

*Patrycja Hrabiec-Hojda*

ORCID: 0000-0001-7893-811X

*Justyna Trzeciakowska*

ORCID: 0000-0003-4917-4493

## Techniki wyszukiwania informacji w mediach społecznościowych dla celów białego wywiadu

SŁOWA KLUCZOWE:

*media społecznościowe, SOCMINT, Social Media Intelligence,  
OSINT, biały wywiad*

STUDIA I ANALIZY

### Wstęp

Media społecznościowe w 2019 roku będą miały ponad 2,77 mld użytkowników na całym świecie, z czego większość będzie korzystała z social mediów tylko przez telefon komórkowy<sup>1</sup>. Najpopularniejszą platformą społecznościową od lat pozostaje Facebook. W Polsce jest to trzecia najpopularniejsza strona, zaraz za Google i YouTube<sup>2</sup>.

Największym wyzwaniem społeczeństwa informacyjnego już teraz jest opanowanie rosnącej wykładniczo liczby informacji oraz znalezienie sposobów na jej efektywne przetwarzanie i analizowanie. Media społecznościowe przyczyniają się do jeszcze szybszego wzrostu informacji i wzmagają informacyjną opresję. Przeciętny użytkownik Facebooka publikuje dziennie nawet 10 postów<sup>3</sup>. Wyobraźmy sobie, jaka liczba informacji

<sup>1</sup> *Number of Social Media Users Worldwide from 2010 to 2021 (in Billions)*, <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/> (dostęp: 26.01.2019).

<sup>2</sup> Według rankingu Alexa z 30.01.2019.

<sup>3</sup> *Digital 2018 Q4 Global Digital Statshot*, <https://www.slideshare.net/DataReportal/digital-2018-q4-global-digital-statshot-october-2018-v2> (dostęp: 27.01.2019).

pojawia się tylko na jednej platformie społecznościowej dziennie, jeśli aż 26% całej ludzkości z niej korzysta. Taki przyrost informacji wymaga coraz lepszych narzędzi i technik ich przetwarzania. Szczególnie, że jak pokażemy w kolejnej części artykułu, większość publikowanych treści pozostaje poza zasięgiem głównego narzędzia do pozyskiwania informacji, czyli wyszukiwarki.

Poniższy artykuł traktować będzie o metodach wyszukiwania i pozyskiwania informacji w źródłach, jakimi są portale społecznościowe. Proponowane techniki wyszukiwawcze wynikają z wieloletniej praktyki autorek w obszarze infobrokeringu oraz białego wywiadu na potrzeby biznesu. Opiszemy trzy możliwe podejścia do pozyskiwania informacji: bezpośrednią, z wykorzystaniem narzędzi oraz z wykorzystaniem wyszukiwarki. Narzędzia opisane w artykule są na chwilę powstawania tego tekstu bezpłatne dla użytkowników. Nie będziemy odnosić się do narzędzi proponowanych przez duże firmy i będących rozwiązaniami płatnymi. Nie będziemy również poruszać tematu analizy i weryfikacji pozyskiwanych danych.

W 2012 roku po raz pierwszy użyte zostało określenie SOCMINT (*Social Media Intelligence*), czyli wykorzystanie mediów społecznościowych do pozyskania informacji w ramach tzw. białego wywiadu (OSINT). W niniejszym artykule skupimy się na technikach i narzędziach, jakie wykorzystuje się w praktyce researchu. „Pozyskiwanie danych ze źródeł otwartych nie opiera się jednak na zdobywaniu dostępu do nich kanałami oficjalnymi ani na gromadzeniu informacji, do których dostęp jest przez użytkowników zastrzeżony. Opiera się na założeniu, że wiele przydatnych i znaczących informacji jest dostępnych publicznie”<sup>4</sup>. Informacje dostępne na platformach społecznościowych są umieszczane dobrowolnie, z dużą samoekspresją użytkowników i często jeszcze większym brakiem rozważań. Internauci publikują wiele informacji na swój temat, których w realnym świecie by tak łatwo nie ujawnili. Należy nadmienić, że rośnie również zagrożenie związane z celowo tworzonymi fałszywymi kontami, których celem jest między innymi szerzenie nieprawdziwych informacji, tzw. dezinformacji. Szerzej temat ten poruszono na końcu artykułu.

