Tomasz Kłys

WYKORZYSTANIE TECHNIK BIAŁEGO WYWIADU W PROCESIE WYKRYWCZYM

Rozwój technologiczny przyczynił się do istotnych zmian w codziennym funkcjonowaniu społeczeństwa. Coraz szersze wykorzystanie Internetu oraz urządzeń mobilnych do komunikacji, stały się początkiem problemów związanych z bezpieczeństwem użytkowników. Z drugiej strony mamy do czynienia z zupełnie nowymi metodami pozyskiwania danych z otwartych źródeł informacji. To właśnie informacje stanowią jeden z elementów zainteresowania służb odpowiadających za ochronę bezpieczeństwa i porządku publicznego, a także agencji wywiadowczych. Analizując zagadnienie związane z pozyskiwaniem danych w Internecie, należy zwrócić uwagę na pojęcie białego wywiadu (ang. OSINT – open source intelligence). Wykorzystanie białego wywiadu do celów wykrywczych jest dla organów ścigania nowym wyzwaniem, ale właściwie przeprowadzona praca wywiadowcza może przynieść bardzo dobre rezultaty.

Czym jest biały wywiad?

Polskie ustawodawstwo nie reguluje pojęcia białego wywiadu. W literaturze spotkać można jednak wiele definicji, które definiują czym jest wykorzystanie otwartych źródeł informacji. Tego rodzaju informacje mogą więc zostać określone jako ciąg danych, które pochodzą z jednego lub więcej jawnych źródeł, podlegające procesowi ewaluacji z uwzględnieniem czasu publikacji informacji oraz ich zawartości[1]. OSINT uznawany jest także za wynik przeprowadzenia pewnych czynności w stosunku do informacji, które są specjalnie poszukiwane, porównywane ze sobą co do treści, celem wybrania najważniejszych dla odbiorcy procesu.

Poza wyjaśnieniem znaczenia terminu otwarte dane, w literaturze zagranicznej wytycza się także cztery zasadnicze etapy ich analizowania.
Zaliczają się do nich:

1) zebranie danych surowych (ang. Open Source Data, OSD), pochodzących z pierwotnego źródła (tj. publikacji drukowanych, mediów, stron internetowych, fotografii) i przedstawienie ich w najprostszej formie,
2) poddanie analizie zebranych danych (ang. Open Source Information, OSINF), poddanie ich pewnym zabiegom edytorskim, zebranie ich w jeden dokument i przekazanie osobom zarządzającym w celu dalszego rozpowszechnienia,
3) zaplanowane uzyskanie informacji, tzw. biały wywiad (ang. Source Intelligence, SINT) – przekazanie danych wyselekcjonowanej grupie odbiorców i zgodnie z zasadami określonymi przez składającego zapytanie (każda służba policyjna czy wywia-

[1] K. Mroziewicz, Czas pluskiew, Warszawa 2007, s. 334.

dowcza wypracowuje we własnym zakresie metodykę postępowania z danymi i informacjami),

4) weryfikacja informacji, tzw. zweryfikowany, potwierdzony biały wywiad (ang. Validated Open Source Inteligence, OSINT-V) – potwierdzenie stopnia poziomu pewności informacji na podstawie różnych źródeł[2].

Źródła pozyskiwania informacji

Społeczeństwo informacyjne pojęcie poszukiwania informacji kojarzy głównie z wykorzystaniem podstawowej funkcjonalności Internetu, czyli wyszukiwarek www. Użytkownik może za ich pomocą znaleźć wszelkiego rodzaju informacje na temat konkretnej instytucji lub osoby. Chodzi między innymi o profile na serwisach społecznościowych lub aukcyjnych. Danych poszukiwać można na forach, blogach, serwisach prowadzonych przez przedsiębiorstwa, rejestrach domen WHOIS, mapach, zdjęciach satelitarnych lub lotniczych. Adresy e-mail pozwalają na wyszukiwanie powiązań pomiędzy osobami, przedsiębiorstwami, domenami oraz adresami IP. Coraz częściej z Internetu korzysta się przy pomocy urządzeń mobilnych, takich jak smartfony. Do komunikacji mogą one wykorzystywać technologię Bluetooth, natomiast do określania przybliżonej lokalizacji stosowana jest antena GPS. W trakcie działań wywiadowczych mogą one dostarczać ogromnej ilości aktualnych informacji.

Biały wywiad nie opiera się jednak w całości na wykorzystaniu zasobów Internetu.

Otwarte źródła informacji najczęściej wykorzystywane w pracy organów ścigania można podzielić na kilka kategorii:

1) media tradycyjne:
   – prasa drukowana (np. dzienniki, czasopisma branżowe, dokumenty rządowe),
   – telewizja informacyjna, rozgłośnie radiowe,
   – literatura (książki, publicystyka, analizy, śledztwa dziennikarskie),
2) usługi komercyjne:
   – podmioty gospodarcze, które za opłatą przygotowują sprofilowane raporty i analizy,
   – wydawnictwa marketingowe,
3) szara literatura (ang. grey literature) – analizy, informacje będące do dyspozycji tylko poprzez wyspecjalizowane kanały, generowane przez środowiska akademickie, organizacje państwowe i pozarządowe,
4) bazy danych i katalogi[3].

---

[2] B. Sromczyńki, P. Waszkiewicz, *Biały wywiad w praktyce pracy organów ścigania na przykładzie wykorzystania serwisów społecznościowych*, „Prokuratura i Prawo" 2014, nr 5, s. 149

[3] G. Dobrowolski, W. Filipkowski, M. Kisiel-Dorohnicki, W. Rakoczy, *Wsparcie informatyczne dla analizy otwartych źródeł informacji w Internecie w walce z terroryzmem. Zarys problemu*, [w:] *Praktyczne elementy zwalczania przestępczości zorganizowanej i terroryzmu*, pod red. L. K. Paprzyckiego, Z. Rau, Warszawa 2009, s. 281–282.

