia, przedstawiające konkretne przedmioty lub zwierzęta, potrzebowały więcej czasu, aby rozpoznać je na ostrzejszych zdjęciach. Tak samo sytuacja miała się, jeśli dane osoby rozpoczynały od zdjęcia bardziej rozmazanego, które później stawało się coraz ostrzejsze. Podobne zależności istnieją

[121] Tamże, s.7.

także w pracy analitycznej, w której po długotrwałym przetwarzaniu danych niejasnych i wieloznacznych, potrzeba świeżego spojrzenia, aby dostrzec prawidłowe zależności i odpowiedzi lub być może dojść nawet do innych wniosków niż na początku.

Aby zaradzić opisywanym problemom, konieczne jest m.in. wskazywanie możliwości i ograniczeń, jakim podlegają analizy wywiadowcze, stawianie na procedury obejmujące korzystanie z różnych punktów widzenia, a także cykliczne ponowne badanie kluczowych problemów od podstaw, aby uniknąć efektu podejścia inkrementacyjnego.

Konstrukcja ludzkiego umysłu, w celu ułatwienia przetwarzania danych, korzysta z zaprogramowanych w pamięci wzorców powiązań pomiędzy danymi. Do takiego wzorca odwołuje się umysł ludzki w momencie przywołania określenia z nim związanego. Stanowi to jednocześnie duże ułatwienie w zapamiętywaniu złożonych zestawów informacji, jak i niebezpieczeństwo użycia typowego wzorca podczas analizy danych, powiązanych z nim w niewielkim stopniu. Ten rodzaj „dowiązywania" informacji do już istniejących wzorców jest jednym z trzech sposobów na zapamiętywanie nowych informacji[122] – jest to nauka poprzez łączenie. Dwie inne metody to: uczenie się „na pamięć" oraz korzystanie z urządzenia mnemonicznego, czyli wiązanie zapamiętywanych informacji z wzorcem, np. tworzenie słowa z pierwszych liter zapamiętywanej serii wyrazów lub kojarzenie zapamiętywanych informacji z określonym miejscem lub czynnością, która pozwoli na łatwiejsze ich przypomnienie.

1.17.2. Rodzaje błędów poznawczych, mających wpływ na efekty rozpoznania otwartoźródłowego

W ciągle zmieniającym się środowisku, jakim jest sytuacja na świecie, analitycy muszą nauczyć się rozpoznawać sposoby rozumowania, którymi się posługują i zdać sobie sprawę, że istniejące w pamięci wzorce mogą negatywnie wpływać na ich ocenę sytuacji, która może być inna każdego dnia. To właśnie dzięki tym wzorcom analitycy mogą wypełniać luki w posiadanych informacjach, tworząc w ten sposób osądy, dotyczące badanego tematu, dlatego powinni oni używać sposobów na badanie swoich osądów i weryfikowanie ich za pomocą kreatywnego myślenia. Poniżej przedstawiono zestawienie błędów poznawczych, które autor zidentyfikował jako mogące mieć negatywny wpływ na wyniki prowadzonego OSINT-u:

[122] Tamże, s.25.

- **Błąd konfirmacji**, zwany także efektem potwierdzenia (ang. *confirmation bias*) ma podobną zasadę działania do opisywanego powyżej efektu lepszego dostrzegania tego, czego się spodziewamy. Polega na tym, że ludzie mają tendencję do faworyzowania idei, które są zbieżne z ich przekonaniami lub wcześniejszymi hipotezami. Jeśli jakaś osoba podejmie już jakąś decyzję, będzie szukała informacji ją popierających, ignorując lub umniejszając jednocześnie wartość informacji, które stoją z nią w sprzeczności. Może to skutkować pomijaniem istotnych faktów, a co za tym idzie, błędnymi wynikami analiz, dlatego osoba badająca dostępne informacje powinna być jak najbardziej zdystansowana od tematyki, którą się zajmuje. Wśród sposobów radzenia sobie z efektem potwierdzenia można wyróżnić chociażby technikę konkurujących hipotez, opisywaną m.in. w książce „Psychology of Intelligence Analysis" R. J. Heuera. Możliwe jest także zweryfikowanie przez analityka czy nie padł ofiarą tego błędu poznawczego, np. poprzez zadanie sobie pytania „czy wnioski potwierdziły moje wcześniejsze przypuszczenie i dlaczego?". Możliwe jest także zweryfikowanie przez analityka wyniku wyciągnięcia przez niego odwrotnych wniosków.

- **Efekt lustrzanego odbicia** (ang. *mirror-imaging effect*) jest błędem zaburzającym obiektywność osób prowadzących analizę na podstawie niekompletnych danych. Jego istotą jest bazowanie przez analityka na własnych doświadczeniach, światopoglądzie i zasadach kulturowych. Poprzez odniesienie tych aspektów do analizowanych obiektów, dochodzi do błędnego przypisania im takich samych cech i oczekiwania od nich postępowania, opartego na podobnych wartościach. Efekt lustrzanego odbicia był przyczyną błędnego określenia zagrożeń, związanych z atakiem na Pearl Harbor w 1941 roku czy na World Trade Center w 2001 roku. W pierwszym przypadku Amerykanie nie potrafili przyjąć do wiadomości, że możliwe jest zaatakowanie przez Japończyków kogoś, kogo potencjał jest o wiele większy, co mogło skutkować jedynie porażką. Coraz większa część świata przyjmuje tzw. kulturę zachodnią, co może prowadzić do błędnego wysnuwania wniosków, że wszelkie grupy rządowe lub pozarządowe,

będą kierowały się zachodnimi wartościami podczas planowania swoich działań[123].

- **Efekt kadrowania** (ang. *framing effect*) jest błędem poznawczym, który polega na różnym interpretowaniu faktów, na podstawie różnego sposobu ich pierwotnej prezentacji, np. w formie pozytywnych lub negatywnych danych. Na podstawie różnego przedstawienia tych samych danych możliwe jest przez to uzyskanie różnych wyników ich analizy.

- **Egotyzm atrybucyjny** (ang. *self-serving bias*) oznacza tendencję do wyjaśniania swojego zachowania prze pryzmat czynników działających na korzyść ich samooceny. W przypadku sukcesów lub pozytywnych zdarzeń, ludzie zwykle przypisują większe znaczenie czynnikom wewnętrznym, zależnym od nich, jak np. swojej osobowości czy inteligencji, a w przypadku poszukiwania przyczyn negatywnych zdarzeń, większe znaczenie przypisywane jest czynnikom zewnętrznym, niezależnym od danej osoby.

- **Efekt dostosowania i zakotwiczenia** (ang. *adjustment and anchoring effect*) jest błędem poznawczym, który wynika z pierwotnego przyjęcia jakiejś informacji (np. idei, liczby lub właściwości) jako decydującej o dalszym przebiegu rozumowania. Informacja ta może wynikać np. z kolejności w jakiej zapoznajemy się z danymi lub z pierwszych, zgrubnych określeń lub obliczeń. Różne, najczęściej bezwiednie przyjęte szacowania lub wrażenia powodują zaburzenie dalszych działań w kierunku początkowej „kotwicy".[124]

- **Błąd dostępności**, zwany także heurystyką dostępności (ang. *availability bias* lub *availability heuristic*) jest rodzajem uproszczonego wnioskowania, które bazuje na przypisywaniu większego znaczenia elementom, które są łatwe do przywołania (na przykład przez to, że wydarzyły się niedawno i wspomnienie o nich jest jeszcze silne albo wystąpiły wiele razy, przez co lepiej zapadły w pamięć) lub są silnie nacechowane emocjonalnie w psychice danej osoby. Może być to spowodowane informacjami zasłyszanymi ostatnio w doniesieniach medialnych,

[123] L. Witlin, *Of Note: Mirror-Imaging and Its Dangers*, SAIS Review of International Affairs, Johns Hopkins University Press, Volume 28, Number 1, Winter-Spring 2008, s. 89-90.
[124] A. Tversky; D. Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, Science, New Series, Vol. 185, No. 4157. (27.09.1974), s. 1124-1131.

które spowodują podświadome przypisywanie większej częstotliwości lub prawdopodobieństwa wystąpienia danego zdarzenia, ponieważ informowały o nim media[125].

- **Iluzja grupowania** (ang. *clustering illusion*) oznacza naturalną tendencję człowieka do zauważania wzorców i zbiorów danych w miejscach, w których losowe dane wydają się być ułożone w sposób nie-losowy. Jednymi z najpowszechniejszych przykładów są próby odnajdywania wzorów przypominających konkretne przedmioty lub osoby w układzie chmur, gwiazd czy nawet plam. Iluzja ta nasila się szczególnie w przypadku posiadania małej ilości analizowanych danych.

W zakresie identyfikacji błędów poznawczych można też wyróżnić korzystanie z nieprawidłowych analogii, które są spowodowane osobistymi przekonaniami lub brakiem odpowiedniej wiedzy w danym zakresie. Już sama identyfikacja błędów poznawczych, jakie mogą zostać napotkane podczas analizy jest dużym krokiem naprzód w kierunku minimalizacji ich wpływu na końcowy wynik rozpoznania.

[125] A. Tversky, D. Kahneman, *Availability: A heuristic for judging frequency and probability*, Cognitive Psychology, Volume 5, Issue 2, wrzesień 1973, s. 207-232.

1.18. Synteza i poprawne wyciąganie wniosków w celu uniknięcia błędów poznawczych

Wyciąganie wniosków z przeprowadzonej analizy i formułowanie ich jest tematem, który jest podatny na zaburzenia, gdy wykonują go osoby bez przygotowania analitycznego i nadzoru w tym zakresie. Jest to widoczne szczególnie w śledztwach, prowadzonych przez tzw. internetowych detektywów (ang. *Internet sleuths*), którzy nie wiedzą lub nie przejmują się odpowiedzialnością za wykonane analizy i wyciągnięte wnioski. Dlatego też w ich działaniach najlepiej widoczne są zagadnienia, związane z przedmiotowym zagadnieniem.

Zachowania obserwowane po ataku bombowym podczas maratonu w Bostonie w 2013 roku pokazują, że jednym z elementów, które wpływają na zaburzenia poprawnego formułowania wniosków z rozpoznania otwartoźródłowego jest efekt podczepienia, który został już przywołany w rozdziale XX. Działania w grupie powodują, że osoby wierzą w określone rzeczy tylko dlatego, że inni tak mówią lub robią. Stąd też biorą się błędne przekonania, że znalezione zostały mocne dowody w śledztwie lub że wnioski z niego są poprawne, kiedy są one publikowane i rozpowszechniane przez dużą liczbę osób. Innymi błędami poznawczymi, które można było zaobserwować w trakcie internetowych śledztw po zamachu w Bostonie były m.in. błąd (heurystyka) niedostępności czy egotyzm atrybucyjny. Podczas formułowania wniosków formułowane były kryteria, które nie mają silnych podstaw do bycia decydującymi parametrami w tego typu analizie. Przykładem takiego kryterium było np. wysuwanie podejrzeń wobec osób, które na zdjęciach nie patrzyły w stronę maratonu lub miały inny kolor skóry. Skutkiem nieprawidłowych analiz były niesłuszne oskarżenia, pogróżki w Internecie i nawet proces o bezpodstawne publikowanie niesłusznie podejrzewanych osób na pierwszej stronie gazety. Opublikowanie zdjęć osoby niesłusznie podejrzewanej o zamachy w Paryżu i w Nicei w latach 2015-2016 przez hiszpańską gazetę *La Razón* na pierwszej stronie jest przykładem, że nawet w organizacjach, które powinny weryfikować publikowane przez siebie oskarżenia, zdarzają się analizy, których obalenie jest banalnie proste dla osoby z chociażby niewielkim doświadczeniem w rozpoznaniu otwartoźródłowym.

W przywołanej już wcześniej książce Richardsa J. Heuera „Psychology of Intelligence Analysis", jej autor definiuje listę kontrolną, która ma pomóc analitykom

radzić sobie z wpływem błędów poznawczych na ich działania. Obejmuje ona następujące punkty[126]:

- Zdefiniowanie problemu – zadanie odpowiednich pytań, określenie odbiorców prac i zdefiniowanie jak dokładne informacje mogą zostać uzyskane w ramach prac analitycznych w zadanym czasie.

- Identyfikację prawdopodobnych hipotez – spisanie wszystkich możliwych hipotez, które mogą być wypracowywane np. przy współpracy z ekspertami lub podczas burzy mózgów, a następnie określenie listy, która będzie możliwa do przetworzenia w zadanym czasie. Podczas tej czynności ważne jest spisywanie także hipotez, co do których nie ma jasnych dowodów na tym etapie. Możliwe, że wśród nich pojawi się także hipoteza o próbie zmylenia przeciwnika.

- Zbieranie informacji – pozyskiwanie informacji nie tylko z już dostarczonych zestawów danych, ale także poszukiwanie ich poza nim, także przy pomocy innych komórek organizacyjnych. W przypadku należy wstrzymać się od jakichkolwiek osądów i wniosków, a jeśli analityk uzna, że zna już odpowiedź na jakieś pytanie, powinien poszukać informacji, które mogłyby zmienić jego zdanie na ten temat.

- Weryfikację hipotez – ważniejsze są te dowody, które zaprzeczają danej hipotezie, niże te, które ją potwierdzają, więc należy się skłaniać raczej ku odrzucaniu hipotez, niż na ich wspieraniu. Tutaj bardzo ważna jest świadomość błędów poznawczych, ponieważ mogą one wpłynąć na interpretację posiadanych dowodów.

- Wybranie najprawdopodobniejszej hipotezy – wskazanie tej hipotezy, która ma najmniej obalających ją dowodów, a także kilku innych, które są mniej prawdopodobne, wraz z argumentami, które zostały wykorzystane.

- Ciągłe monitorowanie – weryfikacja czy nie pojawią się nowe fakty, które mogą wpłynąć na prawdopodobieństwo poszczególnych hipotez lub zmienić postrzeganie całego tematu.

[126] R. J. Heuer Jr., *Psychology of Intelligence Analysis*, Center For The Study Of Intelligence, Central Intelligence Agency, 1999, s. 173-177.

W celu stworzenia lepszego środowiska do pracy dla analityków, Heuer sugeruje także zakres działań dla kierownictwa, które powinno wspierać badania wpływu błędów poznawczych na wyniki rozpoznania, szkolenia z zakresu analizy informacji (w tym także wsparcie, udzielane przez mentorów i emerytowanych pracowników), umożliwienie i wspieranie alternatywnych sposobów myślenia oraz takie prowadzenie analiz, aby pracownicy nie bali się publikować mniej prawdopodobnych scenariuszy, dokonywać ponownego rozpatrywania niektórych spraw i informować o niepewności wyników.

Innym sposobem ograniczenia wpływu błędów poznawczych jest określenie kryteriów dla posiadanych danych – ich użyteczności, wiarygodności, pewności oraz kompletności, a także takie metody jak:

- **Analiza „Co, jeśli?”** – dla analiz wskazujących na brak możliwości wystąpienia konkretnego zdarzenia, przyjęcie założenia, że to zdarzenie jednak miało miejsce i weryfikacja jak mogło do tego dojść.
- **Metoda sześciu kapeluszy** – metoda Edwarda de Bono, w której analitycy przyjmują konkretne role, określone kolorami wirtualnych kapeluszy. Oprócz osoby prowadzącej eksperyment (niebieski kapelusz), są jeszcze osoby o nastawieniu: neutralnym (biały), emocjonalnym (czerwony), kreatywnym (zielony), optymistycznym (żółty) i pesymistycznym (czarny).[127]

[127] *Cognitive biases – acaps technical brief*, https://www.acaps.org/sites/acaps/files/resources/files/acaps_technical_brief_cognitive_biases_march_2016.pdf, dostęp online 21.10.2022 r.

1.19. Inne czynniki ludzkie wpływające na proces prowadzenia wywiadu otwartoźródłowego

Osoby pracujące nad tematami, które niosą ze sobą dużą wagę emocjonalną, powinny także zwrócić uwagę na inny aspekt psychologiczny, mogący mieć wpływ na przebieg prowadzonego przez nich rozpoznania, a mianowicie zjawisko traumy zastępczej lub zapośredniczonej (ang. *vicarious trauma* lub *vicarious traumatization*, skr. VT), znanej także jako wtórny zespół stresu pourazowego (ang. *secondary traumatic stress disorder*, skr. STSD). Jak zauważa Nico Dekens w swoim artykule „Vicarious trauma and OSINT – a practical guide"[128], szczególnie narażeni na ten rodzaj traumy są osoby, które mają częsty kontakt z materiałami takimi jak filmy, zdjęcia lub nagrania dźwiękowe dotyczące np. zbrodni wojennych, przestępstw wobec dzieci lub aktów terrorystycznych. Osoby prowadzące śledztwa OSINT-owe także są narażeni na tego typu traumę, szczególnie jeśli pracują same, w pomieszczeniach, w których są oddzieleni od innych osób i skupiają się głównie na swojej pracy. Oznakami pogłębiającej się traumy wtórnej mogą być np.: uczucie przygnębienia, wahania nastroju, bezsenność lub koszmary senne, niechęć społeczna, gniew, problemy seksualne czy większa wrażliwość na przemoc. Aby zapobiegać negatywnemu wpływowi emocji na psychikę, warto skorzystać z takich technik jak: wyłączanie dźwięku podczas oglądania materiałów video, nieużywanie słuchawek, oglądanie materiałów video w wersji czarno-białej zamiast kolorowej, a także oglądanie ich w obecności innych osób lub poza biurem, gdzie jest dużo innych dźwięków. Powinno się także unikać przynoszenia pracy do domu, co uniemożliwi odcięcie się od tematyki powodującej negatywne skutki emocjonalne. W przypadku zauważenia oznak traumy nie wolno ich lekceważyć, ponieważ w pewnym momencie może okazać się ona zbyt silna, by sobie z nią poradzić. Sposobami na tzw. "oczyszczenie umysłu" są m.in.: zmiana komputera lub chociaż tła w systemie operacyjnym przy pacy nad różnymi sprawami, zmiana ubrania po pracy lub cykliczne wychodzenie na świeże powietrze, aby jak najbardziej oddzielić świat pracy nad traumatyzującymi materiałami od czasu odpoczynku.

[128] N. Dekens, *Vicarious trauma and OSINT – a practical guide*, https://osintcurio.us/2020/06/08/vicarious-trauma-and-osint-a-practical-guide/ – dostęp online 30.10.2022 r.

Wnioski

Prowadzenie analiz i formułowanie wniosków w ramach wywiadu otwartoźródłowego jest tak samo narażone na błędy ludzkie (a szczególnie błędy poznawcze, wynikające ze sposobów ludzkiego rozumowania) jak każdy inny rodzaj wywiadu. Dlatego też osoby wykonujące OSINT powinny mieć świadomość niebezpieczeństw wynikających z tych błędów oraz być wspierane w przeciwdziałaniu im, głównie poprzez odpowiednie szkolenie i wsparcie ekspertów. Błędy te mogą mieć wpływ na pomijanie istotnych aspektów lub nadawanie znaczenia aspektom nieistotnym, czego skutkiem może być zestaw nieprawidłowych wniosków, na podstawie których zostaną podjęte nieodpowiednie decyzje.

W przypadku działań analitycznych podczas rozpoznania na poziomie organizacji, która nie zadbałaby o szczegółową weryfikację publikowanych wniosków z analiz, skutkiem może być nie tylko negatywny wpływ na wyniki lub reputację danej organizacji, ale także niepotrzebne odkrycie własnych działań przed ewentualnymi osobami, których takie rozpoznanie mogłoby dotyczyć. Dobór odpowiednich kryteriów dla wyciągania wniosków, a także gruntowna ich analiza przez doświadczonych analityków, ekspertów dziedzinowych oraz osób o innym punkcie widzenia ma zatem ogromne znaczenie w wykonywaniu analiz. Jest to jeszcze ważniejsze, jeśli od tych decyzji zależy zdrowie i życie ludzkie.

W celu weryfikacji, czy podczas formułowania wniosków nie zostały popełnione błędy, można zadać sobie serię pytań, weryfikujących metodologię działania[129]:

- Co wydaje mi się, że wiem?
- Dlaczego wydaje mi się, że wiem?
- Kiedy to co wiem, może nie być prawdą?

Odpowiedzi na te pytania, a także świadomość faktu, że wszyscy są podatni na błędy poznawcze, a także że one nadal oddziałują na sposób myślenia pomimo świadomości ich istnienia i bycia podatnym, ma szansę sprawdzić poprawność toku rozumowania i pewność wyciąganych wniosków.

[129] B. Brown, *Cognitive Bias and Critical Thinking in Open Source Intelligence (OSINT)*, wystąpienie podczas Circle City Con 2014 – https://www.youtube.com/watch?v=bWjEgd-KSHY – dostęp online 19.10.2022 r.

Podatność ludzkiego umysłu na błędy poznawcze, może także w pewnych przypadkach być wykorzystana do obrony, tzn. takiego sterowania wynikami rozpoznania, prowadzonego w stosunku do jakiejś osoby lub organizacji, aby prowadzący rozpoznanie doszli do mylnych wniosków, a co za tym idzie – nie zdobyli planowanych informacji o celu. Dodatkowym atutem sprawnie przeprowadzonej akcji mylenia przeciwnika, wykonującego rozpoznanie otwartoźródłowe, będzie jego przekonanie, że posiada on poprawne informacje, co utrudni mu późniejsze przeprowadzenia ataku w stosunku do danej osoby czy organizacji. Zawsze jednak istnieje niebezpieczeństwo, że działania takie zostaną wykryte przez prowadzącego rozpoznanie, dlatego też ewentualne wykorzystanie tej techniki powinno być poprzedzone gruntowym planowaniem i przygotowaniem scenariuszy, mających na celu zmylenie rozpoznającego.

ROZDZIAŁ 4

WYKORZYSTANIE WYWIADU OPARTEGO NA OTWARTYCH ŹRÓDŁACH W ZAKRESIE BEZPIECZEŃSTWA SYSTEMÓW TELEINFORMATYCZNYCH ORAZ BEZPIECZEŃSTWA OSOBOWEGO I BIZNESOWEGO

Uwagi wstępne

W obecnych czasach coraz większa ilość informacji, zarówno dotyczących działań w sferze prywatnej, jak i służbowej, jest przechowywana oraz często możliwa do odnalezienia w Internecie. Sprzyja temu zarówno rozwój technologiczny, jak i chęć dzielenia się informacjami, dotyczącymi życia zarówno osobistego, jak i służbowego na portalach społecznościowych, blogach oraz w korespondencji przesyłanej drogą elektroniczną. To także dotyczy informacji odnoszących się do systemów teleinformatycznych, które przez swój styk z Internetem i wykorzystywanie ogólnodostępnych mechanizmów zdobywania szczegółowych informacji o tych systemach (jak np. mechanizm przejrzystości certyfikatów SSL/TLS – ang. *certificate transparency*), daje duże możliwości ich otwartoźródłowej analizy. Wiele osób i organizacji nie zdaje sobie sprawy ze znaczenia tych informacji i możliwości wykorzystania ich w celu naruszenia bezpieczeństwa osobowego lub firmowego. Wraz z ciągłym rozszerzaniem się bazy danych, jaką jest niewątpliwie Internet, rośnie znaczenie umiejętności wyszukiwania i analizy zawartych tam informacji. Dzięki odpowiednim narzędziom i technikom OSINT-owym, możliwe jest przygotowanie przez przestępców (lub inne grupy o charakterze ofensywnym, np. haktywistów[130]) ataków na pojedyncze osoby lub całe organizacje. Przykładem mogą być tutaj ataki socjotechniczne, które polegają na takiej manipulacji, aby przekonać atakowaną osobę poprzez e-mail lub telefon (w formie rozmowy lub SMS-ów) do wykonania pożądanych działań, mających

[130] Haktywista – potocznie: haker działający z pobudek społecznych, dla pożytku publicznego – za: Słownik Języka Polskiego, sjp.pl.

przynieść zysk atakującemu. Ataki socjotechniczne prowadzone są także ramach fizycznego atakowania firm lub innych organizacji poprzez wejście na ich teren i kontakt bezpośredni z osobami, które poddawane są manipulacji, w celu umożliwienia dostępu do pożądanych miejsc. W celu skutecznego przeprowadzenia ataku socjotechnicznego, którego podstawą są podatności nie systemów informatycznych, ale ludzkiego mózgu, konieczne jest w pierwszej kolejności zebranie jak największej ilości szczegółowych informacji o celu, które później przekuwane są w odpowiednie „legendy" (ang. *pretext*), czyli historie, które mają oszukać atakowanego i przekonać go do wypełnienia zamierzonych przez atakującego zamierzeń.