<sup>4</sup> P. Karasek, *Analiza informacji z mediów społecznościowych jako narzędzie wspierające kontrolę bezpieczeństwa w procedurach migracyjnych*, „Przegląd Bezpieczeństwa Wewnętrznego” 2018, nr 19, s. 200.

## Czym są media społecznościowe i co czyni je fenomenem?

Początki mediów społecznościowych szacuje się na 1997 rok. Powstał wtedy amerykański portal pod nazwą SixDegrees. Jako pierwszy wprowadził model „kręgów znajomych” i pozwalał na podtrzymywanie relacji ze znajomymi i rodziną. Kolejnym portalem był MySpace w 2000 roku. Dopiero cztery lata później powstał dzisiejszy lider, czyli Facebook, a zaraz za nim Twitter<sup>5</sup>. W Polsce, zanim prym objął globalny Facebook, mieliśmy własne platformy, jak najpopularniejsza Nasza Klasa, o której wciąż warto pamiętać pod kątem pozyskiwania informacji, oraz bardziej niszowe, mniej znane Grono.net. Wielu użytkowników tych platform dziś nawet nie pamięta, że posiadali tam konta.

Minęło ponad 20 lat od powstania pierwszych mediów społecznościowych, a można przypuszczać, że nadal nie każdy rozumie wagę pozyskiwania informacji za ich pośrednictwem. O popularności social mediów jako źródeł informacji w pewnych kręgach może świadczyć fakt, że informacja szybciej roznosi się na Twitterze lub portalu Wykop.pl niż ukazuje się w portalach informacyjnych. Do grup użytkowników świadomie pozyskujących informacje z mediów społecznościowych należą na pewno dziennikarze, politycy i coraz częściej osoby związane z mediami i marketingiem. Pomimo stale rosnącego zainteresowania i promowania social mediów w biznesie można się spotkać z opinią, jakoby korzystanie z nich było „zbyt młodzieżowe”, „zbyt niepoważne”.

Ponieważ publikowanie w mediach społecznościowych niektórym jeszcze kojarzy się zbyt młodzieżowo, nie zawsze dopuszcza się myśl o pozyskiwaniu informacji z tego zasobu jako źródła godnego uwagi. O ile niektóre dyscypliny naukowe, takie jak antropologia, socjologia, marketing czy politologia, zaczęły badać przekazy publikowane w social mediach oraz ich wpływ na decyzje zakupowe, o tyle temat SOCMINT-u czy samych metod pozyskiwania informacji jest poruszany bardzo rzadko lub prawie w ogóle. W polskiej literaturze ukazują się publikacje, które zarysowują istnienie takiej metody pozyskiwania danych. Temat „białego wywiadu” również nie należy do zagadnień powszechnie opisanych ani zbadanych w polskim piśmiennictwie naukowym. W literaturze obcojęzycznej temat jest lepiej opracowany, głównie pod kątem użycia SOCMINT-u w bezpieczeństwie międzynarodowym i walce z terroryzmem.

---

<sup>5</sup> K. Merrell, *The History of Social Media: Social Networking Evolution!*, <https://historycooperative.org/the-history-of-social-media/> (dostęp: 27.01.2019).

Nieliczne artykuły pokazują, jak można pozyskać informacje o konkurencji, preferencjach klientów czy o poglądach politycznych obywateli.

W mediach społecznościowych można znaleźć informacje praktycznie na każdy temat. Najogólniej podzielić można je na kilka grup:

1. Informacje o osobach (użytkownikach) – celem jest ustalenie jak największej liczby danych o konkretnej osobie: od statusu jej związku, pracy, zainteresowań, po relacje z innymi i poglądy<sup>6</sup>.
2. Informacje o firmach/organizacjach:
  - informacje publikowane przez firmy/organizacje w określonych celach;
  - informacje publikowane przez osoby postronne, na przykład klientów na temat współpracy z firmą;
  - informacje publikowane przez osoby związane obecnie lub kiedyś z daną firmą;
  - informacje o produktach, markach, usługach;
  - informacje pozwalające rozwiązać problem osobisty, rodzinny lub związany z życiem zawodowym – informacje poradnikowe;
  - informacje związane z bieżącą sytuacją w kraju, gospodarce czy innym interesującym użytkownika obszarem;
  - informacje o zjawiskach, zagadnieniach, tematach interesujących dla użytkownika.