Strategia prowadzenia białego wywiadu

- Określić cel poszukiwania;
- Wybrać właściwe słowa kluczowe;
- Określić geograficzny zasięg wyszukiwania (z perspektywy tematu i źródeł informacji);
- Określić czasowy zakres wyszukiwania;
- Określić źródła informacji;
- Zgromadzić informacje;
- Przeprowadzić selekcję oraz analizę.

Obserwując praktykę śledczą służb specjalnych, można zaryzykować stwierdzenie, że wykorzystywanie otwartych źródeł informacji, przede wszystkim zasobów Internetu, jest nieodłącznym elementem działań wykrywczych. Niestety, faktem jest, że działanie takie nie jest obce również zorganizowanym grupom przestępczym oraz ugrupowaniom terrorystycznym, które adaptują techniki charakterystyczne dla białego wywiadu i używają ich w swojej działalności[4].

Analiza mediów społecznościowych

Media społecznościowe na całym świecie cieszą się ogromną popularnością. Najpopularniejszą platformą społecznościową od wielu lat jest Facebook. Liczba umieszczanych na Facebooku treści jest tak duża, że ich analiza często wymagać będzie zastosowania specjalnych narzędzi i technik ich przetwarzania. Obecnie sporym zainteresowaniem cieszy się również Instagram, X (dawniej Twitter), LinkedIn. Od 2012 roku stosowane jest pojęcie SOCMINT (ang. Social Media Intelligence), czyli wykorzystywanie mediów społecznościowych do pozyskiwania informacji w ramach białego wywiadu. W czasach, gdzie 1/4 ludzkości komunikuje się za pomocą mediów społecznościowych, należy poznać mechanizmy ich funkcjonowania, bowiem to właśnie od tego elementu zależeć będzie sukces związany z pozyskaniem odpowiednich informacji[5].

W celu pozyskania informacji z mediów społecznościowych można użyć jednej z poniżej opisanych technik:

1) szukanie bezpośrednio w mediach społecznościowych za pomocą wbudowanej wewnętrznej wyszukiwarki oraz opcji wyszukiwania zaawansowanego,
2) szukanie z wykorzystaniem zewnętrznych narzędzi, które poprzez wpięcie API do mediów społecznościowych pozwalają na pobieranie danych i ich ustrukturyzowanie, lub narzędzi, których zadaniem jest stworzenie zaawansowanej kwerendy wyszukiwawczej do Google,

---

[4] K. Jarczewska-Walendziak, *Wykorzystanie otwartych źródeł przez służby śledcze*, Toruńskie Studia Biologiczne 2017 nr 1, s. 139

[5] P. Hrabiec-Hojda, J. Trzeciakowska, *Techniki wyszukiwani informacji w mediach społecznościowych dla celów białego wywiadu,* Studia Politologiczne vol. 54, 2019, s. 179

3) szukanie w wyszukiwarce z wykorzystaniem zaawansowanych operatorów i techniki Boolean string[6].

Należy mieć na uwadze fakt, że informacje podawane za pośrednictwem serwisów społecznościowych mają charakter deklaratywny i wcale nie muszą znajdować odzwierciedlenia w rzeczywistości. Pewnym ograniczeniem są również ustawienia prywatności lub publikacja treści w zamkniętych grupach.

W mediach społecznościowych każdego dnia pojawia się ogromna liczba informacji, dlatego do ich analizy wykorzystywane jest specjalistyczne oprogramowanie i aplikacje. Za ich pomocą możliwe staje się monitorowanie oraz agregowanie danych ze wskazanych platform[7].

Media społecznościowe są niezwykle popularnym kanałem informacji. Liczba informacji umieszczanych za ich pośrednictwem, w większości pozostaje poza zasięgiem wyszukiwarek. Umiejętność pozyskiwania danych z mediów społecznościowych oraz znajomość oprogramowania wspierającego pobieranie i analizę takich materiałów, stanowi niezwykle ważną umiejętność podczas realizacji działań wywiadowczych.

Wykorzystanie wyszukiwarek internetowych

Wyszukiwarka stanowi narzędzie internetowe, które umożliwia użytkownikom znajdowanie informacji w sieci. Jest to program, który wyszukuje i identyfikuje pozycję w bazie danych, które odpowiadają słowom kluczowym lub znakom określonym przez użytkownika. Najczęściej stosowana jest do wyszukiwania określonych witryn w sieci.

Wyszukiwarki działają poprzez przeszukiwanie setek miliardów stron przy użyciu własnych robotów indeksujących. Te roboty indeksujące są powszechnie nazywane botami wyszukiwarek lub pająkami. Wyszukiwarka porusza się po Internecie, pobierając witryny internetowe i podążając za linkami, które tam napotkają, aby odkryć nowe, powiązane z nimi strony.

Użytkownik wprowadza słowa kluczowe lub frazy kluczowe do wyszukiwarki i otrzymuje listę wyników treści, które zdaniem wyszukiwarki najlepiej pasują do wprowadzonego zapytania. Są wyświetlane w postaci listy stron internetowych, obrazów, filmów lub innych danych, dostępnych online.

Google Hacking jest to technika pozyskiwania danych wrażliwych osób i instytucji z wykorzystaniem do tego celu odpowiednio sformułowanych zapytań skierowanych do wyszukiwarki Google. Pozwala na uzyskanie informacji o parametrach konfiguracyjnych serwerów, w tym serwerów www oraz innych urządzeń sieciowych, wyszukiwanie stron oraz dokumentów bezpośrednio niedostępnych, odtwarzanie struktury witryn internetowych lub struktury sieci wewnętrznej, pozyskiwanie informacji celowo zabezpieczonych przed dostę-

[6] P. Hrabiec-Hojda, J. Trzeciakowska, *Techniki wyszukiwani informacji w mediach społecznościowych dla celów białego wywiadu,* Studia Politologiczne vol. 54, 2019, s. 180

[7] Tamże s. 186

pem (ang. *paywall*), a także pozyskanie nazw użytkowników i haseł oraz innych oznaczeń identyfikujących użytkowników lub informacji o nich[8].