1.20. Zagrożenia dla bezpieczeństwa infrastruktury teleinformatycznej

Zagrożenia, dotyczące infrastruktury teleinformatycznej, wynikające z możliwości jej analizy poprzez techniki OSINT-owe, nie zawsze muszą wiązać się ze złą konfiguracją elementów tejże infrastruktury czy nadmiarowego udostępniania informacji. W wielu przypadkach zbierane informacje są i muszą być ogólnodostępne, ponieważ wynika to z zasad działania serwisów internetowych i urządzeń podłączonych do sieci. Rekonesans sieciowy należy zatem traktować na dwóch płaszczyznach: informacji koniecznych i informacji nadmiarowych. Te pierwsze zazwyczaj są drobnymi elementami, branymi pod uwagę podczas analizy infrastruktury i nie dają możliwości przeprowadzenia prostego ataku – należy jedynie mieć ich świadomość. Do tego rodzaju informacji możemy zaliczyć dane pochodzące z rejestrów certyfikatów cyfrowych, informacje zapisane w rejestrach DNS, dane o oprogramowaniu wykorzystywanym w ramach serwisów internetowych czy adresację IP. Dane, których obecności i dostępności powinniśmy unikać, są informacje nadmiarowe, które mogą wynikać ze złej konfiguracji urządzeń i oprogramowania, wycieków lub kradzieży danych dostępowych i konfiguracyjnych, a także braku nadzoru nad udostępnianymi w Internecie informacjami przez pracowników lub ich najbliższych. Mogą one stanowić bezpośrednie zagrożenie dla bezpieczeństwa infrastruktury, ponieważ zawierają wskazówki, które z konfiguracyjnego punktu widzenia bezpieczeństwa nie powinny być upubliczniane.

Do grupy informacji, których ogólnodostępność jest z góry założona i konieczna do prawidłowego funkcjonowania założonych mechanizmów w ramach usług dostępnych w Internecie są certyfikaty cyfrowe SSL/TLS. To na podstawie ich dostępności budowane jest zaufanie do danego serwisu, jednak w związku z dostępną historią ich wystawiania, a także dostępności informacji o subdomenach, dla których są wystawiane, możliwe jest zbudowanie informacji o nieudostępnionych publicznie serwisach, które być może nadal działają, a których zabezpieczenia nie są na tyle dojrzałe, żeby oprzeć się atakom. Przykładem takich serwisów mogą być testowe wersje serwisów internetowych oraz usługi wyłącznie do użytku przez pracowników lub współpracowników danej organizacji (przykładowe zestawienie takich subdomen przedstawiono na Rys. 13). Z nazewnictwa samych subdomen także możliwe jest wywnioskowanie jakich informacji i jakiego oprogramowania można się pod danym adresem spodziewać, np. dla certyfikatu, wystawionego dla subdomeny „testpanel.domena.com”, można oczekiwać, że pod tym adresem znajdzie się niegotowa jeszcze do wdrożenia wersja panelu zarządczego,

a w przypadku subdomeny „mysql.domena.com” można spodziewać się serwera baz danych. Dlatego w przypadku udostępniania serwisów testowych w Internecie, należy wziąć pod uwagę, że ich zabezpieczenia powinny być na tyle dojrzałe, żeby były one gotowe na weryfikację ich przez użytkowników, którzy będą próbowali uzyskać do tych serwisów dostęp.

39361701	2016-10-02	2013-01-15	2014-01-15	nwk1-g10-vpn-asa3.apple.com	nwk1-g10-vpn-asa3.apple.com	C=US, O=Apple Inc., OU=Apple IST Certification Authority, CN=Apple IST CA 1
39353239	2016-10-02	2014-02-12	2014-07-07	uklon5-asavpn-fw2.euro.apple.com	uklon5-asavpn-fw2.euro.apple.com	C=US, O=Apple Inc., OU=Apple IST Certification Authority, CN=Apple IST CA 1
39342723	2016-10-02	2014-03-05	2014-07-07	ivpn.apple.com	ivpn.apple.com	C=US, O=Apple Inc., OU=Apple IST Certification Authority, CN=Apple IST CA 1
39338500	2016-10-02	2009-07-13	2012-09-14	plmdfmtest.asia.apple.com	plmdfmtest.asia.apple.com	C=US, O="Entrust, Inc.", OU=AND ADDITIONAL TERMS GOVERNING USE AND RELIANCE, OU=CPS CONTAINS IMPORTANT LIMITATIONS OF WARRANTIES AND LIABILITY, OU=www.entrust.net/CPS is incorporated by reference, OU="(c) 2008 Entrust, Inc.", CN=Entrust Certification Authority - L1B
39334538	2016-10-02	2011-08-11	2013-08-10	st13p03sa.apple.com	st13p03sa.apple.com	C=US, O="VeriSign, Inc.", OU=VeriSign Trust Network, OU=Terms of use at https://www.verisign.com/rpa (c)06, CN=VeriSign Class 3 Extended Validation SSL SGC CA
39332407	2016-10-02	2012-07-09	2014-05-31	gsp16-ssl.ls.apple.com	gsp16-ssl.ls.apple.com	C=US, O="Entrust, Inc.", OU=www.entrust.net/rpa is incorporated by reference, OU="(c) 2009 Entrust, Inc.", CN=Entrust Certification Authority - L1C
39321688	2016-10-02	2012-09-29	2013-09-28	mrlnbackgroundcheck-staging.apple.com	mrlnbackgroundcheck-staging.apple.com	C=US, O="Entrust, Inc.", OU=www.entrust.net/rpa is incorporated by reference, OU="(c) 2009 Entrust, Inc.", CN=Entrust Certification Authority - L1C
39321309	2016-10-02	2012-09-20	2013-09-27	attachet.apple.com	attachet.apple.com	C=US, O="Entrust, Inc.", OU=www.entrust.net/rpa is incorporated by reference, OU="(c) 2009 Entrust, Inc.", CN=Entrust Certification Authority - L1C
39318266	2016-10-02	2014-02-21	2014-07-07	uklon5-asavpn-fw1.euro.apple.com	uklon5-asavpn-fw1.euro.apple.com	C=US, O=Apple Inc., OU=Apple IST Certification Authority, CN=Apple IST CA 1
39298941	2016-10-02	2012-07-10	2013-07-10	remoteadvisor.apple.com	ara.apple.com arade.apple.com araes.apple.com arafr.apple.com arait.apple.com araja.apple.com arajp.apple.com aranl.apple.com arapt.apple.com jvs.apple.com remoteadvisor1.apple.com remoteadvisor2.apple.com remoteadvisor.apple.com sws.apple.com	C=US, O="VeriSign, Inc.", OU=VeriSign Trust Network, OU=Terms of use at https://www.verisign.com/rpa (c)06, CN=VeriSign Class 3 Extended Validation SSL SGC CA

Rys. 13 – wycinek zestawienia subdomen apple.com, dla których wystawiono certyfikaty. Widoczne nazewnictwo sugerujące m.in. punkty dostępu do usług VPN lub środowisko przedprodukcyjne (staging).
Źródło: crt.sh – dostęp online 04.11.2022 r.

Sposobem enumeracji, czyli pozyskania jak największej ilości subdomen możliwych do pozyskania z otwartych źródeł, jest także wykorzystanie narzędzi i technik przedstawionych w rozdziale 1.13. Umożliwiają one zdobywanie informacji w sposób zarówno pasywny, jak i aktywny. Jednak o ile pozyskanie listy subdomen zaliczyć można bardziej do rekonesansu sieciowego, wykonywanego w ramach testów penetracyjnych, o tyle zdobywanie informacji skojarzonych z domenami, stanowi istotny element rozpoznania otwartoźródłowego w przypadku firm. To właśnie rodzajem danych, których zakres i szczegółowość należy mieć na uwadze, są dane powiązane z rekordami DNS. W przypadku domen polskich, dostępność danych osoby rejestrującej daną domenę uzależniona jest od tego czy rejestracja dokonywana jest na osobę prywatne czy na firmę. W pierwszym przypadku dane osobowe abonenta domeny są niewidoczne w rejestrach WHOIS, gdyż nie pozwalają na to przepisy prawa o ochronie danych osobowych, natomiast w drugim przypadku, kiedy podmiotem rejestrującym jest firma, dane są w pełni widoczne, aby „zapewnić należyte poszanowanie praw użytkowników Internetu,

właścicieli znaków towarowych, praw autorskich i innych dóbr prawnie chronionych”[131]. W przypadku domen globalnych dane abonenta są widoczne, niezależnie od tego czy jest osobą fizyczną czy firmą. Daje to możliwości zdobycia informacji o osobach pracujących lub powiązanych z daną organizacją, ich adresów e-mail i numerów telefonów, a także techniczne informacje, takie jak wykorzystywane serwery nazw. Istnieje co prawda możliwość ukrycia danych abonenta domeny globalnej, jednak informacją, która zawsze pozostaje jawna, jest adres e-mail abonenta.

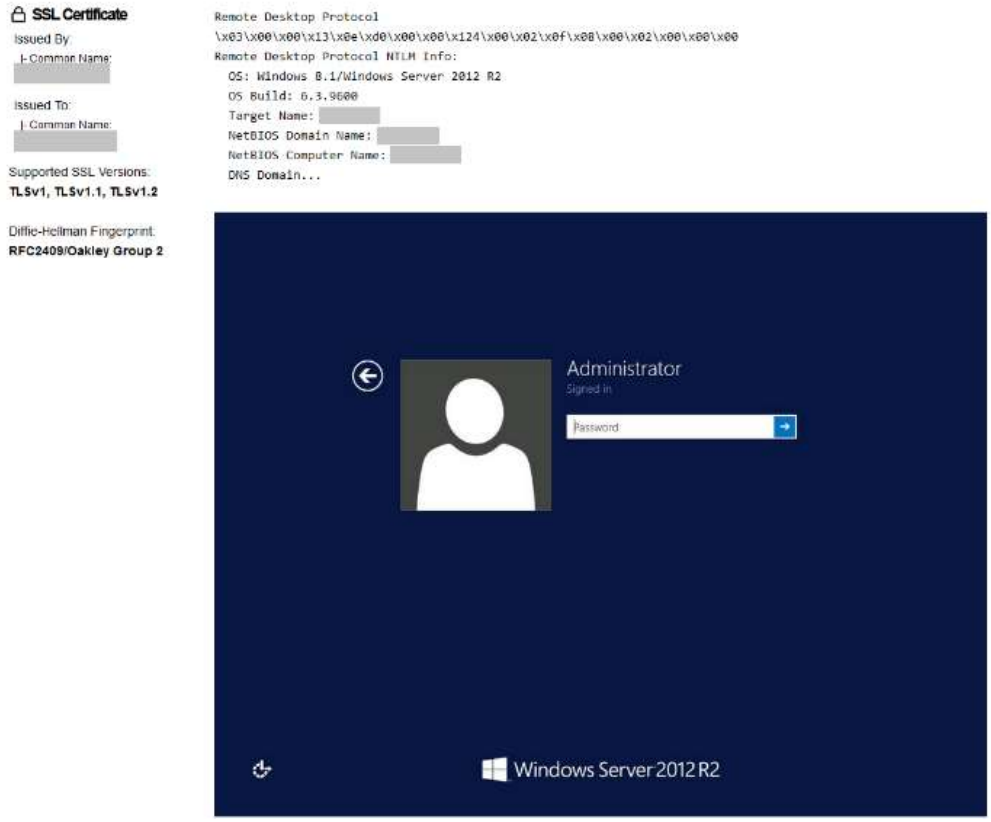
Znaczenie informacji technicznych, związanych z analizowaną domeną, także może być elementem rozpoznania OSINT-owego. Dzięki analizie adresu IP, który jest skojarzony z domeną, możliwe jest określenie firmy hostującej oraz lokalizacji odpowiadającej danemu adresowi – z dużym prawdopodobieństwem, ale bez pewności, gdyż powiązanie miejsca z adresem IP zawsze może dawać błędne wyniki. Ilość informacji dostępnych na podstawie adresu IP może być także bardzo duża. Przykładem śledztwa OSINT-owego, które wykorzystało tego typu dane, jest poszukiwanie informacji o niemieckiej służbie Bundesservice Telekommunikation, gdzie po odpytaniu bazy RIPE możliwe było nie tylko uzyskanie danych osobowych, numerów telefonu, faksu oraz adresu e-mail. Dzięki poprawnej analizie otrzymanych wyników autorka śledztwa doszła do wniosku, że dane osobowe wyglądają na sztucznie generowane, a konstrukcja adresu e-mail wskazuje na brak zgodności z typowymi adresami administracji niemieckiej, co z kolei dawało podejrzenie, że jednak dane te nie należą do typowej komórki organizacyjnej. Tym bardziej, że domena, w której były założone skrzynki e-mail także nie miała typowej konstrukcji, a o jej istnieniu nie wiedzieli nawet pracownicy rządowego CERT-u, czyli komórki odpowiedzialnej za monitorowanie cyberbezpieczeństwa w tym obszarze.

Niebezpieczeństwem, szczególnie związanym z infrastrukturą, jest wykrywanie błędnie skonfigurowanych serwerów lub innych urządzeń (np. IoT), podłączonych do Internetu. Za pomocą wyszukiwarek, takich jak Shodan lub Zoomeye, możliwe jest wychwycenie elementów konfiguracji, które wskażą podatne elementy w sieci. Mogą być to przykładowo systemy operacyjne z wystawioną publicznie usługą zdalnego pulpitu, przez co możliwe jest w najlepszym przypadku zebranie informacji o wersjach systemów

[131] NASK – Krajowy Rejestr Domen – WHOIS FAQ – https://www.dns.pl/whois/FAQ – dostęp online 07.11.2022 r.

operacyjnych i nazw użytkowników, a w najgorszym – możliwość uzyskania dostępu do tych systemów poprzez słabo zabezpieczone konta użytkowników lub administratorów.

Rys. 14 – Wyniki wyszukiwania interfejsów graficznych, dostępnych na publicznych adresach IP – ilości i rozkład na świecie oraz najczęściej wykrywane porty. Źródło: shodan.io – dostęp online 17.11.2022 r.



Rys. 15 – Przykład wyniku wyszukiwania, przedstawiającego panel logowania do maszyny z systemem Microsoft Windows Server (zasłonięto dane, mogące identyfikować maszynę oraz znalezione organizacje). Źródło: shodan.io – dostęp online 17.11.2022 r.

Wyszukanie w Internecie informacji, dotyczących rodzaju i wersji wykorzystywanego oprogramowania, wraz z możliwością znalezienia ich podatności i nierzadko gotowych *exploitów*, czyli opracowanych sposobów na wykorzystanie tych podatności w celu przeprowadzenia skutecznego ataku na dany komponent, powoduje, że publikowanie tego typu informacji jest niewskazane. W bazie skatalogowanych rodzajów słabości i luk sprzętu i oprogramowania – MITRE CWE[132], istnieje opis dla tego rodzaju luki: CWE-205 „Observable Behavioral Discrepancy”[133], w opisie której zaznaczono, że w idealnym przypadku produkty powinny udostępniać tak mało informacji o swoim działania, jak to tylko możliwe. W przeciwnym przypadku atakujący mogą wykorzystać te informacje do usprawnienia lub zoptymalizowania swoich ataków. Dotyczy to także nazwy i wersji oprogramowania.

[132] MITRE Common Weakness Enumeration – https://cwe.mitre.org
[133] CWE-205 – https://cwe.mitre.org/data/definitions/205.html – dostęp online 22.11.2022 r.

Innym przykładem niebezpieczeństw związanych z możliwymi do wyszukiwania kanałami dostępu jest bezpośrednie podłączenie do Internetu m.in. drukarek, kserokopiarek wielofunkcyjnych, sterowników przemysłowych, elementów „inteligentnych domów", paneli sterowania maszynami i większymi elementami infrastruktury, często z domyślnymi danymi logowania lub bez jakichkolwiek zabezpieczeń. Do niebezpiecznej sytuacji mogło dojść w na początku 2021 roku, kiedy to włamywacz dostał się do słabo zabezpieczonego systemu komputerowego w stacji uzdatniania wody na Florydzie[134]. Kanał dostępu został przez niego prawdopodobnie zidentyfikowany poprzez Shodan lub podobne narzędzie. Hasło do logowania do systemu z wykorzystaniem oprogramowana Team Viewer, umożliwiające zdalny dostęp do maszyn w sieci, uzyskał natomiast z dostępnych w Internecie baz wycieków danych[135]. Dodatkowymi problemami w tym systemie był brak jakiegokolwiek firewalla i jednakowe hasła dostępu, co nie powinno mieć miejsca z punktu widzenia zabezpieczeń infrastruktury. Haker zwiększył zawartość wodorotlenku sodu ponad stukrotnie, co na szczęście zostało wychwycone bardzo szybko przez operatora systemu, który ustawił z powrotem prawidłowe wartości. Z kolei na początku 2022 roku, także za pomocą podobnych narzędzi, służących do rekonesansu w Internecie, nastolatkowi z Niemiec udało się wyszukać interfejs oprogramowania TeslaMate, służącego do podglądu parametrów samochodów Tesla[136]. Po przeskanowaniu otwartych portów i zidentyfikowaniu kanału dostępu, umożliwiającego wysyłanie dowolnych zapytań do danego samochodu, możliwe było uzyskanie klucza API Tesli, dającego dostęp m.in. do funkcji włączania świateł i klaksonu, otwarcia samochodu oraz do jego uruchomienia przy pomocy funkcji Keyless Driving. Do uzyskania wspomnianego klucza API konieczne było podanie loginu i hasła, jednak okazało się, że dostęp był możliwy z wykorzystaniem bardzo popularnej kombinacji admin:admin.

Pozostawianie domyślnych danych dostępowych do urządzeń podłączonych do sieci jest kolejnym, obok tych samych haseł do wielu punktów dostępu, przewinieniem

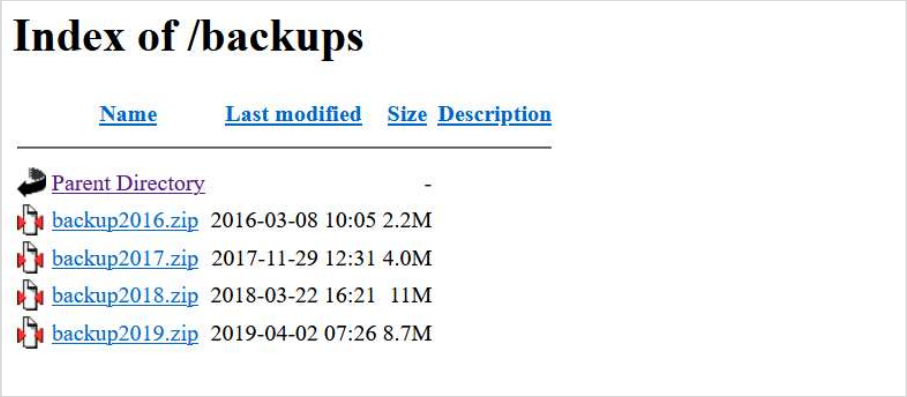
[134] Zapis wideo z konferencji prasowej w temacie włamania do stacji uzdatniania wody na Florydzie – https://www.youtube.com/watch?v=MkXDSOgLQ6M – dostęp online 14.11.2022 r.
[135] *Przez TeamViewer-a aż do systemu kontroli uzdatniania wody. Hacker zmienił parametry chemiczne wody.* – https://sekurak.pl/przez-teamviewer-a-az-do-systemu-kontroli-uzdatniania-wody-hacker-zmienil-parametry-chemiczne-wody – dostęp online 14.11.2022 r.
[136] *Nastolatek uzyskał zdalny dostęp do ponad 25 Tesli na całym świecie. Lokalizowanie samochodów, otwieranie drzwi, wystartowanie auta…* – https://sekurak.pl/nastolatek-uzyskal-zdalny-dostep-do-ponad-25-tesli-na-calym-swiecie-lokalizowanie-samochodow-otwieranie-drzwi-wystartowanie-auta – dostęp online 14.11.2022 r.

w zakresie zabezpieczania infrastruktury. Wiele urządzeń jest możliwych do zaatakowania w ten sposób, ponieważ domyślne loginy i hasła dostępowe albo sposoby ich generowania są obecne w instrukcjach w Internecie. Temat ten dotyczy m.in. urządzeń sieciowych, punktów dostępu bezprzewodowego, drukarek i urządzeń wielofunkcyjnych, sprzętu IoT oraz kamer CCTV. Dzięki serwisom takim jak Insecam.org możliwe jest znalezienie kamer z otwartym dostępem. Także poprzez odpowiednie zapytania w wyszukiwarce Google, Shodan albo Zoomeye można zidentyfikować kamery, które można w kolejnych krokach zbadać pod kątem możliwości nieautoryzowanego dostępu do nich. Nawet jeśli okaże się, że dostęp jest zabezpieczony hasłem, to w wielu przypadkach hasła są pozostawiane przez użytkowników w formie domyślnej, co umożliwia podgląd, a czasami także i posłuch – jeśli kamera jest wyposażona także w mikrofon.

Szczegóły konfiguracyjne serwerów www wystawionych w Internecie lub intranecie także mogą zostać wykorzystane w ramach rozpoznania otwartoźródłowego do znalezienia nieprawidłowości, mogących mieć wpływ na obniżenie poziomu bezpieczeństwa infrastruktury sieciowej. Jednym z możliwych uchybień jest konfiguracja serwerów, umożliwiająca przeglądanie ich zawartości poprzez przeglądarkę internetową. Dzieje się to wówczas, gdy brakuje zabezpieczeń przeciwko tzw. listingowi plików na serwerze. Można temu zapobiec poprzez ustawienia konfiguracyjne dla danego katalogu głównego danej domeny lub każdego podkatalogu w niej dostępnego. Nawet w przypadku braku takiego zabezpieczenia konfiguracyjnego możliwe jest, że nie uda się wyświetlić listy katalogów i plików w badanym serwisie, kiedy w danym katalogu obecny będzie np. plik index.html lub index.php, których zawartość zostanie pobrana domyślnie przez serwer i wyświetlona użytkownikowi zamiast zawartości danego katalogu. Lepsze jednak jest zawsze podejście systemowe, zapewniające odpowiednią konfigurację dla wszystkich katalogów w ramach danego serwisu www. To zagrożenie jest tym poważniejsze, że możliwość wyszukania tego typu błędnie skonfigurowanych serwisów daje zwykła wyszukiwarka, jak Google czy Bing, gdyż w przypadku typowego listingu zawartości katalogu przez serwer www (np. Apache) zawiera ona frazę „Index of” przed nazwą wyświetlanego katalogu. Zatem użycie zapytania:

intext:”Index of”

wskaże wszystkie wyniki, które na wyświetlanej stronie mają tego typu ciąg znaków, a zatem najprawdopodobniej posiadają konfigurację, umożliwiającą tego typu dostęp do zasobów danego serwera.



Rys. 16 – Przykładowy listing plików w niezabezpieczonym przed tą formą dostępu katalogu. Źródło: opracowanie własne.