Media społecznościowe na przestrzeni lat intensywnie się zmieniały. Ewolowały poszczególne platformy, ich popularność, grupy użytkowników, dochodziło też do przejęć, które intensywnie wpłynęły na zmianę globalnej mapy mediów społecznościowych. W 2011 roku autor nieaktualizowanego już dzisiaj bloga Oxyweb opublikował interaktywną mapę mediów społecznościowych, która pokazuje zmiany, jakie zaszły w strukturze platform społecznościowych na świecie w latach 2008–2011. Nie trzeba zaawansowanej analizy, by zauważyć, że zniknęło w tym czasie sporo lokalnych mniejszych platform, przejętych w części krajów przez najbardziej ekspansywnego gracza – Facebooka. Dziś najpopularniejsze obok niego platformy to Instagram<sup>7</sup>, Twitter ze swoją specyfiką restrykcyjnego limitowania długości wpisów, LinkedIn dedykowany profesjonalistom na skalę niemal globalną.

Oprócz tego moglibyśmy wskazać na obecność mniej lub bardziej popularnych platform społecznościowych, które wyróżniają się także ze

<sup>6</sup> Należy wziąć pod uwagę, że informacje udostępniane są przez użytkowników deklaratorywnie, a czasem część informacji świadomie przedstawiana jest w sposób nieprawdziwy.

<sup>7</sup> Właścicielem portalu Instagram jest firma Facebook.

względu na specyficzną grupę użytkowników, jak choćby ResearchGate dedykowany osobom związanym ze światem nauki. Warto też zwrócić uwagę na platformy przygotowane przede wszystkim dla użytkowników z określonych krajów: rosyjski odpowiednik Facebooka – V Kontakte, działające w Chinach WeChat – najpopularniejszy komunikator oraz ChinaWebo – odpowiednik Twittera. W krajach niemieckojęzycznych popularniejszą od LinkedIn siecią dla profesjonalistów jest Xing. Na uwagę zasługują także platformy społecznościowe, które dziś nie cieszą się już największym uznaniem internautów, ale wciąż stanowią repozytorium danych, których na próżno szukać w najpopularniejszych sieciach. W przypadku Polski można wskazać na przykład portal nk.pl (dawna Nasza Klasa) lub GoldenLine.

Nie jest naszym celem wyliczanie kolejnych platform społecznościowych i podkreślanie faktu ich istnienia. Chcemy jednak zwrócić uwagę na to, że przygotowując się do wyszukiwania informacji w mediach społecznościowych, należy zastanowić się nie tylko nad tym, jakich informacji, o kim i o czym potrzebujemy, ale także nad tym, które media społecznościowe powinniśmy przeszukać, aby zwiększyć swoje szanse na dotarcie do kluczowych dla nas informacji. Przy tym wszystkim trzeba zachować równowagę, bo nie zawsze więcej znaczy lepiej.

Żyjemy w czasach mediów społecznościowych, ponad 1/4 ludzkości komunikuje się za ich pomocą. Co takiego dają media społecznościowe, że internauci chcą z nich korzystać? Fenomen kryje się w łatwości dostępu do informacji o innych, poczuciu bycia na bieżąco, bezpieczeństwie komunikacji i w końcu łatwości w kreowaniu swojego wizerunku. Media społecznościowe dają też poczucie bycia w grupie, dla części osób jest to również sposób na podniesienie własnej pewności siebie.

## **Kluczem do wyszukiwania jest zrozumienie social mediów**

Dobry research zaczyna się od zrozumienia tematu, na jaki będziemy szukać informacji, ustalenia słów kluczowych i początkowych źródeł. W przypadku mediów społecznościowych dochodzi jeszcze zrozumienie, jak one funkcjonują, jakie funkcjonalności działają na poszczególnych portalach, jakie obowiązują zasady itd.