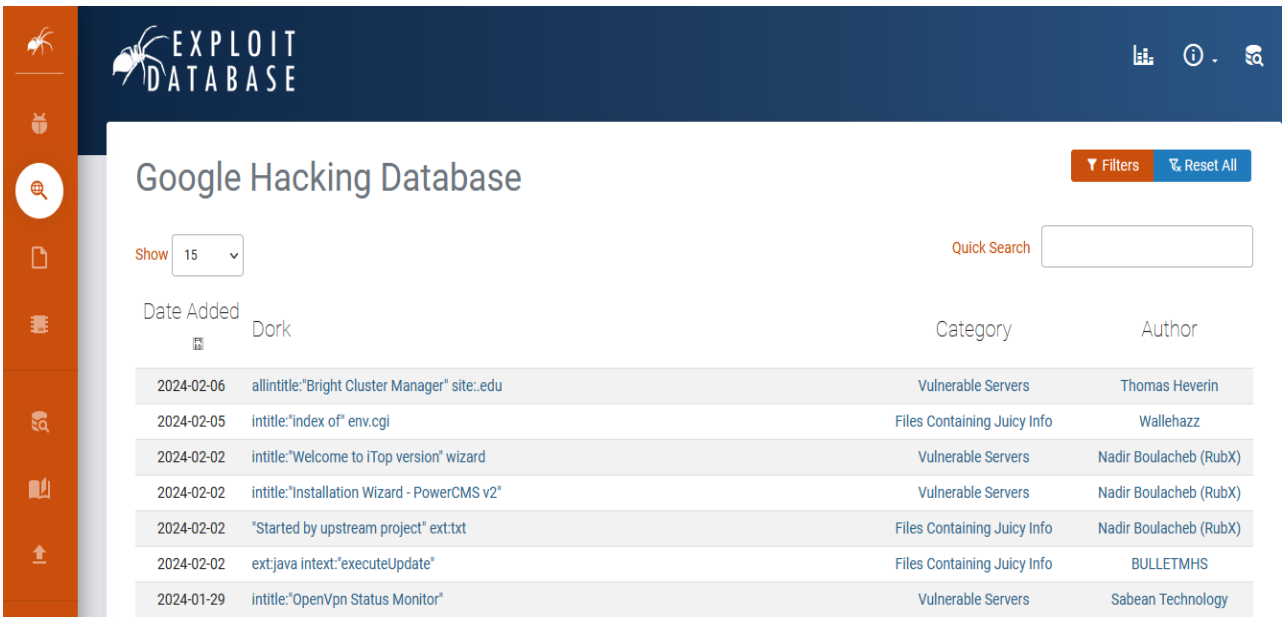
Przykładowe operatory i znaki specjalne wyszukiwania

- **Cudzysłów „”** – wywołuje on konkretną poszukiwaną przez nas frazę. Wpisując: „biały wywiad”, otrzymamy wyniki dotyczące tej właśnie frazy. Jeżeli nie użyjemy apostrofów, Google może każdego słowa szukać osobno lub też je pomijać. Z cudzysłowami mamy pewność, że będzie szukać dokładnie wymaganej przez nas frazy.
- **Gwiazdka *** – używając gwiazdki w danym zapytaniu, maskujemy elementy, których nie-znamy. Jest to symbol wieloznaczny. Gwiazdka musi znajdować się na końcu wyszukiwanego hasła. Gdy wpiszemy np. biał* – Google jako pierwszy wynik podpowie nam Białystok. Można użyć gwiazdki do poszukiwania adresu e-mail, żeby znaleźć wszystkie wiadomości wysłane do lub z domeny zewnętrznej. Co ważne, jeżeli nasz symbol wygeneruje ponad 100 słów wyniku, nasze działanie zostanie zakończone niepowodzeniem.
- **site:** – wyszukujemy informacje znajdujące się na konkretnej interesującej nas stronie internetowej. Wpisujemy więc w pierwszej kolejności np. site:onet.pl – i dalej frazę, którą będziemy szukać.
- **inurl:** – dzięki temu zapytaniu będziemy wyszukiwać jedno podane słowo w adresie URL strony, np. inurl:detektyw da nam wyszukiwania stron, które w adresie URL mają słowo detektyw.
- **allinurl**: – jest to wyszukiwanie podobne do poprzedniego, pozwala jednak w adresie URL znaleźć kilka potrzebnych nam słów kluczowych, np. allinurl:biały wywiad,
- **intitle:** – używając go, znajdziemy jeden interesujący nas wyraz w nazwach stron internetowych, a nie tak jak powyżej – w całym długim linku URL, np. intitle:tajemnica,
- **allintitle:** – operator ten umożliwia znalezienie więcej niż jednego interesującego wyrazu w nazwie strony internetowej, np. allintitle:tajemnica detektyw,
- **filetype:** – stosujemy go, jeżeli interesują nas konkretne pliki, np. dokumenty. Szukając wszystkich interesujących nas PDF-ów o wybranym tytule, najpierw wpisujemy filetype:PDF,
- **AND** – szukamy wiadomości, które zawierają dwa podane przez nas hasła, np. tajny AND raport,
- **OR** – szukamy wiadomości, które zawierają obydwa poszukiwane hasła lub przynajmniej jedno z nich, np. tajne OR ściśle tajne,
- **NOT** – znajdź wiadomości, które nie zawierają podanego słowa, np. detektyw NOT Kraków,
- **from:** – znajdź wiadomości wysłane z podanego konta, gdy np. interesują nas wszystkie wiadomości wysłane przez użytkownika danego forum o konkretnym nicku, wpisujemy: from:tutajumieszczamytennick[9].