Dodatkowo, możliwe jest modyfikowanie zapytania, aby uzyskać bardziej efektywne wyniki wyszukiwania i tak na przykład fraza zapytania:

intext:”Index of” inurl:”wp-content/downloads”

wskaże wszystkie serwisy, stworzone przy użyciu popularnego oprogramowania Wordpress (na co wskazuje fragment ścieżki wp-content/downloads, unikalna dla tego oprogramowania). Oczywiście konfiguracja umożliwiająca listowanie katalogów i plików nie jest zabroniona z punktu widzenia bezpieczeństwa, jednak należy mieć pełną świadomość i kontrolę nad tym, które zasoby są w ten sposób udostępnione.

Zasoby, dostępne w ramach serwisów www oraz inne informacje, które mogą okazać się cenne z punktu widzenia OSINT-u, mogą znajdować się w dwóch plikach o typowych nazwach: robots.txt oraz security.txt, które często znajdują się w głównych katalogach serwisów internetowych. Przy poprawnej konfiguracji bezpieczeństwa nie powinny one stanowić zagrożenia, jednak niekiedy okazuje się, że umieszczone w nich informacje dają pogląd na konstrukcję danej infrastruktury lub dostęp do nienależycie zabezpieczonych danych. Plik robots.txt, który zawiera w sobie informacje dla botów indeksujących, może zarazem przekazywać informacje o zasobach, których administrator danej witryny nie chce publikować, a jednak samo umieszczenie ich w tym pliku daje osobom analizującym dany serwis www wskazówki, gdzie można szukać nienależycie zabezpieczonych danych. Typowym przykładem takiego zasobu może być ścieżka /backup, która nie powinna być indeksowana ze względu na fakt, że przeciętni

użytkownicy danej witryny nie powinni podczas wyszukiwania otrzymać informacji, że taki zasób istnieje, a jednocześnie w przypadku złego zabezpieczenia takiego katalogu (np. opisywanego wyżej braku wyłączenia listowania katalogu oraz braku odpowiednich zabezpieczeń), może dać dostęp niepowołanym osobom do przechowywanych tam danych, w tym przypadku – kopii bezpieczeństwa. Plik security.txt, który jest proponowanym standardem, umożliwiającym publikowanie polityki bezpieczeństwa witryny, mającej ułatwiać badaczom zgłaszanie luk w jej zabezpieczeniach. W domyślnej konfiguracji nie stanowi zagrożenia dla bezpieczeństwa, jednak zawarte w nim informacje mogą stanowić jeden z elementów zdobywania wiedzy o danej organizacji.

Niekiedy nieprawidłowo skonfigurowane serwisy internetowe, zawierające nadmiarowe dane, mogące mieć wpływ na bezpieczeństwo danej witryny lub szerszego zakresu infrastruktury, zostają zastąpione poprawną wersją danego serwisu, jednak zawsze należy pamiętać, że istnieją serwisy, przechowujące archiwalne wersje stron, np. Internet Archive[137]. Jeśli nie jest możliwe usunięcie zarchiwizowanych wersji serwisów z widocznymi uchybieniami, należy wprowadzić zabezpieczenia przed potencjalnymi atakami, wykorzystującymi te informacje.

[137] https://archive.org

1.21. Zagrożenia dla bezpieczeństwa osobowego

Internet, jako medium komunikacji międzyludzkiej oraz przechowywania danych, stanowi coraz większą część życia codziennego wielu osób, a zatem informacje, które są w nim umieszczane powinny być też chronione w taki sam sposób, jak fizyczna własność. Tymczasem poprzez zjawisko *oversharingu*, czyli nadmiernego udostępniania informacji w sieci, bezpieczeństwo osobowe staje się bardziej zagrożone. Media społecznościowe stanowią ogromną bazę informacji osobistych, jak zdjęcia, filmy, opisy planowanych działań czy preferencje w różnych dziedzinach. Brak odpowiednich ustawień bezpieczeństwa w serwisach społecznościowych umożliwia dostęp do prywatnych danych osobom nieuprawnionym, z których część może wykorzystać je do działań niezgodnych z prawem. Profile użytkowników, którzy nie ograniczyli dostępności publikowanych przez siebie danych do grona znajomych, czyli osób śledzących ich profil (lub jeszcze lepiej jedynie ograniczonej liczby zaufanych osób), dzielą się tak naprawdę swoimi danymi ze wszystkimi w Internecie. Także coraz częściej to rodzice rozpoczynają umieszczanie zdjęć, filmów i innych informacji na temat swojego dziecka w Internecie, co może w przyszłości mieć wpływ na jego bezpieczeństwo i komfort działania w sieci. Według opublikowanych w 2020 informacji[138], około 40% polskich rodziców zamieszcza w Internecie treści dotyczące ich dzieci. Zjawisko takie określane jest mianem „*sharenting*" (od złożenia angielskich słów „*share*" – dzielić się i „*parent*" – rodzic).

W ramach grup w mediach społecznościowych udostępniane są niekiedy także dane osobowe w postaci zdjęć dowodów osobistych, praw jazdy oraz innych dokumentów zawierających różne wrażliwe dane. Nie zawsze jest to wynikiem działania w celu zaszkodzenia innej osobie, a jedynie np. chęcią oddania dokumentów osobie, która je zgubiła. Niestety brak świadomości konsekwencji takiego działania, które mogą obejmować chociażby próby wzięcia kredytu na opublikowane dane, powoduje duże zagrożenie dla bezpieczeństwa osób, do których dane dokumenty należą. Także zdjęcia innych obiektów, jak np. znalezione klucze do samochodu lub mieszkania, mogą stanowić pewne zagrożenie w przypadku zidentyfikowania ich pochodzenia i próby dorobienia duplikatów kluczy jedynie na podstawie zdjęcia.

Dokumentami, które także niekiedy publikowane są nadmiarowo w Internecie, są bilety lotnicze. Zawierają one nie tylko informacje w formie możliwej do bezpośredniego

[138] https://www.gov.pl/web/niezagubdzieckawsieci/sharenting – dostęp online 14.11.2022 r.

odczytania przez człowieka, ale także kody (paskowe i 2-wymiarowe), zawierające dodatkowe informacje o posiadaczu biletu lub szczegółach lotu. Tego typu informacje często umożliwiają zalogowanie się w serwisie przewoźnika i uzyskanie dodatkowych informacji o osobowych, danych o preferencjach wybranych podczas lotu (np. rodzaju posiłków lub konieczności wsparcia dla osoby niepełnosprawnej). Głośną sprawą pokazującą zakres możliwych do uzyskania danych na podstawie jednego zdjęcia była sprawa byłego premiera Australii, Tony'ego Abbotta, który opublikował na Instagramie właśnie fotografię swojego biletu lotniczego. Na podstawie odczytanych z biletu danych możliwe było uzyskanie dostępu do serwisu linii lotniczych, w którym oprócz możliwych do odczytania danych identyfikacyjnych posiadacza biletu, możliwe było także uzyskanie dodatkowych danych, takich jak np. jego data urodzenia, ze względu na uchybienia bezpieczeństwa w zakresie umieszczania ich w kodzie strony internetowej rezerwacji i zabezpieczenie jedynie poprzez brak ich wyświetlania w graficznej reprezentacji strony.

Innymi informacjami osobistymi, których publikacja może mieć negatywny wpływ na poziom bezpieczeństwa osobowego, jest nadmiarowe udostępnianie informacji dotyczących planów wyjazdów wakacyjnych, które mogą zostać wykorzystane przez niepowołane osoby do próby włamania do mieszkania danej osoby. Oczywiście w tym celu konieczne jest także ustalenie adresu danej osoby, jednak może to być możliwe na wiele sposobów – od analizy wpisów na grupach dyskusyjnych dotyczących lokalnych społeczności po analizę miejsca wykonania zdjęć, pokazujących mieszkanie danej osoby lub widok z jej okna. Także publikowanie zdjęć z wakacji może mieć wpływ na poziome bezpieczeństwa, jednak w tym przypadku częściej ma to miejsce w przypadku osób poszukiwanych oraz ich partnerów, którzy z dużą dozą beztroski dzielą się interesującymi zdjęciami z miejsc odległych zazwyczaj od kraju, w którym dana osoba jest poszukiwana. Dzięki analizie takich zdjęć udało się m.in. odnaleźć poszukiwanego Europejskim Nakazem Aresztowania mężczyznę, którego partnerka publikowała na Facebooku zdjęcia, pokazujące jednocześnie twarz poszukiwanego oraz widok miejsca, w którym się zatrzymali.[139]

Wzrost znaczenia narzędzi internetowych w życiu prywatnym powoduje także, że niektóre dane osobiste są przechowywane w chmurze, co samo w sobie jest bardzo

[139] *Poszukiwany za przekręt na 50 mln zł złapany, bo jego młoda partnerka chwaliła się życiem w mediach społecznościowych* – https://www.antyradio.pl/News/Poszukiwany-za-przekret-na-50-mln-zl-zlapany-bo-jego-mloda-partnerka-chwalila-sie-zyciem-w-mediach-spolecznosciowych-38083 – dostęp online 18.11.2022 r.

użyteczną funkcjonalnością. Wykorzystywanie jednak tego typu narzędzi bez świadomości zagrożeń, płynących z braku zabezpieczenia umieszczanych w chmurze danych przed dostępem osób nieuprawnionych. Przykładem mogą być tutaj serwisy do przechowywania dokumentów jak np. Docer[140], gdzie możliwe jest publikowanie dowolnych dokumentów, co sprawia, że możliwe są tam do znalezienia zeskanowane wersje dowodów osobistych, legitymacji, certyfikatów, świadectw, biletów oraz innych dokumentów, zawierających dane osobowe i inne wrażliwe informacje.

Brak świadomości osób, które udostępniają tego typu dokumenty w celu ich łatwiejszego udostępnienia instytucjom lub posiadania dostępu do nich w dowolnym momencie, powoduje efekt nadmiarowego udostępniania danych, co może mieć wpływ na obniżenie poziomy bezpieczeństwa tych osób. Podobnymi serwisami są tzw. pastebiny – schowki internetowe, w których można umieszczać dokumenty tekstowe w celu ich udostępnienia np. poprzez serwisy posiadające ograniczenia w ilości przesyłanych danych. Najbardziej znanym serwisem tego typu jest pastebin.com[141]. Zdarza się, że w publicznie dostępnych „schowkach" dostępne są całe książki adresowe, zawierające imiona, nazwiska, numery telefonów, a także adresy e-mail listy kontaktów.

Kontrolowanie zakresu dostępności danych, publikowanych w internetowych serwisach, jest konieczne w celu zapewnienia odpowiedniego poziomu bezpieczeństwa osobowego. Domyślne ustawienia zabezpieczeń profili w mediach społecznościowych, aplikacjach oraz innych serwisach mogą powodować udostępnianie swoich informacji publicznie. Cykliczne aktualizacje oprogramowania w zakresie paneli ustawień mogą także powodować bezwiedną zmianę poziomu bezpieczeństwa konta, stąd ich cykliczne przeglądanie przez użytkowników jest wskazane w celu utrzymania pożądanego poziomu zabezpieczeń.

[140] https://docer.pl
[141] https://pastebin.com

1.22. Zagrożenia dla bezpieczeństwa biznesowego i operacyjnego

Zagrożenia, wynikające z nadmiarowego publikowania nadmiarowych informacji, często bez świadomości ich wagi, można także rozpatrywać w zakresie bezpieczeństwa organizacji lub operacji (np. wojskowej), kiedy to właśnie ich dotyczą. Wiele informacji publikowanych jest samych pracowników (lub uczestników danej operacji) lub ich rodziny. W ramach programów prowadzonych w armii amerykańskiej, można spotkać się z poradnikami dla rodzin żołnierzy[142], wskazującymi czym jest bezpieczeństwo operacji i jakich zasad należy przestrzegać, aby nie narazić żołnierzy, przebywających na misji na dodatkowe niebezpieczeństwo. Amerykański Departament Obrony, podczas swoich szkoleń[143] w tym zakresie wskazuje m.in. przypadek przechwyconej w Iraku jednej z niezliczonych kopii poradnika, mówiącego o tym, że 80% informacji wywiadowczych dotyczących wojska może być uzyskane poprzez analizę legalnych źródeł, takich jak strony internetowe, blogi czy gazety.

Aplikacją, która zyskała niechlubną sławę w zakresie nadmiarowego udostępniania danych o znaczeniu wojskowym jest Strava[144]. Jest to pokłosie głośnej sprawy z 2018 roku, kiedy to udostępniona została mapa, pokazująca trasy przemieszczania się użytkowników tej aplikacji, służącej domyślnie do śledzenia swoich wyników, związanych z uprawianiem sportu (głownie bieganiem) i dzielenia się nimi z innymi użytkownikami tego oprogramowania. Na tzw. mapie termicznej (ang. *heatmap*) można zobaczyć które trasy są bardziej uczęszczane przez użytkowników Stravy, a które mniej. Niestety mapa ujawniła też dokładne obrysy i ścieżki poruszania się w bazach wojskowych na całym świecie oraz na terenie niektórych ambasad, gdyż wśród użytkowników aplikacji znajdowali się także żołnierze na służbie.

---

[142] Przykładami są tutaj: *OPSEC - A Guide For Family and Friends Presented by 1st Information Operations Command (Land) Vulnerability Assessment Detachment Army OPSEC Support Element* – https://www.usace.army.mil/Portals/2/docs/Careers/OPSEC-A%20Guide%20for%20Families%20and%20Friends.pdf (dostęp online 18.11.2022 r.) lub *Operations Security (OPSEC) Guidance for Family Members* – https://www.2ndmardiv.marines.mil/Portals/47/Docs/Operations%20Security%20Guidance%20for%20Family%20Members.pdf (dostęp online 18.11.2022 r.), a także poradniki takie jak *OPSEC 101 for Military Spouses: Help Keep Your Family and Service Member Safe* – https://themilitarywifeandmom.com/opsec-for-military-spouses (dostęp online 18.11.2022 r.).

[143] *DoD OPSEC for Families* – https://www.slideshare.net/DepartmentofDefense/opsec-for-families – dostęp online 18.11.2022 r.

[144] *Fitness app Strava lights up staff at military bases* - https://www.bbc.com/news/technology-42853072 – dostęp online 18.11.2022 r.

Rys. 17 – widok obrysu bazy Bagram na mapie aplikacji Strava.
Źródło: https://www.strava.com/heatmap – dostęp online 24.04.2022 r.

Podobne rodzaje zagrożeń, wynikających z nadmiarowego ujawniania informacji przez aplikację społecznościową zostały w maju 2020 roku naświetlone przez grupę dziennikarzy śledczych – Bellingcat. Tym razem jednak źródłem były zdjęcia publikowane przez użytkowników aplikacji Untappd, której celem było dzielenie się wrażeniami z degustacji różnego rodzaju odmian piwa. Użytkownicy Untappd, wśród których, jak się później okazało, było także wielu przedstawicieli wojska, mogli do swoich relacji dodawać także lokalizację i zdjęcia marek wypitego piwa, na których w pewnym momencie oprócz butelek można było także dostrzec widoki z baz wojskowych, ich wyposażenie (w tym także m.in. myśliwce F-16), karty kredytowe oraz inne dokumenty o typowo wojskowym charakterze, które nie powinny być publikowane. Dodatkowym zagrożeniem okazało się ujawnianie lokalizacji i nazw baz wojskowych, z czego można było odczytać w jakich okresach i do jakich miejsc podróżował personel wojskowy w ramach ćwiczeń i misji. Z listy miejsc, które odwiedzili, można było także wytypować często odwiedzane miejsca o charakterze cywilnym, jak na przykład restauracje w pobliżu baz wojskowych. Także na podstawie zdjęcia, które wykonał i opublikował znudzony żołnierz amerykańskiej piechoty morskiej podczas ćwiczenia na pustyni Mojave, możliwe było zidentyfikowanie w jakiej jednostce służy i gdzie aktualnie

się ona znajduje[145]. W tym przypadku skończyło się jedynie na teoretycznym „unicestwieniu jego jednostki". Podobny przypadek miał miejsce podczas ćwiczenia francuskiej marynarki wojennej Polaris 21, gdzie pomimo ciszy radiowej, załoga jednego z okrętów znalazła się na tyle blisko brzegu, że możliwe było opublikowanie przez jednego z jej członków filmiku na portalu Snapchat. Zespół „czerwony", którego zadaniem było wirtualne atakowanie jednostek, zidentyfikował miejsce opublikowania filmiku za pomocą technik OSINT-owych i wysłał 14 rakiet, ponieważ zdobyte koordynaty były na tyle dokładne, że możliwe było zaatakowanie celu[146]. Przypadki te pokazują, że zagrożenia dla operacji, związane z nadmiarową publikacją zdjęć i filmów z misji lub ćwiczeń, stanowią realny i coraz poważniejszy problem, któremu należy aktywnie przeciwdziałać.

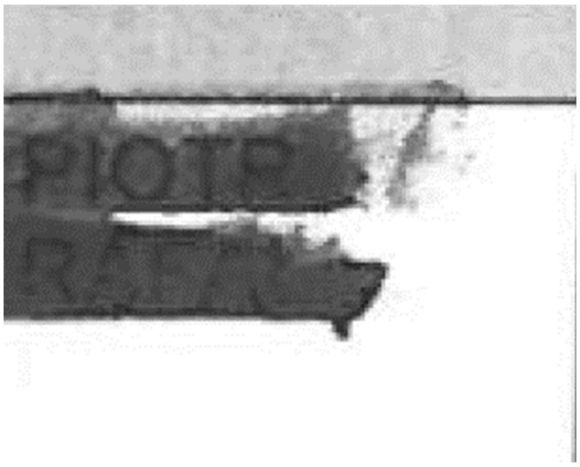
W przypadku zagrożeń dla bezpieczeństwa organizacji cywilnych, za publikowanie nadmiarowych informacji odpowiedzialni są w głównej mierze pracownicy, którzy na swoich prywatnych kontach udostępniają informacje, mogące mieć wpłynąć na bezpieczeństwo firmy, chociaż zdarzają się także przypadki, że nadmiarowe informacje udostępniane są na stronach i mediach społecznościowych samych firm, gdzie zapobiec temu powinny odpowiednie procedury. W przypadku prywatnych kont pracowników udostępniane mogą być np. informacje o rozpoczęciu pracy w nowej firmie, w ramach których mogą pojawić się zdjęcia nowego identyfikatora, który będzie wskazówką w przypadku przygotowywania się osób, mających na celu zaatakowanie firmy poprzez fizyczny atak z wykorzystaniem socjotechniki (zarówno w przypadku kontrolowanych testów bezpieczeństwa, jak i faktycznych ataków). Na podstawie pozyskanego wzoru identyfikatora przygotowywane są duplikaty, które pomimo braku posiadania funkcjonalności karty dostępowej (jeśli identyfikator taką spełnia), będą stanowiły elementy uwiarygadniający atakowane osoby jako pracowników danej firmy lub organizacji, co znacznie ułatwi im poruszanie się po jej terenie.

Nowi pracownicy, a zatem tacy, którzy nie posiadają jeszcze odpowiedniego przeszkolenia w zakresie ochrony informacji wrażliwych dotyczących swojej nowej pracy lub tacy, którzy nie przyswoili ich jeszcze na tyle dobrze, aby nie popełnić błędu

[145] *A Lance Corporal's Phone Selfie Got His Marine Unit 'Killed' at 29 Palms* – https://www.military.com/daily-news/2020/01/07/lance-corporals-phone-selfie-got-his-marine-unit-killed-29-palms.html – dostęp online 18.11.2022 r.
[146] Wpis na portalu Twitter: https://twitter.com/Marsattaqueblog/status/1582039213591040000 – dostęp online 18.11.2022 r.

w zakresie nadmiarowego udostępniania informacji, mogą także udostępniać zdjęcia z wnętrz pomieszczeń firmowych, ułatwiając tym samym zidentyfikowanie informacji pożądanych zarówno przez konkurencję (np. dotyczący projektów, którymi się aktualnie firma zajmuje), hakerów (w zakresie wykorzystywanych technologii oraz loginów i haseł zapisanych w widocznych miejscach) lub wspomniane już osoby, przygotowujące fizyczny atak socjotechniczny (w zakresie rozkładu pomieszczeń). Wśród zdjęć mogą znaleźć się także przypadkowo ujęte na nich dokumenty, zawierające dane wrażliwe lub poufne z punktu widzenia polityki firmy.

Umieszczanie w Internecie nadmiarowych informacji z dokumentów firmowych może odbywać się nie tylko poprzez wykonywanie ich zdjęć, ale także przez nieodpowiednie przygotowywanie dokumentów zeskanowanych lub wydrukowanych do formatu PDF. Dane, które powinny zostać zasłonięte na publikowanych dokumentach mogą być nieodpowiednio zaciemnione przed zeskanowaniem, np. poprzez użycie flamastra, który przed zeskanowaniem daje efekt tekstu niewidocznego dla nieuzbrojonego ludzkiego oka, natomiast po zeskanowaniu i konwersji do odcieni szarości efekt ten zanika na tyle, że możliwe jest bezproblemowe odczytanie zasłoniętego tekstu.

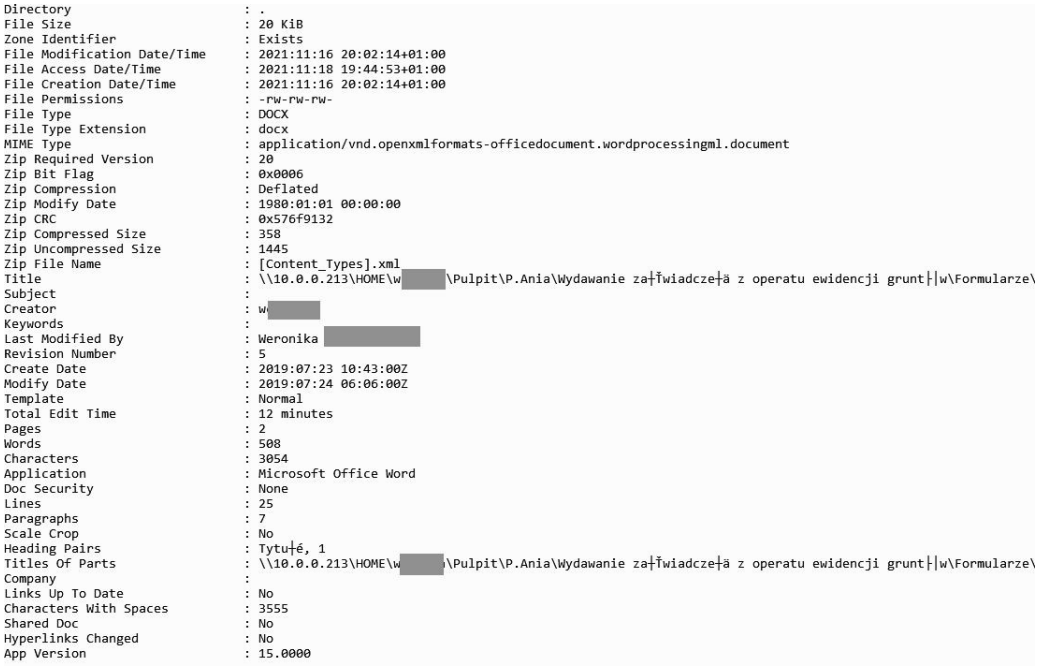


Rys. 18 – Przykład fragmentu dokumentu zaciemnionego przed zeskanowaniem, co nie uniemożliwia odczytania jego zawartości w formie cyfrowej. Źródło: Internet.