Pierwszym ważnym krokiem jest zrozumienie, czym różni się hashtag od słowa kluczowego. Hashtag to słowo lub zlepek słów poprzedzony znakiem # – na przykład #artykuł. Jest to ważne z punktu widzenia funkcjonowania wyszukiwarek. W niektórych platformach społeczno-

wych szukanie według hashtagów daje inne wyniki niż szukanie według słów kluczowych. Hashtagi są też istotne ze względu na ustalanie grup powiązanych ze sobą na przykład zainteresowaniem, uczestnictwem w wydarzeniu czy manifestacją określonych poglądów lub treści celowo przez autorów wiązanych w cykle tematyczne.

Kolejną istotną kwestią w zrozumieniu mediów społecznościowych jest rozróżnienie, na jakich platformach jakiego rodzaju informacje możemy znaleźć. Jeśli chcemy sprawdzić osobę, z którą mamy nawiązać biznes, to w pierwszej kolejności sprawdzamy ją na platformie biznesowej, na przykład LinkedIn. Jednak ważne jest też, aby sprawdzić, co jest publikowane w kanałach bardziej prywatnych, na przykład na Facebooku, bo może okazać się, że wzorowy wizerunek z LinkedIn nie pokrywa się z wypowiedziami w innych portalach. W social mediach kreowanie własnego wizerunku jest bardzo łatwe i może zwodzić też na manowce.

Nie wszystko, co jest publikowane w social mediach, zostaje zaindeksowane w Google lub innej wyszukiwarce. Wynika to z technicznych zabezpieczeń mediów społecznościowych. Z jednej strony są to ustawienia prywatności na profilach użytkowników, z drugiej są to na przykład informacje, do których dostęp z założenia ma być ograniczony (przykładowo tylko dla członków określonej grupy).

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.
3. szukanie w wyszukiwarce z wykorzystaniem zaawansowanych operatorów i techniki *Boolean string*.

## **Bezpośrednie techniki wyszukiwania**

Technikę wyszukiwania bezpośrednio w social mediach omówimy na przykładzie Facebooka i Twittera, ze względu na ich popularne użytkowanie w Polsce.

## SZUKANIE NA FACEBOOKU

Chcąc wyszukiwać informacje na Facebooku o jego użytkownikach, musimy zacząć od właściwego skonfigurowania konta. Chodzi tutaj przede wszystkim o zmianę ustawień językowych tak, aby jako domyślny język Facebooka wskazany został angielski amerykański (*English US*). Pozwoli to na skorzystanie z wyszukiwarki Facebook Graph Search i na wyszukiwanie informacji o użytkownikach Facebooka nawet w warunkach, kiedy nie mamy kompletu informacji na ich temat. Przydaje się to szczególnie ze względu na fakt, że coraz więcej osób korzystających z Facebooka modyfikuje swoje imiona i nazwiska w taki sposób, by ich wyszukiwanie po podstawowych parametrach było trudniejsze. Aby skorzystać z tych możliwości wyszukiwawczych, konieczne będzie konstruowanie kwerend w języku angielskim.

Przykładowe kwerendy:

- *People named Anna Kowalska* – osoby, które nazywają się Anna Kowalska;
- *People who live in Warsaw* – osoby, które mieszkają w Warszawie;
- *People who work at XYZ* – osoby, które pracują w XYZ;
- *People who live in Warsaw and work at XYZ* – osoby, które mieszkają w Warszawie oraz pracują w XYZ;
- *People who are interested in marketing* – osoby, które interesują się marketingiem;
- *People who checked in Hotel X* – osoby, które oznaczyły swoją lokalizację na Facebooku w hotelu X.

W przypadku wszystkich tych informacji pamiętać należy w pierwszej kolejności, że każda informacja podawana na Facebooku (a także w innych mediach społecznościowych) ma tylko charakter deklaracyjny i nie musi mieć wiele wspólnego z rzeczywistością.