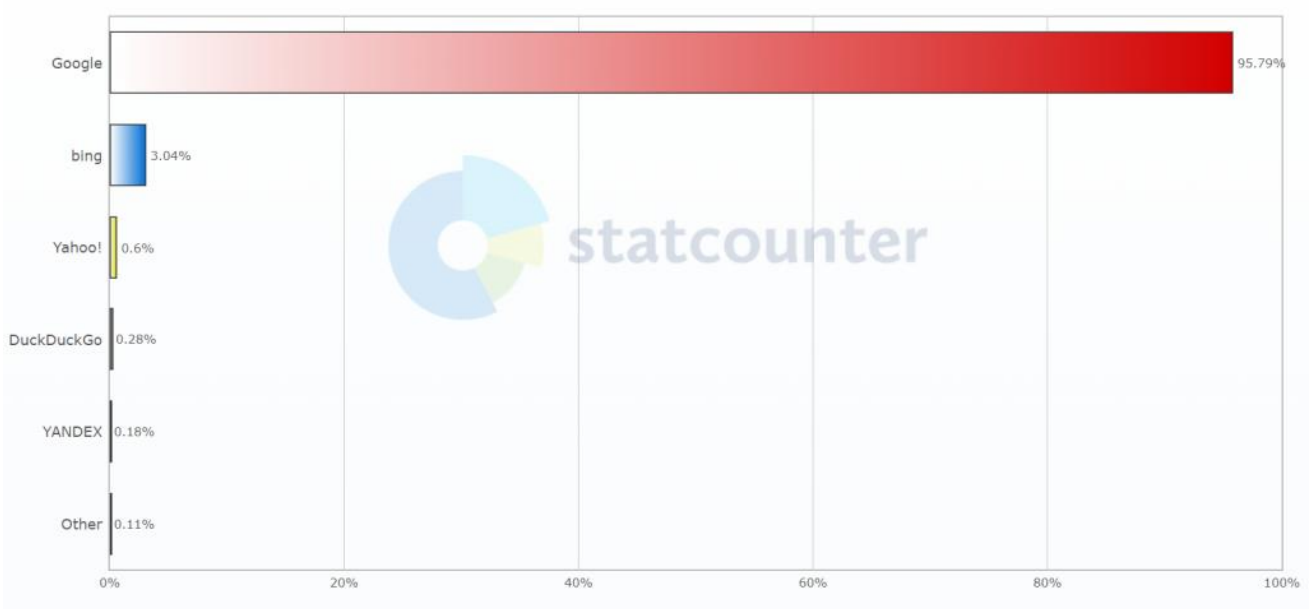
[8] https://cebrf.knf.gov.pl/encyklopedia/hasla/385-definicje/807-google-hacking [dostęp w dniu 06.02.2024r]

[9] https://dkdetektyw.pl/google-hacking/ [dostęp w dniu 06.02.2024 r.]

W Internecie znaleźć można również bazy gotowych zapytań tzw. dorków, które można wykorzystać celem pozyskiwania określonego rodzaju danych.



Należy pamiętać o tym, że poszczególne wyszukiwarki oferują różnorodne mechanizmy pozyskiwania informacji. Ich wyniki mogą być odmienne, pomimo zastosowania tych samych fraz. Największym zainteresowaniem cieszy się jednak Google. W przypadku Polski udział w rynku szacuje się na ponad 95%.

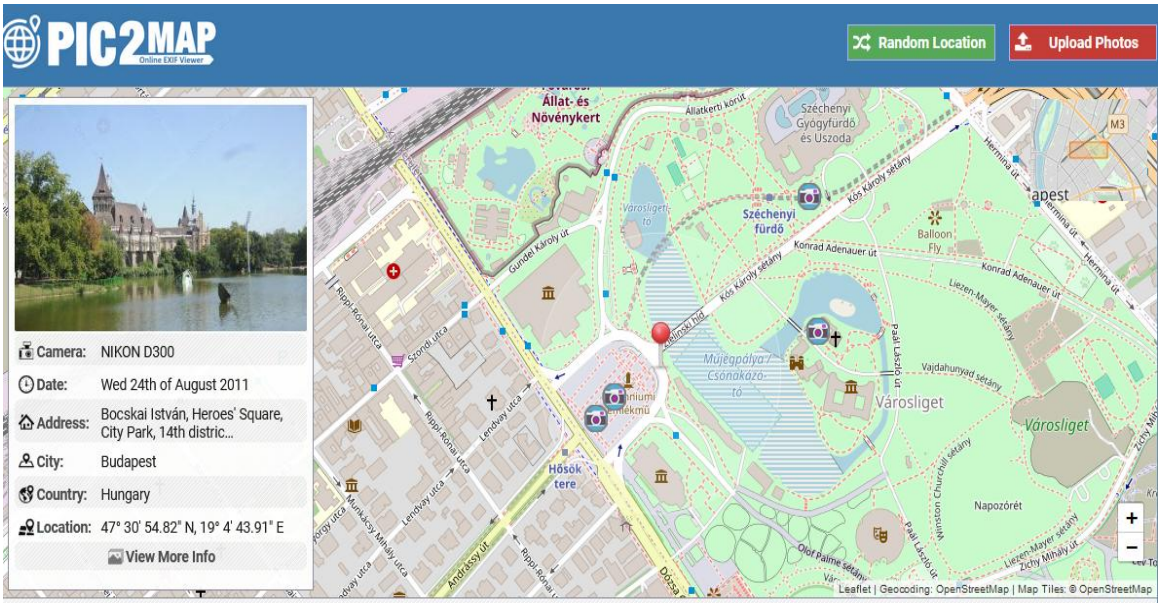


Źródło: https://redseo.pl/blog/ranking-wyszukiwarek-polska-europa-swiat-2023/ [dostęp w dniu 06.02.2024]