Zasłanianie danych wrażliwych może być także w podobny sposób nieodpowiednio wykonane, jeśli do zamazywania tekstu używany jest narzędzie flamaster w oprogramowaniu graficznym. Ma tutaj zastosowanie ta sama zasada – flamaster ten nie

pokrywa w 100% powierzchni, a jedynie daje złudzenie zamazanego tekstu, jednak jego odczytanie możliwe jest poprzez odpowiednią manipulację ustawieniami obrazu w oprogramowaniu graficznym, takimi jak np. kontrast czy intensywność. Innym nieodpowiednio używanym narzędziem, jest zasłanianie danych wrażliwych w plikach PDF poprzez nakładanie dodatkowej warstwy, zawierającej elementy geometryczne, przysłaniające dane na niższej warstwie. Poprzez edycję w ten sposób przygotowanego pliku PDF możliwe jest uzyskanie dostępu do zawartych w nim wszystkich danych, nawet tych przysłoniętych, gdyż warstwy nie są scalone ze sobą. Dlatego powinno się unikać wykorzystywania zarówno kolorowych flamastrów, jak i narzędzi niepokrywających w pełni spodniej warstwy do zaciemniania treści dokumentów przed ich odczytaniem przez osoby, mogące swobodnie pozyskać dany dokument z Internetu.

Zagrożenie publikacji nadmiarowych danych może stanowić także publikowanie danych bez uprzedniego wyczyszczenia ich metadanych, czyli zakresu danych zawierających często informacje o autorze danego dokumentu, wykorzystywanym oprogramowaniu, sprzęcie do skanowania i przygotowywania plików PDF, ścieżkach sieciowych i lokalnych, w których przechowywane były pliki w ramach infrastruktury firmowej lub innych danych, związanych z produkcją danego pliku. Opublikowanie tego typu danych może ujawnić dane pracowników, co jest kluczowe w przypadku publikowania dokumentów przez służby, a także może ujawnić potencjalnym atakującym, że w danej organizacji wykorzystywane jest nieaktualne oprogramowanie, posiadające luki bezpieczeństwa.



```
Directory                       : .
File Size                       : 20 KiB
Zone Identifier                 : Exists
File Modification Date/Time     : 2021:11:16 20:02:14+01:00
File Access Date/Time           : 2021:11:18 19:44:53+01:00
File Creation Date/Time         : 2021:11:16 20:02:14+01:00
File Permissions                : -rw-rw-rw-
File Type                       : DOCX
File Type Extension             : docx
MIME Type                       : application/vnd.openxmlformats-officedocument.wordprocessingml.document
Zip Required Version            : 20
Zip Bit Flag                    : 0x0006
Zip Compression                 : Deflated
Zip Modify Date                 : 1980:01:01 00:00:00
Zip CRC                         : 0x576f9132
Zip Compressed Size             : 358
Zip Uncompressed Size           : 1445
Zip File Name                   : [Content_Types].xml
Title                           : \\10.0.0.213\HOME\w      \Pulpit\P.Ania\Wydawanie za╪Ťwiadcze╪ä z operatu ewidencji grunt╪|w\Formularze\
Subject                         :
Creator                         : w
Keywords                        :
Last Modified By                : Weronika
Revision Number                 : 5
Create Date                     : 2019:07:23 10:43:00Z
Modify Date                     : 2019:07:24 06:06:00Z
Template                        : Normal
Total Edit Time                 : 12 minutes
Pages                           : 2
Words                           : 508
Characters                      : 3054
Application                     : Microsoft Office Word
Doc Security                    : None
Lines                           : 25
Paragraphs                      : 7
Scale Crop                      : No
Heading Pairs                   : Tytu╪é, 1
Titles Of Parts                 : \\10.0.0.213\HOME\w      \Pulpit\P.Ania\Wydawanie za╪Ťwiadcze╪ä z operatu ewidencji grunt╪|w\Formularze\
Company                         :
Links Up To Date                : No
Characters With Spaces          : 3555
Shared Doc                      : No
Hyperlinks Changed              : No
App Version                     : 15.0000
```

Rys. 19 – Przykład opublikowanego w Internecie dokumentu z nieusuniętymi metadanymi (fragmenty wrażliwe zaciemnione przez autora pracy). Źródło: Internet.

Nadmiarowe publikowanie danych nie musi być wyłącznie związane z brakiem odpowiedniego wyszkolenia kadry, ale także z celowym działaniem niezadowolonych pracowników (ang. *disgruntled employees*), którzy mogą chcieć zadziałać na niekorzyść danej firmy lub organizacji ze względu na nieodpowiednie traktowanie lub nieoczekiwane zwolnienie.

Wnioski

Wykorzystanie technik OSINT-owych daje duże możliwości zdobycia danych, publikowanych w Internecie, zarówno przez osoby prywatne, jak i firmy lub organizacje. Jednocześnie znajomość technik wychwytywania nadmiarowych informacji z ogólnodostępnych źródeł w połączeniu ze znajomością typowych ludzkich zachowań, wpływających na obniżenie poziomu bezpieczeństwa, może stanowić bazę do zbudowania bardzo szczegółowej mozaiki wywiadowczej, dzięki której możliwe będzie wykonanie skutecznego ataku, np. na dużą firmę.

Poprzez zwiększenie udziału elektronicznych kanałów komunikacji międzyludzkiej, a także administracyjnej, zwiększa się także powierzchnia możliwych do wykorzystania otwartych źródeł informacji. Zarówno publikowane prywatnie dane, jak i dokumenty służbowe, mogą nieść za sobą dane, które nie powinny być upubliczniane ze względu na możliwość ich łatwego przechwycenia i wykorzystania. Także coraz szersze wykorzystanie technologii w operacjach wojskowych oraz nieodłączność prywatnych smartfonów na polu walki daje powód do powzięcia decyzji w temacie ich wpływu na bezpieczeństwo zarówno misji, jak i pojedynczych żołnierzy.

Należy zaznaczyć, że wiele informacji, dostępnych w Internecie, a związanych z działalnością osób, firm i organizacji, a także konfiguracją infrastruktury, musi się tam znajdować, aby możliwe było działanie w ramach Internetu lub spełnienie wymagań prawnych – dotyczy to m.in. informacji o certyfikatach, danych rejestrowych, ogólnodostępnych kontekstowych baz danych, a także działania typowo marketingowe, bez których niemożliwe byłoby optymalne działanie w sferze ekonomicznej. Znalezienie kompromisu pomiędzy danymi koniecznymi do udostępnianie, a także tymi, które nie wpływają na bezpieczeństwo, a danymi, które mogą takie zagrożenie stanowić, leży w gestii organów, odpowiedzialnych za zapewnienie bezpieczeństwa danego podmiotu.

Przedstawione w niniejszym rozdziale przykłady udostępniania nadmiarowych danych należy zawsze traktować indywidualnie, gdyż dane, które w jednym przypadku będą kluczowe dla zapewnienia bezpieczeństwa (jak w przywołanym przykładzie francuskiego okrętu, z którego opublikowano film na portalu Snapchat), w innym będą zupełnie nieszkodliwe (np. w przypadku publikowania filmików ze znanych miejsc przez blogerów i celebrytów, którzy poszukują jak największego zasięgu swoich filmów),

a nawet mogą być pomocne (np. jeśli śledzimy poczynania bliskich osób, które wysyłają informacje na jakim etapie podróży się znajdują).

Poprzez OSINT możliwe jest zidentyfikowanie nadmiarowego udostępniania informacji i identyfikacja potencjalnych zagrożeń, którym należy przeciwdziałać, aby dane te nie stały się narzędziem w rękach niepowołanych osób, mogących zagrozić bezpieczeństwu osobistemu, operacyjnemu lub biznesowemu. Pomimo, że przywołane przypadki upubliczniania zbyt wielu informacji przez żołnierzy zostały tutaj przedstawione jako potencjalne zagrożenia, to poprzez fakt, że stało się to jedynie podczas ćwiczeń, należy spojrzeć na tę sytuację jako przyczynek do wprowadzenia lepszych szkoleń i zwiększenia poziomu świadomości kadry w zakresie możliwych zagrożeń, płynących z łatwości pozyskiwania ogromnej ilości danych z otwartych źródeł informacji.

Wskazane w niniejszym rozdziale zestawienie zagrożeń i technik ataków, wykorzystujących rozpoznanie otwartoźródłowe nie stanowią zamkniętej listy, a jedynie opracowane przez autora dysertacji zestawienie przekrojowe. Ciągły rozwój technologiczny i zmiany w specyfice działania narzędzi źródeł informacji, a także dostosowywanie indywidualnych technik do odpowiednich działań wywiadowczych powoduje, iż za każdym razem należy podchodzić do analizy zagrożeń indywidualnie, biorąc pod uwagę specyfikę i ekspozycję na możliwe rozpoznanie danej osoby lub organizacji.

ROZDZIAŁ 5

MOŻLIWE DO WPROWADZENIA ZALECENIA BEZPIECZEŃSTWA W ZAKRESIE PRZECIWDZIAŁANIA WYWIADOWI OTWARTOŹRÓDŁOWEMU

Uwagi wstępne

Jak już zostało to wspomniane w poprzednim rozdziale, zagrożenia wynikające z dostępności ogromnej ilości danych w Internecie, których pozyskanie możliwe jest przy pomocy technik OSINT-owych, a zatem legalnych działań, umożliwiających pozyskanie informacji z ogólnodostępnych źródeł, można także traktować w zakresie budowania sposobów ochrony osobistej i organizacyjnej przeciwko takim działaniom.

Bazę do budowania ochrony przed OSINT-em mogą stanowić wszelkie opracowane zestawy wskazówek, czy to wynikające z norm (krajowych i międzynarodowych), czy poradników wojskowych, czy też szkoleń w zakresie świadomości możliwych do wykorzystania technik i narzędzi OSINT-owych. Jako podstawy do opracowywania metod ochrony przeciw rozpoznaniu otwartoźródłowemu mogą być wykorzystywane także wnioski z przeprowadzanych symulacji tego typu rozpoznania, które dadzą pogląd na obszary wymagające wprowadzenia zabezpieczeń lub lepszego monitorowania.

Istotną rzeczą jest, iż działania, będące z jednej strony działaniami aktywnymi, wykorzystywanymi do zdobywania wiedzy o rozpoznawanej osobie lub organizacji, mogą być także wykorzystane do ochrony przed tymi atakami. W pierwszej kolejności konieczne jest przeprowadzenie rekonesansu OSINT-owego własnej osoby lub organizacji w celu znalezienia nadmiarowo udostępnionych informacji (celowo lub przypadkowo, także przez osoby trzecie). Wynikiem takiego rozpoznania będą działania, mające na celu usunięcie nadmiarowych informacji lub ich zaciemnienie. W innych przypadkach stosowane są techniki ostrzegające jakie dane, obecne w Internecie, mogą zostać użyte do przeprowadzenia ataku i przygotowanie się na nie, aby uniknąć efektu zaskoczenia. Dlatego też OSINT zyskuje ważność w zakresie przeciwdziałania atakom

i podniesienia poziomu bezpieczeństwa w odniesieniu zarówno do osób, jak i do systemów teleinformatycznych. Ma on znaczenie także w zapewnianiu bezpieczeństwa państwa, w zakresie którego operują polskie służby wywiadowcze i kontrwywiadowcze, wojskowe służby specjalne oraz inne służby o charakterze policyjnym[147].

[147] L. Wiszniewski, *Rola i znaczenie analizy informacji wywiadowczej w zapewnianiu bezpieczeństwa państwa*, Przegląd Bezpieczeństwa Wewnętrznego, 2020 nr 22 (12), s. 66-83.

1.23. Zasady wynikające z norm, standardów i innych wytycznych

Normy, dotyczące bezpieczeństwa informacji, dają wskazówki, które mogą być brane pod uwagę przy ochronie przed wykorzystaniem narzędzi i technik OSINT-owych. Na ich podstawie można budować zasady, które będą uwzględniały specyfikę danej organizacji i sposób zastosowania wymagań norm, gdyż jedynie sprofilowane polityki bezpieczeństwa mogą stanowić skuteczne narzędzie w przeciwdziałaniu lub ograniczaniu możliwości rozpoznania danej organizacji.

W przypadku normy PN-EN ISO/IEC 27001, definiującej wymagania dla systemów zarządzania bezpieczeństwem informacji oraz powiązanej z nią ściśle normy PN-EN ISO/IEC 27002, opisującej praktyczne zasady zabezpieczania informacji, możliwe jest do określenia grupę polityk bezpieczeństwa, których stosowanie będzie miało wpływ na podwyższenie poziomu bezpieczeństwa w zakresie odporności na rozpoznanie otwartoźródłowe. Polityką, która bezpośrednio będzie miała wpływ na nadmiarowe udostępnianie informacji jest polityka bezpieczeństwa usług sieciowych, która powinna definiować poziom ich zabezpieczeń oraz regularne monitorowanie zdolności do bezpiecznego prowadzenia tych usług. Duże znaczenie w coraz powszechniejszym modelu pracy zdalnej będzie miała także procedura dotycząca telepracy i zagrożeń z nią związanych: oddzielenia prywatnego i służbowego wykorzystania urządzeń, w tym szczególnie wykorzystywania prywatnych kanałów komunikacji i udostępniania danych, dotyczących szczegółów pracy zdalnej (np. zakaz publikowania zdjęć pulpitu i zrzutów ekranu z wideokonferencji na prywatnych mediach społecznościowych). Polityką, która może mieć pośrednie, ale w pewnych przypadkach bardzo duże znaczenie dla zabezpieczenia przed zebraniem wrażliwych danych w ramach OSINT-u, jest także polityka czystego biurka i ekranu, gdyż w przypadku wykonywania zdjęć lub nagrań wideo w miejscu pracy, możliwe jest także przypadkowe uchwycenie informacji tam przechowywanych (np. loginów i haseł, adresacji sieciowej, numerów telefonów i danych personalnych). Polityka ta staje się jeszcze ważniejsza, jeśli stanowisko pracy znajduje się w miejscu, do którego mają dostęp (lub wgląd) osoby postronne.

Do polityk bezpieczeństwa dla systemu zarządzania bezpieczeństwem informacji, zgodnym z wymaganiami PN-EN ISO/IEC 27001, możliwe jest też dołączenie działań OSINT-owych, np. w zakresie prześwietlania kandydatów do pracy na stanowiskach,

wymagających odpowiednich kwalifikacji, a także przy konieczności weryfikacji kandydatów na stanowiska o szczególnym znaczeniu dla bezpieczeństwa firmowego.

OSINT stanowi także bazę dla badaczy bezpieczeństwa i zespołów budujących zabezpieczenia infrastruktury teleinformatycznej, np. w zakresie pozyskiwania IoC (skrót od ang. *Indicators od Compromise*). W ramach zespołów SOC (skrót od ang. *Security Operation Centre*) budowana jest przy wykorzystaniu OSINT-u świadomość, dotycząca aktualnego wachlarza zagrożeń, dotyczącego ich infrastruktury. Dlatego też w europejskim projekcie DiSIEM[148], którego zadaniem jest rozszerzanie możliwości istniejących systemów SIEM (skrót od ang. *Security Information and Event Manangement*) o nowe moduły, powstała koncepcja zebrania źródeł OSINT-owych, dotyczących cyberbezpieczeństwa oraz wdrożenia informacji z nich pobranych do zasilenia systemów SIEM[149].

Jako bazę wskazówek, dotyczących możliwych działań, mających na celu zapobieganie nadmiarowemu udostępnianiu informacji do sfery publicznej, przez co możliwe są one do pozyskania w ramach wywiadu otwartoźródłowego, jest baza *Common Weakness Enumeration* (CWE)[150], przedstawiająca kategorie słabości, które należy brać pod uwagę oraz wskazuje, jak im przeciwdziałać. W ramach listy rodzajów słabości produktów, które należy niwelować, wskazane są rodziny słabości, możliwych do wykorzystania w ramach rozpoznania otwartoźródłowego.

Pierwszą rodziną słabości jest CWE-200: „*Exposure of Sensitive Information to an Unauthorized Actor*" (Udostępnianie wrażliwych informacji nieautoryzowanym podmiotom)[151]. Wskazuje ona, że istnieje wiele rodzajów błędów, które mogą powodować udostępnianie tychże informacji, a ważność danego błędu może się znacznie wahać, w zależności od kontekstu, w którym operuje dany produkt, rodzaju udostępnianych informacji, a także zysków, które może dzięki pozyskaniu tychże wiadomości osiągnąć atakujący. Do przykładowych informacji, których ujawnienie może być niepożądane, należą:

[148] DiSIEM – https://cyberwatching.eu/projects/1040/disiem
[149] DiSIEM Project Deliverable D4.1 – *Techniques and tools for OSINT-based threat analysis* – https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5bb1cd31f&appId=PPGMS – dostęp online 23.11.2022 r.
[150] https://cwe.mitre.org
[151] CWE-200 – https://cwe.mitre.org/data/definitions/200.html – dostęp online 22.11.2022 r.

- prywatne dane, np. wiadomości, dane finansowe, informacje o stanie zdrowia i leczenia, lokalizacje geograficzne lub dane kontaktowe;
- dane o stanie systemu lub środowiska, np. rodzaj systemu operacyjnego i zainstalowanego oprogramowania;
- poufne dane biznesowe oraz własność intelektualna;
- konfiguracja i stan sieci informatycznej;
- kod źródłowy produktu oraz jego stan wewnętrzny;
- metadane, np. logowanie połączeń lub nagłówki wiadomości;
- informacje pośrednie, np. możliwe do podejrzenia przez osoby z zewnątrz podczas interakcji w ramach dwóch wewnętrznych działań.

Jak wskazano w opisie rodziny słabości CWE-200, wrażliwość informacji będzie różna dla różnych stron, z których każda może mieć inne oczekiwania w zakresie zakresu ochrony danych informacji. Do wspomnianych stron można zaliczyć:

- użytkowników produktu;
- osoby lub organizacje, których dotyczą informacje wytwarzane lub wykorzystywane przez dany produkt, nawet jeśli nie są bezpośrednimi użytkownikami produktu;
- administratorów produktu, włączając w to administratorów systemu/systemów i/lub sieci, w których produkt działa;
- twórców produktu.

Drugą rodziną słabości, która wskazuje możliwe do podjęcia działania w zakresie zabezpieczenia przez nadmiarowym udostępnianiem danych jest CWE-552: *Files or Directories Accessible to External Parties* (Pliki lub katalogi dostępne dla zewnętrznych podmiotów)[152]. W jej ramach można wyróżnić m.in. takie rodzaje słabości, jak: umieszczanie plików z wrażliwymi danymi w głównym katalogu aplikacji webowej (CWE-219[153]) lub serwera FTP (CWE-220[154]), udostępnienie repozytorium kontroli

[152] CWE-552 – https://cwe.mitre.org/data/definitions/552.html – dostęp online 22.11.2022 r.
[153] CWE-219 – https://cwe.mitre.org/data/definitions/219.html – dostęp online 22.11.2022 r.
[154] CWE-220 – https://cwe.mitre.org/data/definitions/220.html – dostęp online 22.11.2022 r.

wersji (CWE-527[155]) lub plików kopii zapasowych (CWE-530[156]) podmiotom nieautoryzowanym.

[155] CWE-527 – https://cwe.mitre.org/data/definitions/527.html – dostęp online 22.11.2022 r.
[156] CWE-530 – https://cwe.mitre.org/data/definitions/530.html – dostęp online 22.11.2022 r.

1.24. Stosowanie zasad bezpieczeństwa operacyjnego (OPSEC) i osobistego (PERSEC)

Bezpieczeństwo operacji (OPSEC) jest elementem działań w ramach wywiadu otwartoźródłowego, ale może być także stosowane jako metoda przeciwdziałania mu. Zasady OPSEC zaczęły być formalizowane w ramach armii amerykańskiej po wydarzeniach mających miejsce w trakcie wojny w Wietnamie. Ta tam okazało się, że siły wietnamskie, pomimo braku możliwości odszyfrowania komunikacji pomiędzy jednostkami amerykańskimi, miały wiedzę na temat wszelkich planowanych działań wojsk Stanów Zjednoczonych w południowej Azji. Było to spowodowane zbyt szerokim ujawnianiem danych związanych z tymi działaniami przez same siły amerykańskie. Wynikiem analizy zaistniałej sytuacji była operacja „Purple Dragon" („Fioletowy Smok")[157], której zadaniem było odkrycie, które informacje są nadmiarowo udostępnione i jednocześnie mogą stanowić wartościowe i możliwe do wykorzystania przez wroga wskazówki. Następnym krokiem było zabezpieczenie zidentyfikowanych problematycznych elementów komunikacji. Z dzisiejszego punktu widzenia można porównać zatem działania operacji „Purple Dragon" do rozpoznania OSINT-owego i zabezpieczenia się przed nim.

W dzisiejszej rzeczywistości, gdzie ogromna ilość działań przeniesiona została do sfery cyfrowej, OPSEC obejmuje działania dotyczące zabezpieczenia elementów, które stanowią punkt nadmiernego udostępniania danych, jednak nie w odniesieniu do działań na froncie, ale głównie w Internecie. Analiza ekspozycji swoich danych w sieci i zabezpieczenie się przed tym, wykonane jeszcze przed rozpoczęciem działań, może zapobiec zdemaskowaniu i zniweczeniu zachowania odpowiedniego poziomu skrytości i anonimowości. Osoby przeprowadzające wywiad otwartoźródłowy mogą spotkać podczas swoich działań przeciwników, którzy także będą wykorzystywać OSINT w celu zdemaskowania i identyfikacji tychże osób.

W ramach elementów zachowania bezpieczeństwa operacji należy w pierwszej kolejności zdefiniować poziom OPSEC jaki będzie wymagany podczas prowadzenia działań w sieci. Inny poziom będzie wymagany podczas analizy konkurencji przez firmę, a inny podczas śledzenia sprawców przestępstw przez organy ścigania. Dlatego też inny

[157] National Security Agency – *Purple Dragon. The Origin and Development of the United States OPSEC Program* – https://www.nsa.gov/portals/75/documents/news-features/declassified-documents/cryptologic-histories/purple_dragon.pdf – dostęp online 23.11.2022 r.

będzie zakres koniecznych do przedsięwzięcia kroków, mających na celu wdrożenie odpowiedniego poziomu bezpieczeństwa.

W przypadku chęci uchronienia się przed identyfikacją podczas działań w Internecie, stosowane są często środowiska wirtualne oraz systemy operacyjne typu *live*, które powodują, że każde śledztwo będzie prowadzone z czystej maszyny, bez powiązania z innymi, wcześniejszymi sprawami. Także sposób połączenia się do Internetu jest tutaj sprawą kluczową. Połączenie się ze swojej domowej lub organizacyjnej sieci powoduje możliwość łatwej identyfikacji, dlatego w przypadku konieczności zachowania anonimowości możliwe są do zastosowania połączenia przez VPN, serwery proxy, sieć Tor, a także korzystanie ze stron pośredniczących w uzyskiwaniu dostępu do informacji. Te ostatnie mogą mieć formę specjalnych portali internetowych, które oferują usługę przechwycenia obrazu pożądanej strony internetowej i wyświetlenia jej nam w formie obrazu lub pliku PDF. Oznacza to, że podobnie jak podczas połączenia przez serwer proxy, to nie z maszyny osoby wykonującej rozpoznanie przychodzi połączenie do docelowego serwera, a z adresu strony pośredniczącej. Innym sposobem na przechwytywanie widoku strony internetowej, bez konieczności interakcji z nią, jest skorzystanie z opcji „Kopia", dostępnej obok niektórych wyników wyszukiwania w wyszukiwarce Google, gdzie otrzymujemy stronę internetową w formie, jaką przechwyciło Google podczas ostatniego skanowania. Taki sam efekt uzyskiwany jest w przypadku wspominanego wcześniej operatora wyszukiwania „cache:".