Druga istotna sprawa dla nas, użytkowników korzystających z mediów społecznościowych jako narzędzi do wyszukiwania informacji, to fakt, że barierę, która będzie nas ograniczać w dostępie do informacji, stanowią ustawienia prywatności. Niewiele więc zdziałamy, jeśli chcemy wykorzystać Facebooka do profilowania osoby, której w gronie znajomych nie mamy, a która zadbała o ustawienia prywatności. Chodzi o ustawienie prywatności na konkretnych poziomach: jakie informacje mają być widoczne „tylko dla mnie”, a jakie informacje mogą zobaczyć „znajomi moich znajomych” lub „wszyscy”.

Jeszcze jeden fakt wpływający na dostępność informacji, z którego należy zdawać sobie sprawę w kontekście korzystania z Facebooka, to potęga grup. Grup, których zadaniem jest integrowanie użytkowników

o wspólnych potrzebach, zainteresowaniach itp. Trzeba mieć świadomość tego, że jeśli mamy do czynienia z grupą zamkniętą, nie będziemy mieć dostępu do jej zawartości, w tym aktywności jej członków, dopóki sami jej członkami nie zostaniemy. Tym samym nie poznamy pełnego spektrum zainteresowań/aktywności w mediach społecznościowych profilowanej osoby, a właściwie powinniśmy powiedzieć – konta. To ostatnie wynika bowiem z faktu, że nigdy nie mamy gwarancji, że osoba, której działania nas interesują, korzysta z mediów społecznościowych używając tylko konta założonego na swoje nazwisko. Dość powszechną praktyką jest korzystanie z tzw. avatarów, to jest kont założonych na niekiedy zupełnie fikcyjne dane.

Mówiąc o pozyskiwaniu treści z mediów społecznościowych, warto poruszyć także wątek ich regularnego pozyskiwania, w ramach monitoringu. Popularne, przede wszystkim wśród specjalistów marketingu, są narzędzia do monitoringu Internetu, w tym mediów społecznościowych. Wykorzystywane przede wszystkim do monitorowania, kto, gdzie i jak intensywnie dyskutuje o firmie czy marce, mogą kusić też specjalistów OSINT-u. Sam pomysł nie jest zły, jednak trzeba tutaj zwrócić uwagę na jedno istotne ograniczenie w działaniu tych narzędzi: nie mogą one monitorować zawartości grup działających na Facebooku. Nie dlatego, że ich twórcy nie potrafią opracować takich algorytmów. Z technicznego punktu widzenia jak najbardziej jest to możliwe, jednak w polityce Facebooka zawartość grup ma pozostawać treścią poniekąd ekskluzywną, dostępną jedynie dla członków danej grupy. Na etapie budowania strategii pozyskiwania informacji i jej wdrażania trzeba ten element koniecznie wziąć pod uwagę i albo zainwestować więcej czasu w przygotowanie warsztatu do pozyskiwania danych z Facebooka, albo pogodzić się z niedostępnością części treści.

#### SZUKANIE NA TWITTERZE

Twitter jest jednym z niewielu mediów społecznościowych, w których nie trzeba mieć konta, aby móc w nim szukać. W tym celu należy w dowolną wyszukiwarkę, choćby Google, wpisać kwerendę składającą się z nazwiska znanej osoby, która używa Twittera – na przykład „Robert Lewandowski Twitter”. Wyszukiwarka zwróci nam linki do profili na portalu. Jeśli wyświetlimy czyjeś konto w portalu, to w prawym górnym rogu pojawia się pole wyszukiwania. Można rozpocząć wyszukiwanie. Wpisujemy poszukiwane hasło. Po ukazaniu się wyników wyszukiwania należy wybrać odpowiednią z dostępnych zakładek:

- najlepsze – wyniki posortowane przez wewnętrzny algorytm portalu;



- najnowsze – czyli tweety (posty) zawierające szukane słowo;
- użytkownik – czyli konta mające w nazwie lub w opisie konta szukane słowo;
- zdjęcia opatrzone szukany słowem;
- filmy podpisane szukany słowem;
- news – posty podające linki do newsów zawierających szukane słowo;
- transmisje – posty zawierające nagrania wideo relacjonowane bezpośrednio na Twitterze, również nagrania archiwalne.