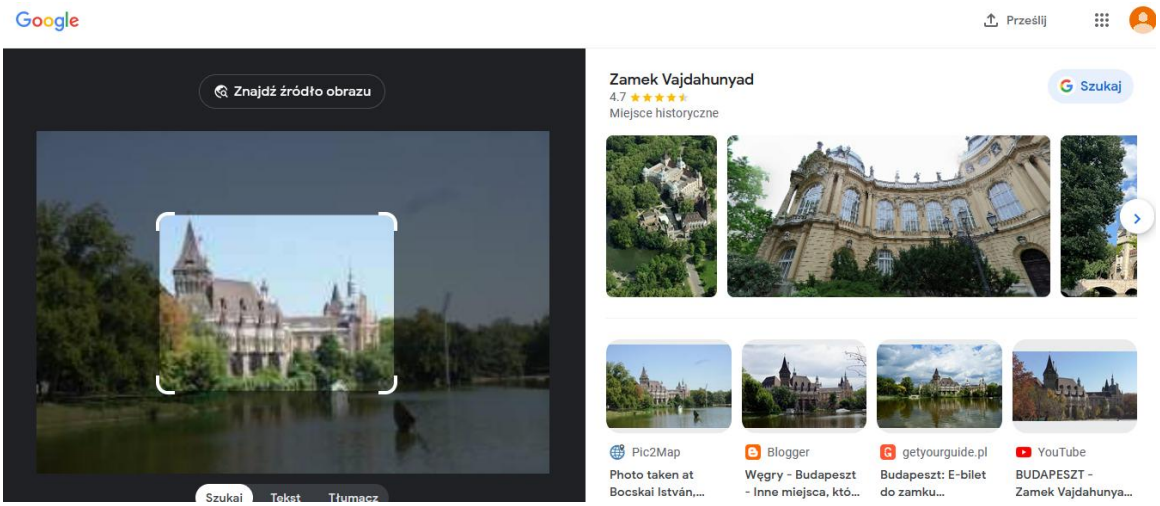
Wyszukiwanie obrazem i metadane

Metadane to ukryte informacje w plikach o tych plikach. Oznacza to, że jeżeli otwieramy plik zdjęciowy, wówczas na ekranie urządzenia widzimy zdjęcie. Natomiast jeśli odczytamy metadane tego pliku, możemy dowiedzieć się, kiedy to zdjęcie zostało wykonane, jakim aparatem, poznać lokalizację oraz odczytać szereg innych informacji.

W celu ustalenia lokalizacji zdjęcia na podstawie metadanych, smartfon lub aparat musi mieć włączone oznaczanie lokalizacji. Niektóre portale społecznościowe lub aplikacje usuwają metadane. Również sami użytkownicy urządzeń w dość łatwy sposób mogą usunąć lub dokonać modyfikacji metadanych. Nie można więc zakładać, że wszystkie zgromadzone takim sposobem informacje będą w pełni odzwierciedlać rzeczywistość. Do uzyskania metadanych wykorzystać można między innymi strony wspomagające wykonanie tego typu czynności.

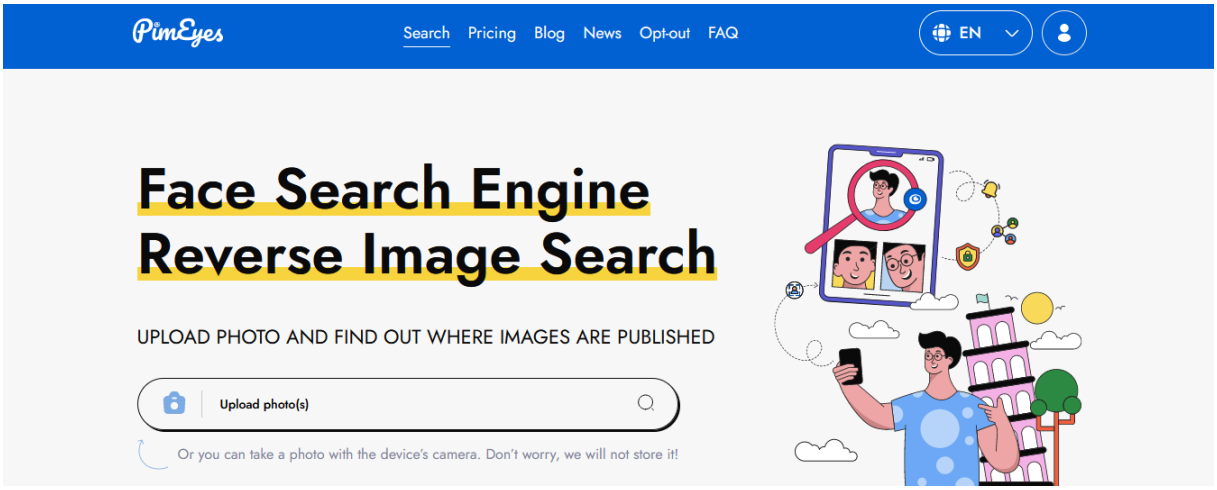


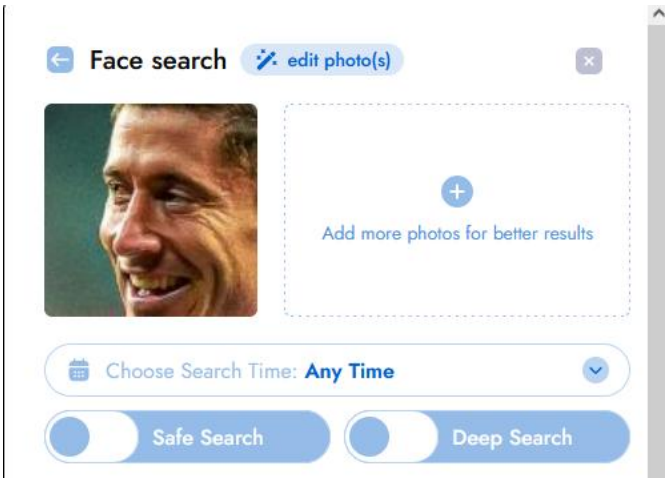
Wykorzystać można również wyszukiwanie obrazem dostępne z poziomu wyszukiwarki internetowej Google.