Kolejną zasadą OPSEC jest oddzielenie kont, wykorzystywanych w ramach śledztw od kont prywatnych, nie tylko poprzez używanie osobnych skrzynek e-mailowych i profili w portalach społecznościowych, ale także przez nieużywanie danych prywatnych do tworzenia loginów, imion, nazwisk, zdjęć i innych elementów kont służbowych. Brak zastosowania takiej zasady spowoduje, że pomimo założenia osobnej skrzynki lub profilu, osoby śledzące zostaną poprzez odpowiednio wykonany OSINT zidentyfikowane i przypisane do ich prawdziwych tożsamości. Temat jest nietrywialny i często drobny element, jak wykorzystanie swojego imienia lub pseudonimu, może doprowadzić do „dekonspiracji". Nawet sposób wypowiadania się w Internecie może doprowadzić do identyfikacji konkretnej osoby, jak to miało miejsce w sprawie Australijczyka, mającego na sumienia przestępstwa wobec małych dzieci, którego zdradziło oryginalne powitanie stosowane w Internecie – „*hiyas*". Natomiast seria drobnych uchybień w stosowaniu zasad OPSEC, spowodowała, że prawdopodobnie

udało się badaczom namierzyć anonimowe konta szefa FBI – Jamesa Comeya. Pierwszym z błędów było przyznanie przez samego Comeya, że w ogóle ma konta na Twitterze i Instagramie oraz określenie, ile aktualnie osób śledzi te konta. Drugim błędem były powiązania anonimowego konta z kontami, należącymi do innych osób z jego rodziny, natomiast kolejnymi – wykorzystanie informacji związanych ze swoją karierą naukową i zawodową do stworzenia pseudonimu dla anonimowego konta oraz specyficzne kategorie wpisów i śledzonych tematów, mające związek konkretnie z upodobaniami i znajomościami Comeya. W ten sposób, za pomocą z pozoru drobnych elementów informacji, możliwe jest zidentyfikowanie prawdopodobnego właściciela danego konta.

W ramach bezpieczeństwa operacji stosuje się anonimowe konta, które określane są mianem tzw. *sock puppets*. Określenie to oznacza w dosłownym tłumaczeniu kukiełkę wykonaną ze skarpety i jest odzwierciedleniem reprezentacji danej osoby w innym wymiarze, tutaj – w Internecie. Zasady tworzenia kont alternatywnych tożsamości sprowadzają się do następujących aspektów:

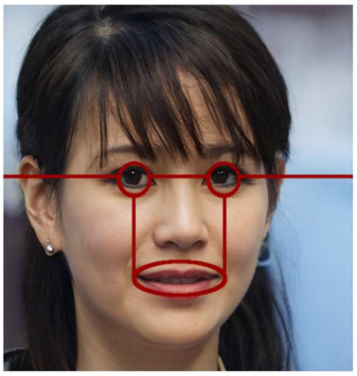
- imię i nazwisko lub pseudonim – niepowiązane z prawdziwą tożsamości i odpowiednie dla prowadzonych działań, co uwzględnia m.in. dobór imienia do wieku i narodowości tworzonej tożsamości;
- płeć i wiek – określenie płci i wieku alternatywnej tożsamości w odniesieniu do potrzeb i specyfiki prowadzonych działań, infiltrowanych środowisk oraz pożądanych wyników;
- zdjęcie profilowe – fotografia, przedstawiająca osobę, której dane zgadzają się z wcześniej określonym imieniem, płcią i wiekiem, przy czym nie wskazująca jednoznacznie, że profil jest sztucznym kontem[158];
- zawód – określony dla potrzeb śledztwa zawód, który będzie mógł ułatwić prowadzone działania i pozostanie nie wykrytym, np. rekruter

[158] Popularnym portalem, służącym do generowania zdjęć profilowych jest thispersondoesnotexist.com, który korzysta z technologii GAN (skrót od ang. *Generative Adversarial Networks* – generatywne sieci współzawodniczące), umożliwiającej tworzenie unikalnych obrazów twarzy. Niestety ograniczenia i znane błędy tego typu mechanizmu (jak stałe położenie źrenic i ust, częste błędy symetrii okularów, kolczyków i błędy w otoczeniu twarzy) powodują, że obraz jest bardzo łatwy do zidentyfikowania jako sztucznie wytworzony. Sposobem na zapobieżenie odkryciu fałszywego zdjęcia profilowego jest stosowanie kombinacji różnych technik, obrazów i efektów graficznych, jak przekształcenia czy łączenie twarzy.

w przypadku śledzenia i nawiązywania kontaktów poprzez portale zawodowe;

- zachowanie – sposób wypowiadania się, interakcji z innymi osobami w Internecie, ilości publikowanych danych i ich rodzaju;
- adres e-mail i numer telefonu – zastosowanie niewykorzystywanego wcześniej adresu e-mail oraz numeru telefonu, a także środowiska ich wykorzystania, jak np. czysta maszyna wirtualna oraz czysty telefon (nowy lub przywrócony do ustawień fabrycznych);
- sieć powiązań – tworzenie odpowiednich sieci znajomych, uwiarygadniających dany profil, np. poprzez wchodzenie w interakcję z prawdziwymi profilami, które jednocześnie są w kręgach, do których dana alternatywna tożsamość ma dołączyć, a z drugiej strony ich właściciele nie weryfikują tożsamości osoby próbującej nawiązać znajomość przed przyjęciem jej do tego grona.

Rys. 20 – Porównanie zdjęć profilowych, wygenerowanych sztucznie na stronie thispersondoesnotexist.com – od lewej: z widocznymi błędami w zakresie symetryczności okularów, kolczyków oraz z widocznymi błędami w otoczeniu twarzy. Źródło: thispersondoesnotexist.com – dostęp online 23.11.2022 r.



Rys. 21 – Analiza stałego położenia oczu i ust na zdjęciu profilowym, wygenerowanym na portalu thispersondoesnotexist.com, z wykorzystaniem narzędzia AmIReal (https://seintpl.github.io/AmIReal – dostęp online 23.11.2022 r.)

Wykorzystanie kont z fałszywą tożsamością jest niezgodne z regulaminem w wielu platformach internetowych. Każdorazowo zatem należy zweryfikować czy wykorzystanie takiego konta nie spowoduje przekroczenie granic, narzuconych także przez podmiot, na rzecz którego wykonujemy śledztwa OSINT-owe.

Zasady ograniczania informacji przekazywanych wraz z docelowo jawnym zakresem danych, czyli np. adresem e-mail, dotyczą także kont prywatnych, nie wykorzystywanych do działań śledczych. W ramach bezpieczeństwa osobistego, czyli zasad PERSEC, można m.in. wskazać częste nadmiarowe udostępnianie informacji o dacie urodzenia, pojawiającej się niekiedy w loginach adresów e-mailowych. Często sposobem na identyfikację przybliżonego wieku osoby, stojącej za danym kontem w Internecie, będzie identyfikacja zestawu platform i serwisów, na których dana osoba posiada konto, gdyż będą one inne dla dzieci, nastolatków, osób dorosłych i osób starszych.

Zagadnieniem, które także wpisuje się w zasady zachowania bezpieczeństwa osobistego, jest konfiguracja ustawień na stronach mediów społecznościowych. Brak ustawienia zabezpieczeń, chroniących konta i zawarte na nich dane przed dostępem osób nieuprawionych może skutkować ujawnieniem ogromnej ilości danych – od daty urodzin (poprzez publiczne udostępnianie wpisów, w których znajomi składają danej osobie życzenia z okazji urodzin) aż do danych dotyczących miejsca zamieszkania (widok z okna lub stałe punkty na trasie do pracy lub szkoły, widoczne na wpisach e mediach społecznościowych o stałych porach) czy statusu materialnego. Z tego względu możliwym rozwiązaniem jest niepodawanie w ogóle prawdziwej daty urodzenia, podanie

daty nieprawdziwej, która będzie elementem mylącym albo odpowiednie zabezpieczenie informacji o dacie urodzenia przed dostępem osób nieuprawionych.

1.25. Rozpoznanie otwartoźródłowe jako element świadomości podatności na atak

Poprzez rozpoznanie otwartoźródłowe zazwyczaj rozumie się zdobywanie informacji o przeciwniku, jednak proaktywne podejście do ochrony infrastruktury teleinformatycznej uwzględnia także działania zespołów atakujących, określanych mianem *red team*, które skupiają się na zasymulowaniu zachowania prawdziwych atakujących. Celem takich zabiegów jest lepsze przygotowanie się na faktyczne ataki. Częścią pracy zespołów atakujących jest także zdobywanie wiedzy eksponowanej przez obiekt atakowany w ogólnodostępnych źródłach, stąd można stwierdzić, że narzędzia i techniki OSINT-owe stanowią ważny element budowania świadomości własnej podatności na atak.

W przypadku badania podatności infrastruktury teleinformatycznej, narzędziami do zdobywania informacji będą głównie bazy informacji o infrastrukturze, a więc m.in. Shodan, Censys, Zoomeye oraz Virustotal. W przypadku chęci aktywnego rozpoznania, możliwe jest skorzystanie ze skanera portów nmap, wraz z jego zestawem skryptów, badających potencjalne występowanie podatności.

Rozpoznanie nie musi jedynie ograniczać się do zakresu sieciowego, ale może także zostać rozszerzone o inne aspekty, jak np. zdobywanie publicznie dostępnych adresów e-mail osób, pracujących w danej organizacji. Pozyskane w ten sposób adresy mogą później posłużyć do ataków phishingowych, co oznacza, że użytkowników należy o tym powiadomić, w celu uniknięcia efektu zaskoczenia. Należy tutaj także zaznaczyć, że nawet brak znalezienia firmowego adresu e-mail pracownika nie sprawia, że nie będzie on celem ataku phishingowego, gdyż możliwe jest w zakresie OSINT-u odnalezienie wzorca tworzenia adresów w danej domenie firmowej (np. w formie: imię.nazwisko@domena), co oznacza, że atakującym wystarczy informacja, że osoba o określonych danych pracuje w konkretnej organizacji (np. poprzez analizę portalu LinkedIn lub artykułów prasowych, w których dana osoba się wypowiadała), aby uzyskać jej adres e-mail.

Niekiedy dane, które mogą zostać wykorzystane w ramach ataku na daną infrastrukturę, atakujący zdobywają ze źródeł oferujących dane z wycieków lub wykradzionych. Korzystanie z tego typu źródeł podczas rozpoznania w ramach testów proaktywnej ochrony infrastruktury powinno zawsze być zaakceptowane przez osobę

decyzyjną, ze względu na legalność pozyskiwania i wykorzystywania tego typu informacji w ramach legalnych działań.

Możliwe jest też wykorzystanie elementów podnoszących świadomość podatności na atak, kiedy atakujący zaczyna już realnie wykonywać rozpoznanie OSINT-owe, przygotowując specjalne pułapki, odpowiadające różnym punktom podlegającym rozpoznaniu. Ten typ pułapek, podobnych do idei *honeypotów*[159], może zostać wykonany np. przy użyciu tzw. *canary tokens*, które poinformują ich twórcę o prowadzonym w stosunku do niego rozpoznaniu oraz pozwolą mu na przygotowanie się do ewentualnej późniejszej obrony. *Canary tokens* mogą przyjmować formę np. dokumentów Word lub Excel, archiwów, obrazków, adresów internetowych itp., umieszczone np. na serwerze dostępnym w Internecie w folderze, który może zostać znaleziony tylko przy wykorzystaniu technik OSINT-owych. Po otwarciu takiego pliku lub odwiedzeniu adresu internetowego zostanie wygenerowane powiadomienie dla ich twórcy, kto, kiedy i z jakiego adresu wszedł w interakcję z danym tokenem. Nazwa „*canary tokens*" odwołuje się do kanarków, które były wykorzystywane w kopalniach do informowania górników o niebezpiecznym stężeniu gazów – stąd analogia do powiadamiania o potencjalnym niebezpieczeństwie w dziedzinie IT.

Wykorzystanie symulowanego rozpoznania otwartoźródłowego i ataków na własną infrastrukturę, bazujących na zdobytej w ten sposób wiedzy jest działaniem, które rozszerza świadomość ekspozycji na możliwe ataki i umożliwia zidentyfikowanie i usunięcie lub przygotowanie się na próby zagrożenia eksponowanym elementom.

[159] Honeypot – system, który ma za zadanie wykrycie nieautoryzowanych prób przełamania zabezpieczeń systemów IT. Jego konfiguracja z jednej strony udaje słabo zabezpieczoną i zawierającą interesujące zestawy danych, co ma wabić atakujących, jednak de facto jest odizolowana od faktycznej sieci.

1.26. Rozszerzenie zasad zabezpieczeń na otoczenie kluczowych osób w organizacjach

Budując zabezpieczenia w ramach danej organizacji, głównym ich założeniem będzie ochrona informacji znajdujących się w niej poprzez środki ochrony fizycznej oraz zabezpieczenia teleinformatyczne, jednak trzecim elementem, który należy brać pod uwagę, jest zabezpieczenie od strony użytkowników, mających dostęp do kluczowych informacji. Osoby te są z jednej strony najczęściej rozpoznawalne w Internecie ze względu na swoją pozycją, a z drugiej mają wysokie uprawnienia, co może powodować, że to właśnie w stosunku do nich będą stosowane techniki OSINT-owe.

Jak wskazuje Christina Lekati w swoim wystąpieniu „Protecting High-Value Individuals: An OSINT Workflow”[160], do kręgu osób, które mogą być zaliczane do kręgu celów o wysokiej wartości (tzw. *High Value Targets – HVT*), należeć będą zarówno prezesi, kadra kierownicza i zarządzająca, osoby o wysokich uprawnieniach, z dostępem do poufnych informacji oraz pracujące nad projektami o wysokiej wrażliwości informacji, ale także politycy, sportowcy, influencerzy i inne znane osoby. To co do nich mogą być przeprowadzane ataki, mające różne cele – od chęci zaszkodzenia lub kradzieży, aż po nękanie przez fanów o niezrównoważonej psychice. Dlatego też w celu ochrony takich osób, a zatem i zasobów, z którymi pracują na co dzień, Christina Lekati wskazuje konieczność wykonania działań tzw. proaktywnego wywiadu, odpowiadającego na pytania: jaki ślad cyfrowy zostawia dana osoba w Internecie oraz zebranie i przeanalizowanie tego typu danych w celu określenia możliwych do przeprowadzenia ataków na osobę kluczową lub jej rodzinę. Należy w tym miejscu zaznaczyć, że tego typu rozpoznanie zawsze musi być wykonywane za zgodą i wiedzą danej osoby. Celem takich działań jest wyeliminowanie elementu zaskoczenia w momencie faktycznego przeprowadzenia ataku na osobę kluczową i zminimalizowanie powierzchni takiego ataku, a same działania, mające na calu przygotowanie się na atak, są niewidoczne dla potencjalnego atakującego, przez co nie zmieni on swojego sposobu ataku. Poprzez techniki OSINT-owe możliwe jest także zbudowanie profilu atakującego oraz wskazanie poprawek w modelu bezpieczeństwa organizacyjnego, co może prowadzić za zapobieżenia działaniom osób atakujących.

[160] C. Lekati – *Protecting High-Value Individuals: An OSINT Workflow* – https://www.youtube.com/watch?v=rE4mORq9T5s – dostęp online 23.11.2022 r.

W trakcie prowadzenia rozpoznania, dotyczącego kluczowych osób należy przygotować odpowiedni plan działania, wskazujący zakres pozyskiwanych danych oraz inne ramy rozpoznania, a następnie przeanalizować zebrane dane w celu uzyskania odpowiedzi na zadane na początku pytania, jak to ma miejsce w standardowym cyklu wywiadowczym. Jako kategorie brane pod uwagę w przypadku rozpoznania w stosunku do kluczowych osób Lekati wskazuje:

- rozpoznawalność – ile i jakiego rodzaju informacje są dostępne w odniesieniu do danej osoby kluczowej: dane osobowe, adresy, bezpieczeństwo w okolicy zamieszkania, nawyki, publikowane informacje o członkach rodziny, plany i sposoby podróżowania itp.
- dostępność – łatwość, z jaką atakujący może dotrzeć do danej osoby i ją zaatakować, w celu osiągnięcia pożądanej korzyści;
- podatność – jaką świadomość bezpieczeństwa ma ochraniana osoba i jakie są słabe strony w jej profilu (np. rodzina, przyjaciele, nawyki itp.);
- *threat intelligence* – informacje o atakującym, jego motywach oraz sposobie działania, co może znacznie usprawnić proaktywną ochronę.

Wszystkie zebrane informacje powinny zostać przedstawione osobie, która była obiektem badania podatności na atak, łącznie z zaleceniami dotyczącymi możliwych do podjęcia kroków oraz zakresem rozpoznania oraz jego ograniczeniami i wyłączeniami. W celu lepszego przekazania wniosków osobom na wysokich stanowiskach należy stosować delikatny, ale konkretny ton raportu[161].

Przykładem sytuacji, w której wysoko postawiona osoba stała się ofiarą ataku poprzez zebrane ogólnodostępne informacje o sobie jest sytuacja z 2008 roku, kiedy to ofiarą prostego jak się później okazało ataku, stała się Sarah Palin, kandydatka na stanowisko wiceprezydenta. Odzyskanie hasła do jej skrzynki e-mailowej było możliwe dzięki odpowiedzi na trzy pytania bezpieczeństwa, co było wówczas metodą dość powszechnie stosowaną przy tego typu działaniach. Odpowiedzi te mogły być z łatwości uzyskane poprzez rozpoznanie otwartoźródłowe Palin, gdyż dotyczyły: daty urodzenia, kodu pocztowego oraz miejsca, gdzie poznała swojego partnera. Student, który był

[161] Tamże.

odpowiedzialny za włamanie, wskazywał, że dzięki dużej liczbie informacji o osobie publicznej, jaką była Sarah Palin, zdobycie tych informacji nie było skomplikowane[162].

Sytuacja ta wskazuje, że wraz z większą ekspozycją kluczowych osób w danej organizacji, rośnie także ich podatność na ataki z wykorzystaniem danych zebranych przy wykorzystaniu OSINT-u, a zatem wzrasta także konieczność ich proaktywnej ochrony.

[162] K. Zetter, *Palin E-Mail Hacker Says It Was Easy* – https://www.wired.com/2008/09/palin-e-mail-ha/ – dostęp online 23.11.2022 r.

Wnioski

Wykorzystanie OSINT-u jako elementu ochrony proaktywnej, a zatem wykonanie symulowanego rozpoznania i wykorzystania zdobytej w ten sposób wiedzy do przetestowania wdrożonych zabezpieczeń zarówno infrastruktury teleinformatycznej, jak i odporności kadry zarządzającej oraz pracowników organizacji, jest jednym ze sposobów zwiększenia poziomu jej bezpieczeństwa i świadomości możliwych przebiegów ataków. Wymagania, dotyczące zabezpieczeń infrastruktury przed rozpoznaniem OSINT-owym mogą być pozyskiwane z wielu źródeł, jednak zawsze wdrożenie polityk bezpieczeństwa, powinno zostać poprzedzone analizą adekwatności wdrażanych zabezpieczeń do specyfiki działania organizacji.

Wdrożenie dobrych praktyk, wynikających z ogólnych zasada zachowania bezpieczeństwa operacji, a także bezpieczeństwa personalnego, jest kolejnym krokiem, przybliżającym do osiągnięcia wysokiego poziomu odporności na rozpoznanie. Niemniej ilość dostępnych danych, które są konieczne do prawidłowego działania infrastruktury sieciowej, a także prowadzenia działań biznesowych i marketingowych sprawia, że części z nich nie jesteśmy w stanie usunąć, a jedynie konieczne jest kontrolowanie ich zakresu i utrzymywanie ich w jak najmniejszym możliwe zakresie, nie wpływającym na możliwości pożądanego działania w Internecie.

OSINT w zakresie identyfikacji możliwych podatności dotyczy nie tylko infrastruktury, ale także sfery pracowniczej, związanej z publikacją informacji w Internecie i innych mediach. Możliwe jest, że przeprowadzony atak będzie połączeniem dwóch obszarów rozpoznania, tzn. informacje o pracownikach mogą być wykorzystywane do prób zaatakowania infrastruktury (np. za pomocą stworzonych loginów na bazie rozpoznanej listy pracowników w portalach społecznościowych) lub na odwrót – uzyskane informacje o infrastrukturze będą wykorzystywane w atakach na pracowników (np. poprzez próbę zdobycia danych logowania do konkretnej platformy przy wykorzystaniu ataku socjotechnicznego, przeprowadzanego za pomocą rozmowy telefonicznej).

Zastosowanie opisanych przez autora niniejszej pracy zabezpieczeń nie wyczerpuje tematu ochrony osób i infrastruktury poprzez wykorzystanie technik OSINT-owych, gdyż ze względu na mnogość scenariuszy ataków i specyfiki działania chronionego podmiotu, należy mieć na uwadze ich rozszerzenie i dostosowanie. Ciągłe

zmiany w sposobie komunikacji, jak powstawanie nowych form przekazu, platform internetowych, mediów społecznościowych oraz zabezpieczeń i metod OSINT-u, wymagają bieżącej analizy otoczenia i dostosowywania wykorzystywanych technik do aktualnych zagrożeń.

ZAKOŃCZENIE

OSINT jako sposób pozyskiwania danych z ogólnodostępnych źródeł oraz analizowania ich w celu stworzenia obrazu wywiadowczego, pozwalającego na stworzenie odpowiedzi na zadane pytania, przeszło dużą transformację od swoich początków do dnia dzisiejszego. Rozwój sieci Internet sprawił, że techniki pozyskiwania danych z gazet, audycji radiowych, filmów i zdjęć, zostały w dużej mierze rozszerzone, a niekiedy nawet zastąpione pozyskiwanie danych z internetowych kanałów komunikacyjnych. Duży dostęp do Internetu i możliwość umieszczania w nim praktycznie dowolnych informacji sprawiły, że musiały być dostosowane także narzędzia i techniki pobierania informacji oraz metody ich odsiewania i analizy. Duży dostęp do otwartych danych w Internecie stawia osoby prywatne, firmy oraz wszelkiego rodzaju organizacje przed koniecznością opracowania metod badania swoich podatności na ataki z wykorzystaniem rozpoznania otwartoźródłowego, jak i aktualnych metod obrony.

Powstanie różnych poziomów Internetu, jakimi są sieć indeksowana oraz nieindeksowana (podzielona dodatkowo na *deep web* i *dark web / dark net*), stawia przed rozpoznaniem otwartoźródłowym dodatkowe wyzwania w zakresie opracowania osobnych technik i narzędzi, służących do pozyskiwania informacji z tych obszarów. Działania w Internecie muszą zostać odpowiednio zaplanowane i przygotowane, w zakresie: opracowania planu i zakresu poszukiwanych informacji, określenia poziomu bezpieczeństwa i anonimowości prowadzonych działań (w tym stosowanych metod OPSEC oraz decyzji w zakresie stosowania rozpoznania: aktywnego, pasywnego lub obu) oraz stworzenia środowiska pracy, dostosowanego do opracowanych potrzeb.

Stosowanie narzędzi OSINT-owych powinno być poparte ich wcześniejszą analizą, gdyż mogą one same w sobie stanowić niebezpieczeństwo w ramach prowadzonych działań. Metody ich pracy powinny zostać gruntownie sprawdzone, ponieważ może się okazać, że uzyskane przy wykorzystaniu narzędzia wyniki nie będą pełne lub nie będą odpowiadały zakresowi poszukiwanych informacji.

OSINT może być stosowany zarówno do działań rozpoznania innych firm i osób, jak i do budowania świadomości o własnych podatności i możliwych wektorów ataków. Działania zespołów badających bezpieczeństwo własnej infrastruktury powinny zawierać także metody OSINT-owe, gdyż ich wykorzystanie nie wymaga dużych nakładów pracy i jest dostępne praktycznie dla każdego użytkownika Internetu. Działania w ramach

programów ochrony firm i organizacji mogą zostać też rozszerzone o rozpoznanie w stosunku do kluczowych pracowników, jak kadra zarządzająca i osoby z wysokimi uprawnieniami dostępu do wrażliwych i cennych informacji. Do rozwoju technik obrony przed wykorzystaniem technik OSINT-owych przeciw danemu podmiotowi można stosować różnego rodzaju wskazówki, wynikające z norm, dotyczących bezpieczeństwa informacji, jak PN-EN ISO/IEC 27001, ale także aktualnych baz podatności, dotyczących namiarowego publikowania wrażliwych danych oraz wszelkich elementów wiedzy, wynikających ze szkoleń w zakresie metod OSINT-u.