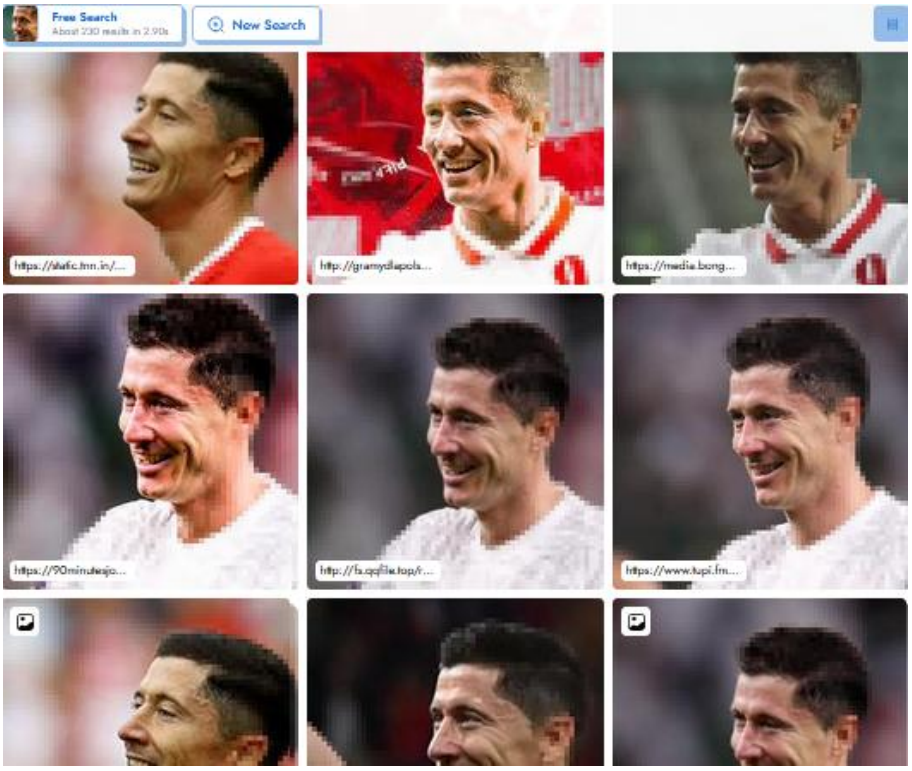


Algorytmy wykrywania twarzy

Za pomocą odpowiednich stron internetowych można również odnaleźć fotografie, na których zachowany został wizerunek konkretnej osoby. PimEyes to wyszukiwarka pozwalająca odnaleźć różne zdjęcia danej osoby na podstawie jednego zdjęcia załadowanego na stronę. Gdy użytkownik prześle do wyszukiwarki PimEyes zdjęcie z wizerunkiem twarzy, wówczas witryna wyświetli obrazy powiązane z innymi miejscami, w których ta osoba się pojawia w sieci – w tym stare filmy, wiadomości, albumy fotograficzne i osobiste blogi. Do przyporządkowania konkretnej twarzy do wizerunku danej osoby PimEyes używa sztucznej inteligencji oraz bazy danych zawierających ponad dwa miliardy zdjęć.





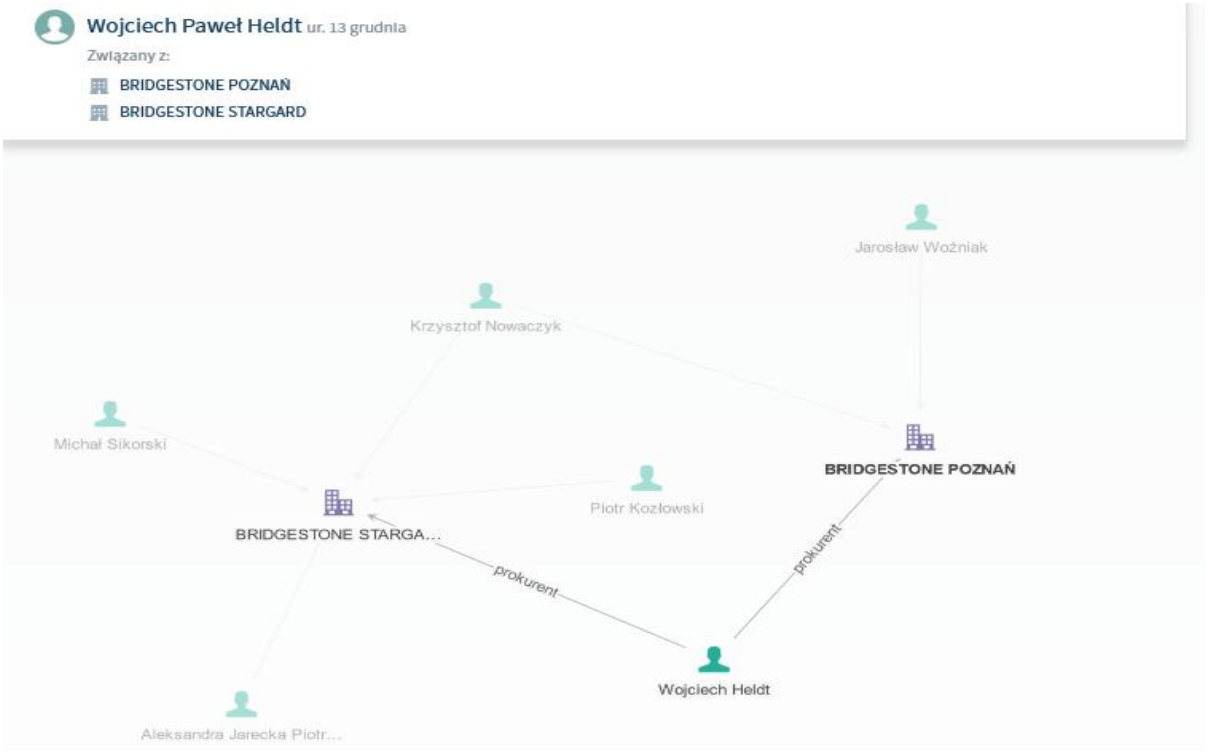


PimEyes potrafi znaleźć zdjęcia każdej osoby, nawet jeśli nosi ona okulary przeciwsłoneczne lub maskę, albo jej twarz jest odwrócona od aparatu, nie uwzględnia jednak wyników pochodzących z serwisów społecznościowych.

Poszukiwanie informacji na temat przedsiębiorstw

W przypadku podejmowania działań wywiadowczych dotyczących przedsiębiorstw, istnieje możliwość szybkiego sprawdzenia danych rejestrowych spółek za pomocą witryny rejestr.io. Analizie poddać można dane dotyczące członków zarządu, wspólników czy kapitał zakładowy. Wystarczy wejść na stronę, wpisać nazwę spółki lub osobę i witryna bardzo szybko pokaże to, czego potrzebujemy. Rejestr.io jest stroną należącą do Fundacji Moje Państwo, która zajmuje się tworzeniem narzędzi do łatwiejszego korzystania z zasobów publicznych.

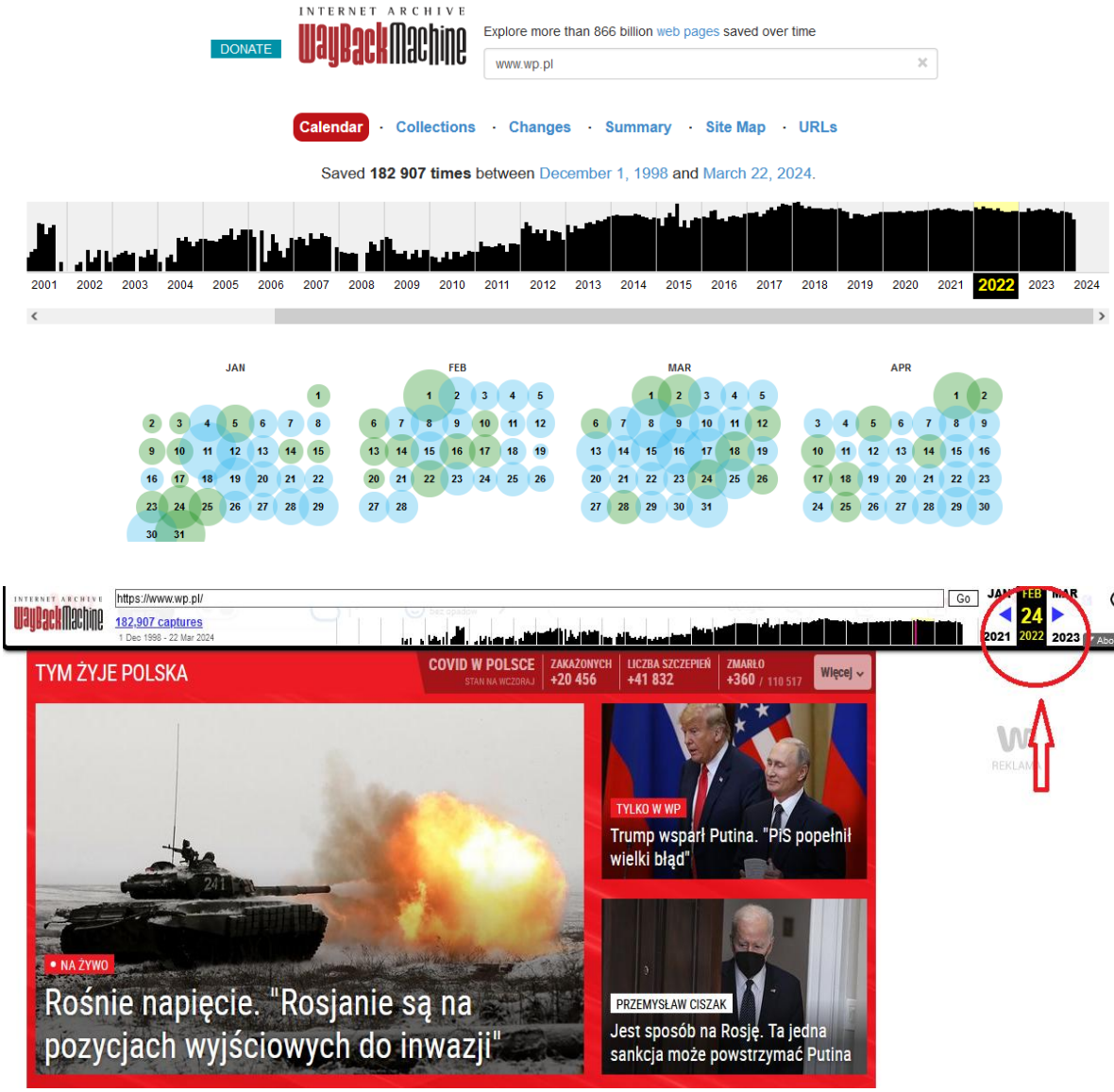
Strona umożliwia przedstawienie powiązań członków organizacji w formie multimedialnego schematu, co może okazać się niezwykle pomocnym narzędziem do przeprowadzania różnego rodzaju analiz. Wystarczy wybrać imię i nazwisko interesującej nas osoby, aby sprawdzić powiązania z innymi organizacjami oraz zajmowane w nich funkcje.