Poprzez identyfikację pewnego obszaru niewiedzy o charakterze naukowym w zakresie możliwości i zagrożeń, wynikających z wykorzystywania wywiadu otwartoźródłowego, możliwe było określenie głównego problemu badawczego: *w jakim zakresie dostępne narzędzia, służące do gromadzenia informacji w ramach wywiadu otwartoźródłowego oraz techniki ich analizy wpływają na bezpieczeństwo systemów teleinformatycznych oraz bezpieczeństwo osobowe, a także jakie są możliwości obrony przed zidentyfikowanymi technikami?*

W wyniku takiego przedstawienia problemu głównego, sformułowano w sensie teoretycznym podstawowy cel badań, przedstawiony w niniejszej dysertacji jako: *identyfikację i ustalenie możliwości uzyskania szczegółowych informacji, wynikających z przeprowadzanego wywiadu otwartoźródłowego, w odniesieniu do systemów teleinformatycznych oraz indywidualnych osób, a także zidentyfikowanie zagrożeń wynikających z tego typu działań oraz metod skutecznej obrony przed przedmiotowym rozpoznaniem*.

W sensie pragmatycznym do celu badań należało: *opracowanie koncepcji identyfikacji i weryfikacji dostępnego zbioru informacji zawierających szczegóły techniczne, osobowe oraz geolokalizacyjne, dotyczące systemów teleinformatycznych oraz osób, a także określenie możliwości wprowadzenia zabezpieczeń przed działaniem zidentyfikowanych technik*.

Tak postawiony cel niniejszej dysertacji został w opinii jej autora w znacznym stopniu osiągnięty. Utrudnieniem w jego wypracowaniu był w pewnym brak formalizacji rozwijanych na bieżąco technik OSINT-u, gdyż ze względu na dynamiczny rozwój sieci Internet oraz ciągłe zmiany w sposobie funkcjonowania serwisów internetowych oraz

mediów społecznościowych, które stanowią nową bazę udostępniania informacji szerszemu gronu odbiorców.

Uzyskane jednak od ekspertów informacje, potwierdzają opinię autora niniejszej dystertacji o poprawnym zidentyfikowaniu zakresu aktualnych możliwości i zagrożeń, wynikających z rozpoznania otwartoźródłowego. Wskazane techniki, narzędzia i kierunki działań, mające na celu wykorzystanie OSINT-u do działań zarówno w sferze aktywnego rozpoznania celu, jak i do wypracowania metod obrony przed nimi w ramach proaktywnego poszukiwania podatnych elementów, dają pogląd na spektrum tematyki OSINT-u w odniesieniu do bezpieczeństwa osobowego i biznesowego.

Na podstawie aktualnie posiadanego stanu wiedzy, wynikającego z badań wstępnych, a także z analizy literatury, możliwe było sformułowania następującej głównej hipotezy roboczej: *zakładam, że w związku z coraz szerszym wachlarzem narzędzi i usług, dostarczających szeroki zakres danych w Internecie oraz poprzez coraz powszechniejszy dostęp do Internetu i znajomości sposobów na wyszukiwanie w nim treści, a także ze względu na fakt, że praktycznie wszystkie aspekty życia osobistego i zawodowego mają swoje odzwierciedlenie w systemach operujących w chmurze, istnieje zwiększające się zagrożenie zarówno dla bezpieczeństwa systemów teleinformatycznych, które te dane przetwarzają, jak i bezpieczeństwa osobowego, które jest bezpośrednio związane z kwestią poufności i integralności przetwarzanych danych. Możliwości i umiejętności użytkowników Internetu dają im sposobność na sprawdzenie jakie dane mogą zdobyć bez narażania się na bezpośrednie niebezpieczeństwo związane z infiltracją źródeł danych*.

W wyniku przeprowadzonych badań, zweryfikowana ona została w dużym zakresie pozytywnie. Określono jakie techniki, dostępne dla każdego użytkownika Internetu, umożliwiają szeroki wachlarz możliwości rozpoznania i uzyskania odpowiedzi na pytania wywiadowcze, sformułowane w stosunku do analizowanego podmiotu. Zagrożenia, wynikające z przeniesienia znacznej części działań osobistych oraz biznesowych do Internetu znacznie zwiększają ekspozycję podmiotów na OSINT, a także stwarzają bardziej bezpośrednie niebezpieczeństwo wykonania ataku z jednoczesnym znikomym ryzykiem, związanym z wykryciem i konsekwencjami działań atakujących.

Wyniki dotyczące problemów szczegółowych, rozwiązanych w ramach prowadzonych badań, zostały zawarte w poszczególnych rozdziałach merytorycznych.

W ramach pierwszego rozdziału merytorycznego autor skupił się na umiejscowieniu OSINT-u w środowisku różnorodnych rodzajów rozpoznania, gdzie współistnieje on m.in. z takimi rodzajami jak: HUMINT, SIGINT, IMINT oraz MASINT. Określone zostały ramy, w jakich musi mieścić się rozpoznanie, aby było uznawane za OSINT w rozumieniu nomenklatury NATO-wskiej oraz amerykańskich dyrektyw wywiadowczych, gdzie w ramach Dyrektywy 301 zdefiniowano OSINT jako „wytworzony na podstawie publicznie dostępnych informacji, które są zbierane, wykorzystywane i rozpowszechniane w określonym wymiarze czasowym odpowiednim odbiorcom, w celu uzyskania konkretnych potrzeb wywiadowczych”[163]. Wskazano także odrębny obszar rozpoznania, którego bazą są media społecznościowe – SOCMINT, które zyskuje na znaczeniu wraz z rozwojem znaczenia portali społecznościowych w rozpowszechnianiu informacji.

W tym rozdziale przybliżono także podział Internetu na sieć indeksowaną i nieindeksowaną (w tym sieć ukrytą, dostępną jedynie z użyciem specjalnego oprogramowania – *dark net*), a także wskazano potrzebę odpowiedniego przygotowania środowiska pracy oraz określenia potrzeb w zakresie prowadzenia rozpoznania aktywnego lub pasywnego przed rozpoczęciem działań. Wskazano także możliwości wyszukiwania informacji tekstowych, graficznych oraz specjalistycznych (technicznych) w oparciu o wyszukiwarki tekstowe oraz graficzne, a także narzędzia specjalistyczne.

W drugim rozdziale merytorycznym skupiono się na procesie analizowania informacji, pozyskanych z zasobów internetowych jako fundamentu wywiadu, szczególnie w aspekcie tworzenia wiedzy wywiadowczej na podstawie zebranych informacji oraz wpływu psychologicznych cech ludzkich na poprawne prowadzenie tego typu procesu. Przedstawiono różne metody rozumowania, które stanowią element procesu analitycznego. Przeanalizowano błędy poznawcze, które wynikają z niedoskonałości oraz sposobów funkcjonowania ludzkiego rozumowania, a także wskazano wpływ nieprawidłowego zakresu zebranych informacji, propagandy oraz dezinformacji. Przedstawiono wzorce, występujące w ludzkim rozumowaniu, które wpływają na poprawność procesu analitycznego. Zidentyfikowane rodzaje błędów poznawczych przedstawiono wraz z ich sposobem oddziaływania na podejmowanie decyzji przez analityków. Wiedza o przedstawionych mechanizmach wpływu błędów poznawczych na

[163] *Intelligence Community Directive Number 301*, National Open Source Enterprise, 2006, https://irp.fas.org/dni/icd/icd-301.pdf – dostęp online 22.03.2022 r. Tłumaczenie własne.

ocenę sytuacji stanowi pierwszy krok do ich uniknięcia przez analityków w procesie cyklu wywiadu.

Dodatkowo wskazano inne czynniki ludzkie, mogące mieć wpływ na proces prowadzenia wywiadu otwartoźródłowego, jak elementy wtórnego zespołu stresu pourazowego, na co są narażone osoby pracujące nad tematami, cechujące się dużą wagą emocjonalną. Przedstawiono metody minimalizowania wpływu tego typu czynników na wyniki prowadzonego rozpoznania i zdrowie samych analityków.

W ramach trzeciego rozdziału merytorycznego wskazane zostały zidentyfikowane metody wykorzystania wywiadu otwartoźródłowego w zakresie bezpieczeństwa systemów teleinformatycznych oraz bezpieczeństwa osobowego i biznesowego. Środowisko działania systemów teleinformatycznych wymaga od nich publikowania zestawu informacji o swojej konfiguracji i aktualnym stanie w celu zapewnienia poprawnego współdziałania w środowisku sieciowym. Zakres jednak tych informacji powinien być przeanalizowany i ograniczony do koniecznego minimum, w celu jak najmniejszej ekspozycji tychże systemów na potencjalne ataki. Informacje, które muszą być dostępne w Internecie (np. dotyczące certyfikatów SSL/TLS, danych rejestrowych i kanałów komunikacyjnych), a stanowią zagrożenie w przypadku ich identyfikacji poprzez OSINT, powinny stanowić podstawę do wzmożonego poziomu bezpieczeństwa w obszarze, w którym zidentyfikowane w ten sposób dane mogą być wykorzystane do ataku.

Dodatkowo w rozdziale tym wskazano zabezpieczenia, które należy wziąć pod uwagę w zakresie bezpieczeństwa osobowego. Dotyczy to głównie nadmiarowego udostępniania informacji przez osoby w Internecie spowodowanego brakiem świadomości konsekwencji tego typu działań oraz nieprawidłowych ustawień bezpieczeństwa w aplikacjach i serwisach internetowych.

Także zagrożenia dla bezpieczeństwa operacji, w większości przypadku prowadzonych w ramach działań wojska lub służb, zostały wskazane i opisane. Metody zabezpieczania nadmiarowego udostępniania danych w zakresie opisanych działań także stanowią element, którego wzrost znaczenia związany jest z rozwojem i przenikaniem sieciowych usług cyfrowych do innych aspektów życia, także zawodowego.

Czwarty rozdział merytoryczny, w którym przedstawiono możliwe do wprowadzenia zalecenia bezpieczeństwa w zakresie przeciwdziałania wywiadowi

otwartoźródłowemu, stanowi zestawienie metod i technik ochrony, zarówno infrastruktury firmowej lub organizacyjnej, jak i osobowej. Wskazano dwa obszary, w których wykorzystywany jest OSINT – w zakresie rozpoznania podmiotu będącego celem, jak i samoweryfikacji zabezpieczeń, poprzez stworzenie programu testów zakresu możliwego pozyskania informacji przy wykorzystaniu technik OSINT-owych i wprowadzenie odpowiednich zabezpieczeń przeciwko potencjalnym atakom z ich wykorzystaniem.

Wskazano normy, zestawienia i techniki, które mogą stanowić bazę do wprowadzania metod zabezpieczeń. Określono zasady, które muszą zostać zachowane, aby zminimalizować ryzyko działań w Internecie, wraz ze wskazaniem ich rozróżnienia dla obszaru bezpieczeństwa biznesowego i osobistego. Przedstawione zostały podstawy i metody rozszerzenia wskazanych zasad zabezpieczeń na otoczenie kluczowych osób w organizacjach, ze względu na ich dostęp do cennych informacji i decyzji, które mogą stanowić obiekt zainteresowania osób lub organizacji atakujących.

Niniejsza praca stanowi jedno z niewielu opracowań w zakresie OSINT-u, zarówno w zakresie prac polskojęzycznych, jak i angielskojęzycznych. Określa ona kierunki postępowania, które w świecie szybkiego rozwoju i zmian, dotyczących łączenia się sfery cyfrowej z życiem osób oraz działalnością firm i organizacji. Może to stanowić bazę do dalszej analizy zidentyfikowanych tematów i wypracowywania metod ochrony przed OSINT-em oraz jego bezpiecznego prowadzenia. Mocną stronę pracy stanowi unikalne zestawienie traktujące tematykę przekrojowo, z ujęciem opinii ekspertów w zakresie OSINT-u i wskazujące techniczne aspekty tematu.

Bibliografia

Literatura:

Bazzell M., *Open Source Intelligence Techniques. Resources for Searching and Analyzing Online Information*, Eighth Edition, 2021

Bean H., *Is Open Source Intelligence an Ethical Issue?*, Research in Social Problems and Public Policy, Volume 19

Chodyński A., *Podnoszenie poziomu bezpieczeństwa. Metody i narzędzia. Wprowadzenie* – Bezpieczeństwo. Teoria i praktyka, 2019, nr 4

Clark R. M., *Intelligence Analysis: A Target-Centric Approach*, CQ Press, 2019

Coffey M. L., *Library application of Deep Web and Dark Web technologies*, School of Information Student Research Journal, 10(1).

Dela P., *Charakterystyka zagrożeń bezpieczeństwa cyberprzestrzeni*, PWN, Warszawa 2022

Dela P., *Elementy propagandy w życiu publicznym*, Studia Politologiczne 54

Dylan H., *Defence Intelligence and the Cold War: Britain's Joint Intelligence Bureau 1945-1964*, OUP Oxford, 2014

Górski S., *O metodach badawczych w naukach społecznych*, Nauki o bezpieczeństwie. Wybrane problemy badań, Wydawnictwo CNBOP-PIB, 2017

Heuer R. J., *Psychology of Intelligence Analysis*, Center For The Study Of Intelligence, Central Intelligence Agency, 1999

Hudnall R. K., *No Safe Haven: Homeland Insecurity*, Omega Press, El Paso, Texas, 2004

Johnson L. M., *Establishing Broadcast Monitoring as Open Source Intelligence: The BBC Monitoring Service during the Second World War*, King's College London, 2013

Król K., *Geoinformation in the invisible resources of the Internet*, Geomatics, Landmanagement and Landscape No. 3, 2019

Lowenthal M. M., *Intelligence: From Secrets to Policy*, wyd. 6, CQ Press, 2014

Lowenthal M. M., Clark R. M., *The Five Disciplines of Intelligence Collection*, CQ Press, 2015

Krizan L., *Intelligence essentials for everyone*, Joint Military Intelligence College, Waszyngton, 1999

Matela K., *Wybrane aspekty systemów wywiadu, obserwacji i rozpoznania (ISR)*, WIEDZA OBRONNA, 2021, Vol. 276 No. 3

Rumsfeld D., *Known and Unknown: A Memoir*, Penguin, 2011

Saleem J., Islam R., Kabir M. A., *The Anonymity of the Dark Web: A Survey*, IEEE Access, vol. 10

Schneider C. J., Schneider D., *World War II*, Infobase Publishing, Nowy Jork 2014

Stróżyk J., *Wybrane problemy międzynarodowej współpracy wywiadowczej. Czy NATO ma wywiad?*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa, 2020

Tversky A.; Kahneman D., *Availability: A heuristic for judging frequency and probability*, Cognitive Psychology, Volume 5, Issue 2, September 1973

Tversky A.; Kahneman D., *Judgment under Uncertainty: Heuristics and Biases*, Science, New Series, Vol. 185, No. 4157. (27.09.1974)

Wiszniewski L., *Rola i znaczenie analizy informacji wywiadowczej w zapewnianiu bezpieczeństwa państwa*, Przegląd Bezpieczeństwa Wewnętrznego, 2020 nr 22 (12)

Witlin L., *Of Note: Mirror-Imaging and Its Dangers*, SAIS Review of International Affairs, Johns Hopkins University Press, Volume 28, Number 1, Winter-Spring 2008

Dokumenty:

ATP 2-33.4. Intelligence Analysis, Styczeń 2020

Berkeley Protocol on Digital Open Source Investigations, HR/PUB/20/2 (advance version)

DiSIEM Project Deliverable D4.1 – Techniques and tools for OSINT-based threat analysis

Feeling the Burden. Ethical Challenges and Practices in Open Source Analysis and Journalism

Intelligence Community Directive Number 301, National Open Source Enterprise, 2006

Joint Chiefs of Staff, Joint Publication 2-0: Joint Intelligence, 2013

National Defense Authorization Act for Fiscal Year 2006

National Security Agency – Purple Dragon. The Origin and Development of the United States OPSEC Program

NATO OSINT Handbook v.1.2

The 9/11 Commission Report

Słowniki:

AAP-6 (2019) PL: Słownik terminów i definicji NATO

Słownik języka polskiego PWN

Słownik języka polskiego sjp.pl

Źródła internetowe:

A Lance Corporal's Phone Selfie Got His Marine Unit 'Killed' at 29 Palms – https://www.military.com/daily-news/2020/01/07/lance-corporals-phone-selfie-got-his-marine-unit-killed-29-palms.html

AltaVista Photo Finder, and how to keep your images "unfound" – http://www.photodude.com/av.htm

AltaVista Photo Finder Has Artists Concerned – https://web.archive.org/web/19990427131502/http://www.searchenginewatch.com/sereport/9811-photofinder.html

Benjamin Brown, *Cognitive Bias and Critical Thinking in Open Source Intelligence (OSINT)* – https://www.youtube.com/watch?v=bWjEgd-KSHY

Cognitive biases – acaps technical brief, https://www.acaps.org/sites/acaps/files/resources/files/acaps_technical_brief_cognitive_biases_march_2016.pdf

K. Wosiński, *Comparison of reverse image searching in popular search engines*, https://research.securitum.com/comparison-of-reverse-image-searching-in-popular-search-engines-osint-hints

D-Day 75: How was the biggest ever seaborne invasion launched?, https://www.bbc.co.uk/teach/d-day-how-was-the-biggest-ever-seaborne-invasion-launched/zrrs7nb

Dark Web Searching, https://osintcombine.com/post/dark-web-searching

DoD OPSEC for Families – https://www.slideshare.net/DepartmentofDefense/opsec-for-families

Central Intelligence Agency, *Factbook on Intelligence*, https://irp.fas.org/cia/product/facttell/index.html

Facebook – Służba Bezpieczeństwa Ukrainy – https://www.facebook.com/SecurSerUkraine/videos/3153483434931349

File types indexable by Google – https://support.google.com/webmasters/answer/35287?hl=en

Fitness app Strava lights up staff at military bases - https://www.bbc.com/news/technology-42853072

How many active sites are there? – https://www.netcraft.com/active-sites

How Much of the Internet is the Dark Web in 2022?, https://techjury.net/blog/how-much-of-the-internet-is-the-dark-web/#gref

How Open-Source Intelligence Is Helping Clear The Fog Of War In Ukraine – https://www.buzzfeednews.com/article/peteraldhous/osint-ukraine-war-satellite-images-plane-tracking-social

Federation of American Scientists, *Intellligence Resource Program, Measurement and Signature Intelligence (MASINT)* – https://irp.fas.org/program/masint.htm

Internet Live Stats: Total number of Websites – https://www.internetlivestats.com/total-number-of-websites

Introducing 15 cm HD: The Highest Clarity From Commercial Satellite Imagery, Chris Formeller, https://blog.maxar.com/earth-intelligence/2020/introducing-15-cm-hd-the-highest-clarity-from-commercial-satellite-imagery

K. Wosiński, *Jak wyszukiwarki radzą sobie z analizą zawartości* obrazów, https://sekurak.pl/jak-wyszukiwarki-radza-sobie-z-analiza-zawartosci-obrazow-osint-hints

July 2022 Web Server Survey – https://news.netcraft.com/archives/2022/07/28/july-2022-web-server-survey.html

Knocking down barriers to knowledge, https://googleblog.blogspot.com/2011/06/knocking-down-barriers-to-knowledge.html

Most popular social networks worldwide as of January 2022, ranked by number of monthly active users (in millions), https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users

Nastolatek uzyskał zdalny dostęp do ponad 25 Tesli na całym świecie. Lokalizowanie samochodów, otwieranie drzwi, wystartowanie auta… – https://sekurak.pl/nastolatek-uzyskal-zdalny-dostep-do-ponad-25-tesli-na-calym-swiecie-lokalizowanie-samochodow-otwieranie-drzwi-wystartowanie-auta

News Watch; A Quick Way to Search For Images on the Web – https://www.nytimes.com/2001/07/12/technology/news-watch-a-quick-way-to-search-for-images-on-the-web.html

Number of global social network users 2018-2027, https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users

Ooh! Ahh! Google Images presents a nicer way to surf the visual web – https://googleblog.blogspot.com/2010/07/ooh-ahh-google-images-presents-nicer.html

Federation of American Scientists, The Interagency OPSEC Support Staff, *Operations Security, Intelligence Threat Handbook*, 1996, https://irp.fas.org/nsa/ioss/threat96/index.html

Operations Security (OPSEC) Guidance for Family Members – https://www.2ndmardiv.marines.mil/Portals/47/Docs/Operations%20Security%20Guidance%20for%20Family%20Members.pdf

OPSEC - A Guide For Family and Friends Presented by 1st Information Operations Command (Land) Vulnerability Assessment Detachment Army OPSEC Support Element – https://www.usace.army.mil/Portals/2/docs/Careers/OPSEC-A%20Guide%20for%20Families%20and%20Friends.pdf

OPSEC 101 for Military Spouses: Help Keep Your Family and Service Member Safe – https://themilitarywifeandmom.com/opsec-for-military-spouses

K. Zetter, *Palin E-Mail Hacker Says It Was Easy* – https://www.wired.com/2008/09/palin-e-mail-ha

Poszukiwany za przekręt na 50 mln zł złapany, bo jego młoda partnerka chwaliła się życiem w mediach społecznościowych – https://www.antyradio.pl/News/Poszukiwany-za-przekret-na-50-mln-zl-zlapany-bo-jego-mloda-partnerka-chwalila-sie-zyciem-w-mediach-spolecznosciowych-38083

C. Lekati – *Protecting High-Value Individuals: An OSINT Workflow* – https://www.youtube.com/watch?v=rE4mORq9T5s

Relevance meets the real-time web, https://googleblog.blogspot.com/2009/12/relevance-meets-real-time-web.html

Report: State of the Web, https://httparchive.org/reports/state-of-the-web

Retired Chicago Firefighter Wrongly Accused of Participating in Capitol Violence, https://www.firefighternation.com/leadership/retired-chicago-firefighter-wrongly-accused-of-participating-in-capitol-violence

Trace Labs, *Search Party Rules of Engagement* – https://www.tracelabs.org/about/search-party-rules

Przez TeamViewer-a aż do systemu kontroli uzdatniania wody. Hacker zmienił parametry chemiczne wody. – https://sekurak.pl/przez-teamviewer-a-az-do-systemu-kontroli-uzdatniania-wody-hacker-zmienil-parametry-chemiczne-wody

Similar Images graduates from Google Labs,
https://googleblog.blogspot.com/2009/10/similar-images-graduates-from-google.html

J. Long*, The Google Hacker's Guide. Understanding and Defending Against the Google Hacker* – https://pdf.textfiles.com/security/googlehackers.pdf

Twitter – @Marsattaqueblog:
https://twitter.com/Marsattaqueblog/status/1582039213591040000

N. Dekens, *Vicarious trauma and OSINT – a practical guide*,
https://osintcurio.us/2020/06/08/vicarious-trauma-and-osint-a-practical-guide

Zapis wideo z konferencji prasowej w temacie włamania do stacji uzdatniania wody na Florydzie – https://www.youtube.com/watch?v=MkXDSOgLQ6M

Spis rysunków

Spis tabel

Wykaz skrótów

API – ang. *Application Programming Interface* – interfejs programistyczny aplikacji.

CAPTCHA – ang. *Completely Automated Public Turing test to tell Computers and Humans Apart* – technika stosowana na stronach internetowych w celu upewnienia się, że inicjatorem danej akcji (wysyłania lub pobierania danych) jest człowiek.

CBIR – ang. *Content-based image retrieval* – pozyskiwanie obrazów na podstawie ich graficznej zawartości.

CERT – ang. *Computer Emergency Response Team* – zespół reagowania na incydenty bezpieczeństwa.

CWE – ang. *Common Weakness Enumeration* – opisana znana kategoria słabości sprzętu lub oprogramowania.

DNS – ang. *Domain Name System* – system nazw sieciowych, umożliwiający rozwiązywanie nazw domen na adresy IP.

GAN – ang. *Generative Adversarial Networks* – generatywne sieci współzawodniczące.

HUMINT – ang. *Human Intelligence* – wywiad bazujący na informacjach pozyskanych ze źródeł osobowych.

IMINT – ang. *Imagery Intelligence* – wywiad bazujący na danych obrazowych.

MASINT – ang. *Measurement and Signatures Intelligence* – wywiad prowadzony na podstawie danych pozyskanych z różnych źródeł, w celu określenia cech specyficznych dla źródeł przechwytywanego sygnału.

OPSEC – ang. *Operations Security* – bezpieczeństwo prowadzenia operacji, głównie stosowane w odniesieniu do działań militarnych.

OSD – ang. *Open Source Data* – zebrane dane w formie nieprzetworzonej jeszcze przez analityków.

OSINF – ang. *Open Source Information* – zebrane dane po etapie ich przetwarzania i filtrowania.

OSINT – ang. *Open Source Intelligence* – wywiad oparty na otwartych, ogólnodostępnych źródłach.

PERSEC – ang. *Personal Security* – bezpieczeństwo osobowe, dotyczące aspektów prywatnego życia osób, głównie stosowane w odniesieniu do żołnierzy i ich rodzin.

SIEM – ang. *Security Information and Event Management* – oprogramowanie służące do zbierania i analizowania zagrożeń bezpieczeństwa sieci.

SIGINT – ang. *Signals Intelligence*) – termin ogólny określający wywiad radiowy i elektroniczny.

SMTP – ang. *Simple Mail Transfer Protocol* – protokół wykorzystywany do komunikacji za pomocą wiadomości poczty elektronicznej.

SOCMINT – ang. *Social Media Intelligence* – wywiad bazujący na informacjach pozyskanych z mediów społecznościowych.

SSL – ang. *Secure Socket Layer* – protokół wykorzystywany do realizacji szyfrowanych połączeń sieciowych, zastąpiony przez swoją rozszerzoną wersję – TLS.

TLS – ang. *Transport Layer Security* – protokół umożliwiający tworzenie szyfrowanych kanałów komunikacyjnych w sieci, następca SSL.

URL – ang. *Uniform Resource Locator* – ujednolicony format adresowania zasobów.

VPN – ang. *Virtual Private Network* – wirtualna sieć prywatna, umożliwiająca stworzenie zabezpieczonego połączenia tunelowego pomiędzy komputerem a serwerem.

Załącznik

AKADEMIA KALISKA

im. Prezydenta Stanisława Wojciechowskiego

WYWIAD EKSPERCKI

Opracował:
mgr inż. Krzysztof Wosiński

pod kierownictwem naukowym:
płk. rez. dr. hab. inż. Piotra Deli

Kalisz 2023

I WPROWADZENIE

Wykorzystanie rozpoznania bazującego na ogólnodostępnych źródłach informacji jest obecnie powszechnie stosowanym podejściem w wielu obszarach działań cywilnych: biznesowym, dochodzeniowym, dziennikarskim, a także w prywatnych działaniach pojedynczych osób. Mnogość i łatwość dostępu do narzędzi, ułatwiających tego typu działania, a także szeroki wachlarz szkoleń i materiałów edukacyjnych, dostępnych w tym zakresie w Internecie powoduje, że konieczna staje się analiza wpływu przedmiotowego zjawiska na bezpieczeństwo zarówno na poziomie technologii, jak i bezpieczeństwa osobowego.

W związku z powyższym, przedmiotem pracy badawczej było określenie zakresu dostępnych narzędzi i technik w zakresie ustalenia możliwości wykorzystania ich przez organizacje cywilne (działania militarne, wykorzystujące rozpoznanie OSINT-owe pozostają poza zakresem niniejszej dysertacji) oraz indywidualne osoby do działań ofensywnych w zakresie bezpieczeństwa. Istotną kwestią, która wymaga uwzględnienia jest także wpływ procesów analitycznych i zasady, które muszą być zachowane podczas ich przeprowadzania (w tym uwzględnienie aspektów psychologicznych, mających wpływ na wspomniane procesy), gdyż ich nieprawidłowe wykonanie może prowadzić do sfałszowanych wyników i także wpływa na wynikowe bezpieczeństwo, szczególnie w aspekcie osobowym.

Analizując wymagania norm międzynarodowych, dotyczących bezpieczeństwa systemów teleinformatycznych, a także wytycznych, zawartych w publikacjach dotyczących bezpieczeństwa prowadzenia operacji oraz bezpieczeństwa osobowego, jako cel pracy obrano także zbadanie wpływu przedmiotowych dokumentów na wprowadzenie skuteczniejszych zabezpieczeń przed działaniem OSINT-u.

Przedmiotem prowadzonych badań była też ocena możliwości zastosowania działań defensywnych, bazujących na znajomości technik rozpoznania otwartoźródłowego.

W celu weryfikacji założonych w tej dysertacji hipotez, zadano ekspertom następujące pytania:

1. Jak ocenia Pani/Pan rozwój możliwości rozpoznania prowadzonego z wykorzystaniem otwartych źródeł (OSINT-u) w ostatnich latach w zakresie wykorzystywania ich przez osoby i organizacje cywilne?

2. Jaki obszar wywiadu otwartoźródłowego uznaje Pani/Pan za najlepiej rozwinięty aktualnie i co Pani/Pana zdaniem stanowi główną podstawę rozwoju przedmiotowego obszaru?

3. Czy uważa Pani/Pan, że wykorzystanie do rozpoznania otwartoźródłowego narzędzi płatnych wpisuje się w zakres przedmiotowych działań, czy też poprzez wprowadzenie ograniczenia finansowego nie należy źródeł płatnych wpisywać w zakres technik OSINT-owych?

4. Proszę o wskazanie jakie narzędzia i techniki OSINT-owe według Pani/Pana dają użytkownikom Internetu możliwość zdobycia informacji, które mogą stanowić zagrożenie dla bezpieczeństwa systemów teleinformatycznych i/lub bezpieczeństwa osobowego?

5. Jakie zasady poprawnego prowadzenia działań OSINT-owych uznaje Pani/Pan za najważniejsze w celu uniknięcia błędnych wniosków z nich płynących?

6. Czy według Pani/Pana oceny indywidualne cechy psychologiczne osoby prowadzącej rozpoznanie i analizę mają wpływ na poprawne prowadzenie rozpoznania OSINT-owego i jeśli tak, to na jakie kwestie psychologiczne należy zwrócić szczególną uwagę w tym zakresie?

7. Jakie zagrożenia dla infrastruktury teleinformatycznej widzi Pani/Pan w kontekście możliwości rozpoznania jej słabości poprzez działania OSINT-owe?

8. Jakie zagrożenia dla bezpieczeństwa osobowego widzi Pani/Pan w kontekście możliwości rozpoznania działań osoby w Internecie poprzez techniki i narzędzia OSINT-owe?

9. Czy Pani/Pana zdaniem narzędzia i techniki OSINT-owe można wykorzystać do działań defensywnych, zarówno w kontekście weryfikacji bezpieczeństwa infrastruktury teleinformatycznej, jak i bezpieczeństwa osobowego oraz czy istnieje równowaga pomiędzy możliwościami ofensywnymi i defensywnymi w przedmiotowym zakresie?

10. W kontekście zasad wynikających z norm i publikacji międzynarodowych, dotyczących bezpieczeństwa teleinformatycznego (jak np. rodzina norm ISO 27000 lub podobne publikacje), a także wytycznych bezpiecznego prowadzenia operacji (OPSEC) i bezpieczeństwa personelu (PERSEC), jak ocenia Pani/Pan ich wpływ na możliwości w zakresie przeciwdziałania wykorzystaniu technik OSINT-owych?

11. Uwagi końcowe.

II SPRAWOZDANIE Z BADANIA EKSPERTÓW

1. Temat badań

Znaczenie wywiadu opartego na otwartych źródłach (OSINT) w zapewnieniu bezpieczeństwa systemów teleinformatycznych i bezpieczeństwa osobowego.

2. Metoda badawcza

Wywiad ekspercki.

3. Cel badań

Uzyskanie opinii ekspertów w celu zebrania materiału faktograficznego i specjalistycznego, dotyczącego badanego zagadnienia. Jako najistotniejszy element z punktu widzenia celu badań uznano posiadanie przez ekspertów specjalistycznej wiedzy w zakresie specyfiki wywiadu otwartoźródłowego, doświadczenia analitycznego oraz wprowadzania i weryfikacji zabezpieczeń przez przedmiotowymi technikami. Grono ekspertów reprezentowane było przez następujące grupy: specjaliści w zakresie bezpieczeństwa i zarządzania informacjami z obszarów: biznesowego, wojskowego oraz policyjnego, a także międzynarodowi trenerzy w zakresie wykorzystania technik i narzędzi OSINT-owych.

4. Czas badań

Badania przeprowadzono pomiędzy kwietniem 2022 r. a sierpniem 2022 r. włącznie.

5. Opis przebiegu badań

Badanie przeprowadzono przy wykorzystaniu kwestionariusza wywiadu eksperckiego, zawierającego 10 pytań oraz punkt dotyczący uwag końcowych,

wymagających wyrażenia przez ekspertów swoich poglądów, sformułowanych z wykorzystaniem swojej wiedzy naukowej i empirycznej. Odpowiedzi na zadane pytania stanowiły pomoc w rozwiązaniu problemu głównego oraz problemów szczegółowych. Badania przeprowadzono w stosunkowo niewielkim przedziale czasu, co pozwoliło na zachowanie stałości środowiska bezpieczeństwa w zakresie technik prowadzenia rozpoznania z wykorzystaniem otwartych źródeł, rodzajów zabezpieczeń oraz wykorzystywanych w tym zakresie technologii.

III WNIOSKI

W odniesieniu do pytania nr 1: ***jak ocenia Pani/Pan rozwój możliwości rozpoznania prowadzonego z wykorzystaniem otwartych źródeł (OSINT-u) w ostatnich latach w zakresie wykorzystywania ich przez osoby i organizacje cywilne?***

W tym zakresie eksperci zgodnie zauważają znaczący wzrost zainteresowania wywiadem otwartoźródłowym w ostatnich latach. Spowodowane jest to wieloma czynnikami, z których można jednak wyróżnić takie, jak: ciągły wzrost zainteresowania i wykorzystywania mediów społecznościowych, powstawanie nowych portali umożliwiających zdobywanie nowych umiejętności poprzez zadania konkursowe, rozwój społeczności internetowych, skupionych wokół tematu OSINT-u oraz pandemię COVID-19, która sprawiła, że wiele wydarzeń (takich jak konferencje, szkolenia i wykłady), zmieniło formę na zdalną, co znacznie ułatwiło zainteresowanym osobom czerpanie z nich nowej wiedzy niezależnie od miejsca zamieszkania i możliwości podróżowania.

Nie bez wpływu jest też zwiększone zainteresowanie mediów sprawami ciekawych śledztw OSINT-owych, co przełożyło się na stworzenie seriali dokumentalnych oraz programów, które w efekcie przyniosły falę jeszcze większego zainteresowani tą tematyką.

Wskazany wzrost zainteresowania dotyczy nie tylko osób prywatnych, które oprócz rozwijania własnych zainteresowań wykorzystują w praktyce zdobyte umiejętności np. do analizy fake-newsów, weryfikacji osób, z którymi zaczynają się spotykać, zatrudnianych opiekunek do dzieci, trenerów itp., ale także firm i instytucji, które w narzędziach i technikach OSINT-owych widzą możliwości sprawdzenia planowanych do zatrudnienia pracowników oraz weryfikacji informacji lub analizy wykroczeń i przestępstw – te możliwości są szczególnie często wykorzystywane np. przez organizacje broniące praw człowieka, agencje prasowe oraz stróżów prawa. Szczególnie w ostatnich czasach można zaobserwować powrót do dziennikarstwa śledczego w zakresie analizy weryfikacji prawdziwości publikowanych (szczególnie w Internecie) informacji.

Niestety eksperci wskazują także na negatywne skutki nagłego wzrostu zainteresowania tematyką OSINT-u, zarówno jeśli chodzi o osoby prywatne, jak i firmy.

Wiele z nich zrównuje OSINT z efektywnym wyszukiwaniem informacji (zazwyczaj ograniczając się do wyszukiwarki Google), pomijając całkowicie wysiłek, który należy włożyć w procedury rozpoznania oraz odpowiednią analizę zebranych informacji, które stanowią bardzo ważny aspekt wywiadu.

W odniesieniu do pytania nr 2: ***jaki obszar wywiadu otwartoźródłowego uznaje Pani/Pan za najlepiej rozwinięty aktualnie i co Pani/Pana zdaniem stanowi główną podstawę rozwoju przedmiotowego obszaru?***

Eksperci mieli tutaj mocno podzielone zdania, jednak najczęściej wskazywaną dziedziną OSINT-u, którą można uznać za najlepiej rozwiniętą jest pozyskiwanie informacji z mediów klasycznych i społecznościowych. Za przyczynę takiej oceny były przez ekspertów wskazywane: rozwój narzędzi do skanowania zasobów Internetu (w tym aplikacji do automatycznego skanowania), duża ilość informacji możliwych do pozyskania z mediów społecznościowych, a także szybki rozwój wspomnianych narzędzi. Praktycznie w każdym tygodniu powstają nowe narzędzia lub udostępniane są aktualizacje już istniejących, rozszerzające ich możliwości. Oprócz samego wyszukiwania informacji, eksperci wskazują także na funkcjonalności automatycznej archiwizacji pozyskanych danych, co znacznie ułatwia pracę. Jednocześnie, pomimo dużego znaczenia narzędzi, wskazany został fakt, że nie implikuje to automatycznie największego znaczenia tej gałęzi rozpoznania otwartoźródłowego, a jedynie jej rozwój.

Drugim najczęściej wskazywanym obszarem, w którym eksperci zauważają znaczny rozwój lub wzrost zainteresowania, jest goelokalizacja. Jako przyczyny takiego stanu rzeczy wskazywane są zarówno aktywne wykorzystywanie tego typu umiejętności w przypadku analizy informacji z obszarów konfliktów i wojen, ale również łatwość przyswajania umiejętności geolokalizacji, które można wykorzystywać nie tylko do działań służbowych, ale także prywatnych.

Wśród ekspertów pojawiła się także opinia, że trudno jest jednoznacznie wskazać konkretny obszar, który jest najlepiej rozwinięty, gdyż tak jak w innych dziedzinach, tak i w zakresie OSINT-u działa wiele osób z szeroką wiedzą na temat rozpoznania, ale specjalizujących się w określonym, wąskim przedmiocie działań OSINT-owych. Obecność wielu ekspertów dziedzinowych w środowisku specjalistów OSINT-owych

stanowi bazę do wykorzystania ich unikalnych zdolności i sposobów pracy w ramach potrzeb artykułowanych przez podmioty rynkowe i wymagania klientów.

W odniesieniu do pytania nr 3: ***czy uważa Pani/Pan, że wykorzystanie do rozpoznania otwartoźródłowego narzędzi płatnych wpisuje się w zakres przedmiotowych działań, czy też poprzez wprowadzenie ograniczenia finansowego nie należy źródeł płatnych wpisywać w zakres technik OSINT-owych?***

Wszyscy eksperci jednogłośnie stwierdzili, że fakt konieczności zapłaty za możliwość skorzystania z narzędzi OSINT-owych nie wyklucza ich z zakresu działań otwartoźródlowych, gdyż fakt, że dane są „otwarte" (ogólnodostępne) nie musi implikować darmowego dostępu do nich, a kryteria finansowe nie powinny być, w przeciwieństwie do prawnych, głównym kryterium dla decydentów.

Dla poparcia tego stanowiska eksperci wskazali sytuacje, w których osoby profesjonalnie zajmujące się rozpoznaniem otwartoźródłowym muszą korzystać z płatnych narzędzi, głównie ze względu na ich efektywność, która pozwala na zaoszczędzenie dużej ilości czasu podczas pracy. Także skalowalność płatnych narzędzi jest ważnym kryterium ich wyboru, gdyż w miarę coraz większego rozrastania się Internetu, analitycy muszą sprostać temu wyzwaniu w celu wydajnego pozyskiwania danych i przetwarzania ich na odpowiedniej jakości dane wywiadowcze. Tego typu narzędzia pozwalają na szybsze i wydajniejsze pozyskiwanie danych niż możliwe by to było w przypadku ręcznego wyszukiwania informacji. Jednym z wskazanych przykładów tego typu usprawnień są dane z wycieków, które z jednej strony mogą być szybciej przeanalizowane w przypadku przeszukiwania wielu rozległych baz danych, a z drugiej strony zmniejszają ryzyko ujawnienia się analityka.

Wiele płatnych narzędzi nie ma swoich darmowych odpowiedników, co sprawia, że ich wykorzystanie staje się koniecznością. Przykładami mogą tutaj być dane historyczne o hostingu stron internetowych, narzędzia do zbierania i przechowywania dowodów czy platformy udostępniające dane o ruchu lotniczym.

Wraz jednak z koniecznością zapłaty za korzystanie z danego oprogramowania, idzie też brak możliwości analizy jej kodu źródłowego, co z kolei uniemożliwia analitykom sprawdzenie skąd i w jaki sposób zbierane oraz jak przetwarzane są dane, na których później budowany jest obraz rozpoznania OSINT-owego. Wiele osób przyjmuje bezkrytycznie działanie „czarnych skrzynek" jako godne zaufania, podczas gdy dają one

tylko wycinek pełnego zakresu informacji, bez dokładnego ich pochodzenia, kontekstu czy wiarygodności źródła. Tego typu narzędzia zdaniem jednego z ekspertów nie powinny być stosowane w działaniach OSINT-owych.

W odniesieniu do pytania nr 4: ***proszę o wskazanie jakie narzędzia i techniki OSINT-owe według Pani/Pana dają użytkownikom Internetu możliwość zdobycia informacji, które mogą stanowić zagrożenie dla bezpieczeństwa systemów teleinformatycznych i/lub bezpieczeństwa osobowego?***

Eksperci wskazali tutaj głównie na niebezpieczeństwa płynące z używania narzędzi, które same w sobie mogą stanowić problem bezpieczeństwa. Po raz kolejny, podobnie jak w odpowiedzi na pytanie 3, wyartykułowane zostało zalecenie weryfikacji sposobu działania narzędzia oraz organizacji lub osoby stojącej za jego wytworzeniem i dystrybucją. Szczególnym obszarem zainteresowania powinny być tutaj informacje, które narzędzia przekazują do firm je obsługujących, a także ile danych na temat analityka udostępniane jest obiektowi rozpoznania podczas prac śledczych. Ten aspekt musi być brany pod uwagę w zakresie wpływu na bezpieczeństwo misji lub osoby prowadzącej śledztwo.

Szczególnie w przypadku działań o dużej poufności i wrażliwości tematu należy zwrócić uwagę na kwestię możliwego przekazywania pozyskanych informacji do organizacji lub osób obsługujących wykorzystywane narzędzie.

Podobnie jak w odpowiedzi na pytanie 3, jeden z ekspertów wskazał na konieczność weryfikacji kodu wykorzystywanych narzędzi. O ile takie programy jak Maltego czy Hunchly mogą być godne zaufania poprzez wiele lat obecności na rynku oraz ich weryfikację przez najlepszych specjalistów, o tyle mniej znane narzędzia o wątpliwej reputacji (jak np. narzędzie Lampyre, którego powiązania ze służbami rosyjskimi nie można wykluczyć[164]) nie powinny być wykorzystywane.

[164] *Be careful what you OSINT with*, https://keyfindings.blog/2020/03/23/be-careful-what-you-osint-with – dostęp online 30.06.2022 r.

W odniesieniu do pytania nr 5: ***jakie zasady poprawnego prowadzenia działań OSINT-owych uznaje Pani/Pan za najważniejsze w celu uniknięcia błędnych wniosków z nich płynących?***

Ogólnym wnioskiem płynącym z odpowiedzi na to pytanie jest stwierdzenie, że nie da się wyróżnić jednej zasady czy metodologii prowadzenia rozpoznania otwartoźródłowego, gdyż zasady różnią się w zależności od organizacji oraz kraju, w którym działania te są wykonywane, m.in. ze względu na różnorodność środowiska czy obowiązujących systemów prawnych (przykładem w przypadku obszaru Unii Europejskiej może być np. dyrektywa RODO).

Wśród opinii ekspertów, które w większości pokrywały się ze sobą, możliwe jest jednak wyodrębnienie trzech głównych, ogólnych wskazówek:

1. Weryfikacja informacji i poddawanie wszystkich pozyskanych w ramach rozpoznania danych pod wątpliwość. W związku z ogromem dezinformacji w Internecie, umiejętność zidentyfikowania jej jest jedną z kluczowych umiejętności podczas określania przydatności i poprawności pozyskanych z otwartych źródeł danych. Korzystanie z wielu różnorodnych źródeł, które mogą zostać sprawdzone pod kątem ich wiarygodności, a także zadawanie dodatkowych pytań i poszukiwanie odpowiedzi na nie w celu pełnego potwierdzenia hipotezy, jaką jest każda informacja znaleziona w Internecie.

2. Krytyczne myślenie i niewyciąganie pochopnych wniosków z pozyskanych informacji. W tym zakresie kluczowym jest skupienie się na faktach i raportowanie wyłącznie ich, bez rozszerzania wniosków o prawdopodobne scenariusze. Zastosowanie pełnego cyklu wywiadu (opisywanego szerzej w rozdziale 0) jest kluczem do wykonania poprawnej analizy. Wynikiem analizy OSINT-owej powinien być raport, na podstawie którego dopiero czytelnik wyciągnie wnioski.

3. Znajomość zasad psychologicznych działania ludzkiego umysłu, w tym głównie błędów poznawczych i utartych ścieżek w myśleniu. Poznanie samego siebie i własnych sposobów oceny analizowanych sytuacji, ograniczeń umysłowych, a przede wszystkim wzorców postrzegania rzeczywistości i zachowania opartego na nich jest podstawą do prowadzenia poprawnych, niezakłóconych i niespolaryzowanych analiz. Należy pamiętać, że wszystkie wnioski mają równą ważność, niezależnie od tego czy popierają, czy obalają one postawioną na początku hipotezę oraz czy wpisują się w nasze przekonania światopoglądowe, czy nie.