Archiwalne treści w Internecie

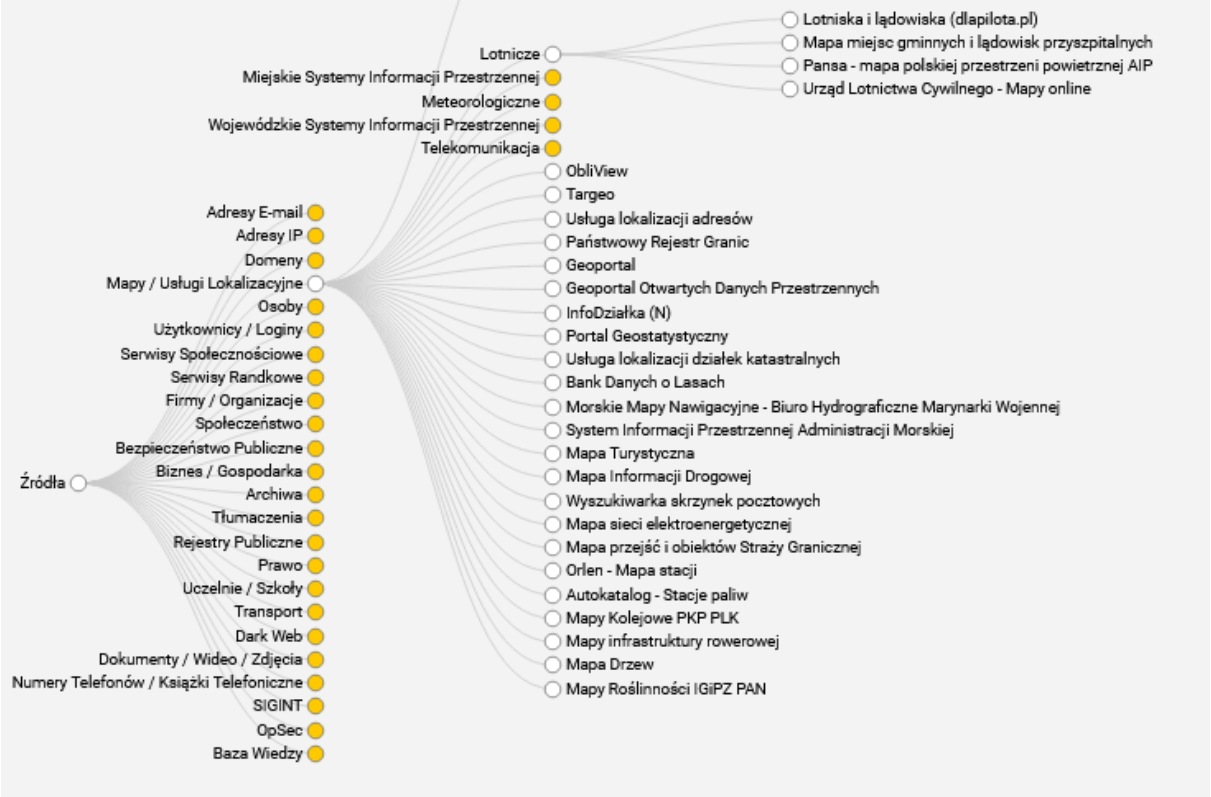
Wayback Machine to narzędzie, które pozwala na przeglądanie zarchiwizowanych wersji stron internetowych. Jest to idealne rozwiązanie dla osób, które chcą przeglądać treści z przeszłości lub dla tych, które potrzebują informacji usuniętych już ze strony internetowej. Hasło „w Internecie nic nie ginie" kojarzy się przede wszystkim z mediami społecznościowymi i udostępnionymi tam treściami i zdjęciami, ale może znaleźć swoje odniesienie także w kontekście stron internetowych.

Web.archive.org umożliwia wprowadzenie nazwy domeny oraz szczegółowego określenia daty, w celu zapoznania się z treścią witryny.



Projekt OSINT Framework

W Internecie znaleźć można wiele narzędzi i witryn wspomagających proces białego wywiadu. Dotarcie do nich ułatwia witryna osintframework.pl, która zawiera dostęp do stron internetowych, które posegregowane zostały na odpowiednie kategorie tematyczne.



Internet jest obecnie kopalnią informacji, które odpowiednio zinterpretowane mogą dostarczyć wiedzy na każdy interesujący poszukującego temat. Ze źródła tego korzystają zarówno podmioty publiczne, jak i pracodawcy czy osoby prywatne. Granice legalnego pozyskiwania danych reguluje prawo krajowe. Ze względu na szybki postęp technologiczny na rynku pojawia się jednakże coraz więcej narzędzi pozwalających na pozyskiwanie w sieci bardziej lub mniej „ukrytych" informacji.

Literatura:

🗁🖋 Dobrowolski G., Filipkowski W., Kisiel-Dorohnicki M., Rakoczy W., *Wsparcie informatyczne dla analizy otwartych źródeł informacji w Internecie w walce z terroryzmem. Zarys problemu*, [w:] *Praktyczne elementy zwalczania przestępczości zorganizowanej i terroryzmu*, pod red. L. K. Paprzyckiego, Z. Rau, Warszawa 2009.

Hrabiec-Hojda P, Trzeciakowska J., *Techniki wyszukiwani informacji w mediach społecznościowych dla celów białego wywiadu,* Studia Politologiczne vol. 54, 2019.

Jarczewska-Walendziak K., *Wykorzystanie otwartych źródeł przez służby śledcze*, Toruńskie Studia Biologiczne 2017 nr 1.

Mroziewicz K., *Czas pluskiew*, Warszawa 2007.

Sromczyńki B., Waszkiewicz P., *Biały wywiad w praktyce pracy organów ścigania na przykładzie wykorzystania serwisów społecznościowych*, „Prokuratura i Prawo" 2014, nr 5.

Źródła internetowe:

1. https://grupainfomax.com/blog/social-media-w-polsce-i-na-swiecie-raport-digital-2023
2. https://cebrf.knf.gov.pl/encyklopedia/hasla/385-definicje/807-google-hacking
3. https://dkdetektyw.pl/google-hacking/
4. https://redseo.pl/blog/ranking-wyszukiwarek-polska-europa-swiat-2023/