W zakresie poprawnego przeprowadzania analiz, zdaniem jednego z ekspertów pomocne może okazać się wykorzystanie wybranego z wielu znanych modeli analitycznych, wśród których można przywołać chociażby:

- ACH (skrót od ang. *Analysis of Competing Hypotheses*) – analiza wielu hipotez, której autorem jest Richards J. Heuer z CIA i która polega na analizie wielu, nawet mało prawdopodobnych hipotez oraz późniejszym obalaniu jak największej ilości z nich, co pozwala jak najbardziej zminimalizować wpływ błędów poznawczych na wyniki pracy analitycznej;
- AMITT (skrót od ang. *Adversarial Misinformation and Influence Tactics and Techniques*) – zestaw taktyk, technik i procesów wykorzystywanych podczas analizy przypadków dezinformacji. Wykorzystuje on m.in. framework MITRE ATT&CK w celu śledzenia i przeciwdziałania szerzeniu nieprawdziwych informacji poprzez wprowadzenie wspólnej taksonomii dla opisywania incydentów. Jest zarówno zestawem modeli dla zespołów zajmujących się obroną (tzw. „blue team"), jak i atakowaniem (tzw. „read team"), a także zawiera w sobie sposoby przeciwdziałania i przykłady dezinformacji. Od 2022 roku jest on częścią DISARM Disinformation TTP (Tactics, Techniques and Procedures) Framework[165].
- CRAAP (skrót od ang. *Currency, Relevance, Authority, Accuracy, and Purpose*) – metoda weryfikacji informacji, wykorzystywana głównie w obszarze akademickim. Polega ona na ocenie informacji w każdym z następujących pięciu zakresów pytań, z których czerpie swoją nazwę:
    - aktualności – kiedy po raz pierwszy pojawiła się dana informacja, czy była od tego czasu aktualizowana lub zmieniana, czy nadal jest aktualna i czy jej aktualność ma znaczenie;
    - związku z tematem – do kogo kierowany jest przekaz, czy ma związek z analizowanym tematem, czy nie jest zbyt ogólna lub zbyt szczegółowa w odniesieniu do analizowanego tematu oraz czy sprawdzono także inne źródła przed wybraniem tej informacji jako znaczącej;
    - autorytetu – jaka osoba lub organizacja jest twórcą / publikatorem informacji, czy posiada odpowiednie doświadczenie / powiązania

[165] *DISARM Disinformation TTP (Tactics, Techniques and Procedures) Framework*, https://github.com/DISARMFoundation/DISARMframeworks/ – dostęp online 04.07.2022 r.

organizacyjne / wykształcenie do tworzenia publikacji w danej tematyce, czy łatwo jest odnaleźć publikacje i dane kontaktowe danego autora oraz co zdradza domena serwisu, w którym publikowane są informacje[166];

- jakości – czy znane są intencje publikacji danej informacji, czy zachowane są zasady poprawnej pisowni i gramatyki, czy cytowane są źródła, czy tekst został sprawdzony przed publikacją, czy jest nacechowany emocjami oraz czy przedstawione poglądy znajdują poparcie w faktach;
- celu – jaki jest cel przyświecający publikacji informacji i czy jest on klarowny, czy tekst miał opisać zdarzenie / przedstawić opinię / nakłonić lub odwieść od czegoś / sprzedać produkt / nauczyć czegoś, czy wydźwięk jest obiektywny czy subiektywny oraz czy zawarte zostały w przekazie określenia lub opinie spolaryzowane pod względem religijnym, ideologicznym, kulturalnym, instytucjonalnym lub osobistym.

W odniesieniu do pytania nr 6: ***czy według Pani/Pana oceny indywidualne cechy psychologiczne osoby prowadzącej rozpoznanie i analizę mają wpływ na poprawne prowadzenie rozpoznania OSINT-owego i jeśli tak, to na jakie kwestie psychologiczne należy zwrócić szczególną uwagę w tym zakresie?***

We wszystkich opiniach, przedstawionych przez ekspertów, pojawiło się stanowisko określające indywidualne cechy psychologiczne osób prowadzących śledztwa jako bardzo ważny aspekt prawidłowego prowadzenia OSINT-u. Eksperci wskazywali na bardzo różnorodne cechy, takie jak: kompetencje komunikacyjne, wytrwałość, oddanie, otwartość umysłu, umiejętność wyciągania wniosków z porażek lub ślepych zaułków w śledztwach i szybkiego dochodzenia do pełni sprawności umysłowej po nich, a także ciekawość, pokorność i umiejętność analitycznego myślenia.

[166] Domena najwyższego poziomu, czyli końcowy ciąg znaków w nazwie domeny serwisu internetowego, określa zazwyczaj jej powiązania i tematykę, np.: .gov – serwisy rządowe, .edu – serwisy naukowe i edukacyjne, .com – serwisy komercyjne, .mil – serwisy wojskowe. Może występować także w powiązaniu z domeną krajową, np.: .edu.pl dla polskich serwisów edukacyjnych lub .com.pl dla polskich serwisów komercyjnych. Domeny .edu, .gov i .mil należą do tzw. sponsorowanych domen najwyższego poziomu, tzn. nie mogą wykupić z nich subdomeny osoby lub organizacje nie należące do konkretnej grupy zawodowej, geograficznej lub etnicznej, tak jak to ma miejsce w przypadku domen komercyjnych. W związku z brakiem nadzoru nad zgodnością treści publikowanych na stronach w domenach komercyjnych, nie ma żadnej pewności, że strony w tych domenach są wiarygodne, nawet pomimo swojego domniemanego powiazania nazwy serwisu z określoną tematyką.

Częścią ludzkiej natury jest także tendencja do błędów poznawczych, czyli postrzegania rzeczywistości w sposób irracjonalny, który wpływa na podejmowanie decyzji i wykonywanie czynności w sposób bezwiednie stronniczy. Obowiązkiem dobrego analityka OSINT-owego jest znajomość różnych błędów poznawczych i pozostawienie emocji oraz podświadomego podejmowania decyzji i wyciągania wniosków poza obrębem swojej pracy. Jedynie praca na faktach i ich weryfikacja przy użyciu określonych metod analitycznych pozwoli na jak największe oddzielenie spolaryzowanych poglądów od prowadzonego śledztwa.

W działaniach OSINT-owych często dobrymi analitykami okazują się osoby o określonych cechach psychologicznych, na przykład osoby z różnymi formami autyzmu, które są w stanie znakomicie poradzić sobie z bardzo szczegółowymi analizami lub odnajdywaniem trudnych do wychwycenia na pierwszy rzut oka powiązań. Także umiejętność odpowiedniej wizualizacji danych jest bardzo pomocna, gdyż pozwala na odnalezienie powiązań, które nie są widoczne w momencie, kiedy dane występują jedynie w formie spisanej.

Poza cechami mającymi pozytywny wpływ na prowadzenie OSINT-u, są oczywiście także cechy negatywne. Jedną z nich jest tendencja do niewystarczająco dobrze prowadzonego rozpoznania w przypadku tematów, które nie są lubiane przez danego analityka lub są dla niego nudzące. Wynikiem tego jest mniej energii wkładane w przeprowadzenie rzetelnego śledztwa niż w przypadku interesujących tematów. Takie podejście, nietraktujące każdego śledztwa z równą ważnością, stanowi niestety podstawę do popełniania błędów i tworzenia niepełnych raportów.

W odniesieniu do pytania nr 7: ***jakie zagrożenia dla infrastruktury teleinformatycznej widzi Pani/Pan w kontekście możliwości rozpoznania jej słabości poprzez działania OSINT-owe?***

Eksperci podzielili swoje opinie w tym temacie na dwie kategorie zagrożeń: dla własnej infrastruktury wykorzystywanej do procesu rozpoznania OSINT-owego oraz dla infrastruktury analizowanej. Druga grupa dotyczy jedynie zagrożeń wynikających z możliwości rozpoznania za pomocą technik OSINT-owych.

W ramach pierwszej z tych kategorii głównym aspektem, któremu instytucje powinny poświęcić więcej uwagi było opracowanie i przestrzeganie zasad bezpieczeństwa operacji (OPSEC) oraz bezpieczeństwa osobistego (PERSEC) przez pracowników obsługujących przedmiotową infrastrukturę. Przykładem nieprawidłowego postępowania może być rozpoznanie, prowadzone z wykorzystaniem mediów społecznościowych, jak np. Facebook, gdzie osoba nieznająca zasad działań rozpoznawczych na tego typu platformie korzysta z własnego profilu, dołącza do znajomych rozpoznawanej osoby i dopiero wtedy zbiera informacje. Podobnym zagrożeniem może być ujawnienie swoich działań podczas rozpoznania infrastruktury sieciowej analizowanej instytucji.

Zagrożenia wynikające z możliwości rozpoznania infrastruktury poprzez wykorzystanie technik OSINT-owych zostały sklasyfikowane przez ekspertów jako z jednej strony podatności techniczne, a z drugiej jako wykorzystanie słabości ludzkich. Te pierwsze zawierają możliwość rozpoznania budowy infrastruktury serwerowej (m.in. lista otwartych portów), serwisów udostępnianych w Internecie, w tym także samych domen, udostępnianie informacji o użytkownikach, które są łatwe do odnalezienia w Internecie z użyciem technik OSINT-owych oraz udostępnianie danych o działaniu firmy, w tym także informacji z klauzulą poufności. Ludzie mogą udostępniać zbyt dużo informacji o działalności i sposobie pracy w firmie, danych geolokalizacyjnych oraz informacji osobistych.

Dodatkowym zagrożeniem dla infrastruktury rozpoznawanej może być częste wykorzystywanie systemu Kali Linux jako platformy do działań OSINT-owych, co zdaniem jednego z ekspertów nie jest dobrą drogą. Duży zasób różnego rodzaju narzędzi, które są preinstalowane w tej dystrybucji Linuksa może wyrządzić wiele szkód, jeśli jest obsługiwany przez osoby bez odpowiedniego doświadczenia w ich stosowaniu. Poprzez skanowanie całej podsieci, agresywnego skanowania wszystkich 65536 portów lub masowego odpytywania serwerów DNS z celu odnalezienia subdomen, należących do badanej organizacji, możliwe jest zablokowanie działania serwerów tak, jak ma to miejsce w przypadku ataku DoS[167].

[167] Atak typu DoS (skrót od ang. *Denial of Service*) – atak sieciowy, polegający na obciążeniu infrastruktury w taki sposób, aby nie mogła ona wykonywać swoich standardowych operacji. Może być przeprowadzony np. poprzez wysyłanie ogromnych ilości zapytań do serwera, co spowoduje bardzo długi czas oczekiwania na ich obsługę lub nawet awarię, która całkowicie uniemożliwi korzystanie z danej usługi.

W odniesieniu do pytania nr 8: ***jakie zagrożenia dla bezpieczeństwa osobowego widzi Pani/Pan w kontekście możliwości rozpoznania działań osoby w Internecie poprzez techniki i narzędzia OSINT-owe?***

Wśród wielu zagrożeń, przywoływanych przez ekspertów, najczęściej wymienianym był *doxing*[168]. Termin ten, wywodzący się od angielskiego określenia „dropping documents", czyli udostępniania (lit. wrzucania) dokumentów w Internecie w celu wyrządzenia szkody drugiej osobie, poprzez ujawnienie zawartych w nich danych osobistych, jak np.: nazwisko, adres, telefon, miejsce pracy lub dane finansowe. Określenie *„doxing"* pojawiło się już w latach 90. XX wieku w środowiskach hakerskich i odnosi się do skrótowego określenia dokumentów w języku angielskim – „docs". W tym samym okresie miały miejsce jedne z pierwszych udokumentowanych przypadków tego typu działań w USA, kiedy 8 lekarzy, powiązanych z klinikami aborcyjnymi, zostało zamordowanych po tym jak aktywiści zdobyli i opublikowali ich dane osobowe, w tym adresy i zdjęcia, na liście osób do zlikwidowania[169]. W 2017 roku, badacze z New York University Tandon School of Engineering (NYU) oraz University of Illinois w Chicago (UIC), opublikowali wyniki badań[170] prowadzonych po raz pierwszy na dużą skalę w zakresie możliwości wyrządzenia szkody poprzez używanie *doxingu*. Wynika z nich, że ofiary tego typu działań są bardziej skłonne do zmiany swoich ustawień prywatności w serwisach internetowych dopiero po tym, kiedy już padną ofiarą ataku, a najczęstszymi powodami *doxingu* są chęć zemsty i wyrównania rachunków, podczas gdy na przykład współzawodnictwo i motywy polityczne odpowiadają jedynie za mniej niż 1% powodów. Spośród przeanalizowanych ponad 5500 plików związanych z *doxingiem*, ponad 90% zawierało adres figuranta, 61% zawierało adres domowy, a 53% adres e-mail. Spośród informacji finansowych, zaledwie 4,3% plików posiadało informacje o numerze karty kredytowej, 2,6% numer ubezpieczenia społecznego[171], a 8,8% inne informacje o charakterze finansowym.

[168] Spotykana jest też pisownia „doxxing".
[169] *How Abortion Providers Are 'Living in the Crosshairs'*, https://www.rollingstone.com/politics/politics-news/how-abortion-providers-are-living-in-the-crosshairs-34307/ – dostęp online 08.07.2022 r.
[170] *Why They Dox: First Large-Scale Study Reveals Top Motivations and Targets For this Form of Cyber Bullying*, https://engineering.nyu.edu/news/why-they-dox-first-large-scale-study-reveals-top-motivations-and-targets-form-cyber-bullying – dostęp online 08.07.2022 r.
[171] Numer ubezpieczenia społecznego w USA jest jednym z podstawowych osobistych numerów identyfikacyjnych. W Polsce jego odpowiednikiem (pod względem ważności) jest numer PESEL.

Wyniki innego badania, które w 2019 roku przeprowadzono na 2120 uczniach szkół średnich w Hong Kongu[172], pokazały, że 12% z nich przyznało się do *doxingu*, przy czym dwiema najpopularniejszymi typami działań, były działania społecznościowe (głównie w zakresie informacji o dacie urodzenia, szkole, statusie związku czy zdjęć i filmów) oraz wrogie (w zakresie informacji osobistych, adresów, numerów kart kredytowych czy stanu posiadania). Oba te badania ukazały skalę i możliwości pozyskania informacji ze źródeł internetowych, a także niebezpieczeństwa, które mogą prowadzić do utraty pracy czy nawet życia przez atakowaną osobę. Najnowsze wyniki badań pokazują[173], że do częstych przypadków należą oskarżenia dotyczące wrażliwych tematów oraz udostępnianie zdjęć, których ofiary wolałyby nie rozpowszechniać.

Wśród innych wskazanych zagrożeń, związanych z wykorzystaniem narzędzi i technik OSINT-owych przeciwko konkretnym osobom, mającym skutki zarówno w Internecie, jak i poza nim. Można tutaj wyróżnić: kradzież tożsamości, uporczywe nękanie czy nawet zjawisko *swattingu* (czyli bezpodstawnego wzywania policji lub służ ratunkowych pod adres nękanej osoby), spotykanego głównie w USA i którego rezultatem były również przypadki śmiertelne[174].

Zagrożenia dla bezpieczeństwa osobowego nie wynikają zazwyczaj z niebezpieczeństwa jakie niesie zostawianie drobnych fragmentów informacji osobistych przez praktycznie każdą osobę korzystającą z Internetu, lecz dopiero poprzez złożenie tych informacji w szerszy obraz czyjegoś życia, nawyków, otoczenia itp. Możliwe jest to dzięki wykorzystaniu odpowiednich technik, procesów i w wielu przypadkach łatwych w obsłudze narzędzi. Dlatego też ważne jest utrzymywanie odpowiedniego poziomu higieny cybernetycznej podczas pracy w Internecie, zadbanie o ustawienia prywatności, brak nadmiarowego udostępniania informacji i używanie umiejętności OSINT-owych w sposób proaktywny, w celu ochrony przed atakującymi, próbującymi korzystać z przedmiotowych technik.

Podobnie, jak to miało miejsce w opiniach wyrażonych w odpowiedzi na pytanie 7, należy zadbać o odpowiedni poziom OPSEC w przypadku prowadzenia śledztw

[172] M. Chen, A. Shann Yue Cheung, K. Ling Chan, *Doxing: What Adolescents Look for and Their Intentions*, International journal of environmental research and public health vol. 16,2 218. 14 Jan. 2019.
[173] *Doxxing in 2022: An Unexpectedly Widespread Cybersecurity Threat*, https://www.safehome.org/family-safety/doxxing-online-harassment-research/ – dostęp online 08.07.2022 r.
[174] *Fatal 'Swatting' Episode in Kansas Raises Quandary: Who Is to Blame?*, https://www.nytimes.com/2017/12/31/us/wichita-swatting-barriss.html – dostęp online 08.07.2022 r.

OSINT-owych, w tym także świadomość zagrożeń płynących z wykorzystywanych w tym procesie narzędzi przed ich pierwszym użyciem (o czym wspomniane było także opiniach dotyczących pytania nr 4).

W odniesieniu do pytania nr 9: ***czy Pani/Pana zdaniem narzędzia i techniki OSINT-owe można wykorzystać do działań defensywnych, zarówno w kontekście weryfikacji bezpieczeństwa infrastruktury teleinformatycznej, jak i bezpieczeństwa osobowego oraz czy istnieje równowaga pomiędzy możliwościami ofensywnymi i defensywnymi w przedmiotowym zakresie?***

Wśród ekspertów można wyróżnić dwa stanowiska w przedmiotowym zakresie: określające OSINT jako sposób rozpoznania jedynie sił przeciwnika oraz takie, które dają możliwość wykorzystania go także w działaniach defensywnych, związanych głównie z działaniem zespołów Red Team, Blue Team oraz Purple Team.

Badacze bezpieczeństwa, specjaliści w zakresie socjotechniki, inżynierowie badający zagrożenia sieciowe i inni osoby, na co dzień zajmujące się profesjonalnie działaniami związanymi z cyberbezpieczeństwem, wykorzystują techniki OSINT-owe w celu proaktywnego identyfikowania zagrożeń i odpowiedniego zarządzania ryzykiem, gdyż te same sposoby, które z jednej strony mogą być wykorzystane do ataku, z drugiej strony mogą być sposobem ochrony infrastruktury. Jedyną różnicą są cele misji, która stoi za ich wykorzystaniem.

W ramach technik OSINT-owych, wykorzystywanych m.in. przez analityków bezpieczeństwa do zbierania informacji i wypracowywania sposobów obrony własnej organizacji (a w tym jej infrastruktury), można wskazać monitorowanie nowych podatności, wykradzionych danych dostępowych do kont firmowych, a także weryfikację informacji udostępnianych w mediach społecznościowych. Tego typu dane znajdują się w raportach takich firm jak Mandiant czy Kaspersky, które monitorują działania określonych grup w Internecie. W czasach, kiedy każdy zostawia drobne fragmenty informacji na forach internetowych, w mapach Google, poprzez udostępnienie krótkiego filmiku na TikToku czy nawet informacji o podróży, którą dana osoba odbyła dawno temu, tego typu dane mogą być wykorzystane do zasilenia procesów śledztw OSINT-owych.

W odniesieniu do pytania nr 10: ***w kontekście zasad wynikających z norm i publikacji międzynarodowych, dotyczących bezpieczeństwa teleinformatycznego (jak np. rodzina norm ISO 27000 lub podobne publikacje), a także wytycznych bezpiecznego prowadzenia operacji (OPSEC) i bezpieczeństwa personelu (PERSEC), jak ocenia Pani/Pan ich wpływ na możliwości w zakresie przeciwdziałania wykorzystaniu technik OSINT-owych?***

W zakresie przestrzegania zasad OPSEC oraz PERSEC jest zdaniem ekspertów niezbędne do zachowania odpowiedniego bezpieczeństwa prowadzonych działań. Korzystanie z takich standardów, jak chociażby ISO 27000 lub mu pokrewne, jest pewnego rodzaju podstawą dla bezpieczeństwa, jednak w związku z faktem, że każda organizacja jest inna, inne są także modele zagrożeń i potrzeby w zakresie ich obsługi. Wszystkie organizacje muszą zaplanować jakie prace muszą wykonać, aby chronić swoje zasoby (w tym te, należące do klientów), poufne informacje oraz całe systemy przed nieuprawnionym dostępem najlepiej, jak tylko mogą. Zakres będzie się oczywiście różnił w zależności od rodzaju organizacji, jej potrzeb, budżetu oraz umiejętności pracowników (zarówno własnych, jak i zakontraktowanych z zewnątrz).

Zachowanie odpowiednich poziomów OPSEC i PERSEC jest konieczne tak samo, jak jednoczesne wykorzystywanie technik OSINT-owych do celów defensywnych (o czym wspomniane było także w opiniach wyrażonych w odpowiedzi na pytanie nr 9), ponieważ zawsze powinno się posiadać możliwości obrony przed wykorzystaniem informacji, pochodzących z rozpoznania i analiz OSINT-owych. Jednocześnie standaryzacja mechanizmów bezpieczeństwa powinna też mieć swoje granice, ponieważ w świecie dynamicznie zmieniającego się środowiska internetowego, zasady te także powinny pozostać elastyczne.

OPSEC i PERSEC mogą pomóc w uchronieniu informacji poufnych przed dostaniem się w niepowołane ręce, jednak sama wiedza o tych zasadach nie wystarczy – muszą one być zintegrowane z procesami, stać się częścią codziennej pracy i elementem zadań każdego pracownika. Dobrą praktyką jest także cykliczna weryfikacja czy przyjęte zasady cały czas działają na odpowiednim poziomie i czy informacje są cały czas bezpieczne.

Ogólnie rzecz ujmując, to nie zasady, wynikające z rodziny norm ISO 27000 czy podobnych zasad, zapewniają bezpieczeństwo, ale tak samo jak ma to miejsce w przypadku zastosowania technik OSINT-owych – wszystko zależy od tego *jak* te

zasady są stosowane, gdyż samo posiadanie certyfikacji nie gwarantuje bezpieczeństwa, jeśli nie idzie za tym zapewnienie ich odpowiedniego funkcjonowania.

Uwagi końcowe

Tylko jeden z ekspertów wypowiedział się w ramach tego punktu kwestionariusza, wskazując na duży zakres analiz środowiska informacyjnego i podmiotów w nim działających, prowadzonych na własny użytek zarówno przez siły zbrojne, jak i instytucje cywilne. W ramach tego typu działań wskazane zostały przykładowo obszary, takie jak: kampanie PR, marketing oraz media, co wskazuje głównie na obszary styku biznesu ze środowiskami konsumenckimi i stanowi podstawę do planowania polityki firmowej wpisującej się w oczekiwania i nastroje rynku.

Wśród wskazanych obszarów znalazło się także zarządzanie zasobami (HR). Tego typu analizy mogą mieć związek nie tylko z możliwymi działaniami w zakresie badania poziomu absencji własnej załogi, efektywności czy satysfakcji z pracy, ale także weryfikacją powiązań między wskaźnikami efektywności w innych firmach lub nawet analizą konkurencji, w zakresie aktywności w Internecie, poziomu płac lub akcji promocyjnych.

Wskazane tutaj zostało także badanie obszaru CIMIC (skrót od ang. *Civil-Military Co-operation*), które stanowią ważny element przygotowania i prowadzenia operacji wojskowych z uwzględnieniem współpracy cywilno-wojskowej. Tego typu działania mają za zadanie stworzyć odpowiednie warunki dla działań militarnych, a także wsparcie środowiska cywilnego dla prowadzonych działań, a odpowiednie rozpoznanie możliwości i sposobów współpracy może okazać się kluczowe dla bezpieczeństwa i powodzenia operacji wojskowych na danym terenie.

IV ŹRÓDŁA INTERNETOWE

Poniżej przedstawiono listę źródeł internetowych, do których odniesienia zawarte są w opracowaniu wywiadu eksperckiego:

Be careful what you OSINT with, https://keyfindings.blog/2020/03/23/be-careful-what-you-osint-with – dostęp online 30.06.2022 r.

DISARM Disinformation TTP (Tactics, Techniques and Procedures) Framework, https://github.com/DISARMFoundation/DISARMframeworks/ – dostęp online 04.07.2022 r.

Doxxing in 2022: An Unexpectedly Widespread Cybersecurity Threat, https://www.safehome.org/family-safety/doxxing-online-harassment-research/ – dostęp online 08.07.2022 r.

Fatal 'Swatting' Episode in Kansas Raises Quandary: Who Is to Blame?, https://www.nytimes.com/2017/12/31/us/wichita-swatting-barriss.html – dostęp online 08.07.2022 r.

How Abortion Providers Are 'Living in the Crosshairs', https://www.rollingstone.com/politics/politics-news/how-abortion-providers-are-living-in-the-crosshairs-34307/ – dostęp online 08.07.2022 r.

M. Chen, A. Shann Yue Cheung, K. Ling Chan, Doxing: *What Adolescents Look for and Their Intentions*, International journal of environmental research and public health vol. 16,2 218. 14 Jan. 2019.

Why They Dox: First Large-Scale Study Reveals Top Motivations and Targets For this Form of Cyber Bullying, https://engineering.nyu.edu/news/why-they-dox-first-large-scale-study-reveals-top-motivations-and-targets-form-cyber-bullying – dostęp online 08.07.2022 r.