Po lewej stronie od wyników wyszukiwania dostępne jest pole o nazwie Dostępne filtry. W filtrach możemy ustalić parametr, taki jak język postów. Jest to szczególnie przydatne, gdy chcemy sprawdzić, co piszą media w innych państwach na temat na przykład polskich polityków czy sytuacji w naszym kraju. W tym polu dostępny będzie również link Wyszukiwanie zaawansowane. Tabela wyszukiwania zaawansowanego pozwala na (opis według kolejności pól):

1. szukanie pojedynczego słowa lub kilku słów;
2. szukanie dokładnego wyrażenia, na przykład „polityka zagraniczna”;
3. wykluczenie słów – aby ta funkcja działała, należy wpisać słowo w pozycji pierwszej – na przykład wpisz w niej słowo polityka, a w pozycji trzeciej słowo zagraniczna; co oznacza: „pokaż wyniki, w których jest słowo polityka, ale nie ma słowa zagraniczna”;
4. wyszukiwanie po hashtagu – bardzo ważne w przypadku szukania w mediach społecznościowych (jak wspomniano, słowo kluczowe i hashtag to nie to samo) – na przykład post o treści: „Jadę pociągiem na weekend #PKP #kochamkolej” – znajdziesz pod słowem pociąg i pod hashtagiem #PKP, ale pod słowem PKP już go nie będzie;
5. określenie języka, w jakim mają być posty – jest to powtórzenie z dostępnych filtrów, jednak daje możliwość stworzenia bardziej skomplikowanej kwerendy;
6. określenie miejsca, z którego został wysłany post – lokalizacja postów ustalana jest na podstawie adresów IP komputerów, z których zostały wysłane, w przypadku urządzeń mobilnych – z pomocą lokalizacji GPS;
7. określenie, jakie ma być zabarwienie emocjonalne wypowiedzi, której szukamy.

Każdy z parametrów dostępnych w tabeli ma również odpowiednik w komendach, które można wpisać bezpośrednio w pole wyszukiwania. Komenda near: pozwala na znalezienie postów ze wskazanej lokalizacji. Komenda -rt pozwala wykluczyć tak zwane retweety, czyli posty udostępnione od innego użytkownika.

Wyszukiwanie informacji na temat osób na Twitterze jest w porównaniu z Facebookiem o tyle bardziej problematyczne, że twórcy portalu nie przewidzieli różnych kont personalnych i – nazwijmy to – firmowych. Finalnie więc na etapie wyszukiwania informacji będziemy musieli zmierzyć się z koniecznością samodzielnego odfiltrowania kont personalnych od tych, które mogą należeć na przykład do firm, czasopism itd.

## Narzędzia ułatwiające zdobywanie i analizę informacji w social mediach

Social media generują ogromne liczby danych, dlatego coraz częściej powstają specjalistyczne aplikacje lub programy komputerowe, których zadaniem jest monitorowanie lub agregowanie danych ze wskazanych platform. „Uzyskiwanie dostępu do informacji oraz ich zdobywanie może następować kompleksowo lub częściowo, w formie procesu zautomatyzowanego, bez konieczności uciążliwego ręcznego »klikania« w treści na portalu społecznościowym”<sup>8</sup>, dlatego dalej prezentujemy kilka wybranych narzędzi ułatwiających pozyskiwanie informacji z mediów społecznościowych. Do opisu wybrano narzędzia dostępne dla każdego i do łatwego znalezienia w sieci.

### STALKSCAN

Wyszukiwanie informacji na Facebooku można uczynić skuteczniejszym poprzez wykorzystanie kwerend w Facebook Graph Search. Gdy już przejdziemy do analizy profilu danej osoby, może jednak okazać się, że mamy do czynienia z jednostką mocno ekstrawertyczną, publikującą sporo treści, bardzo aktywną na Facebooku. Może to spowodować, że analiza profilu takiego użytkownika okaże się dla nas mocno uciążliwa, dlatego z pomocą przyjdzie nam narzędzie StalkScan.

Jest to bardzo prosta w obsłudze aplikacja, której zadaniem jest pokazanie w kawałkach profilu użytkownika Facebooka. Wklejamy link do analizowanego profilu lub podajemy numer użytkownika Facebooka (tzw. ID). Jeśli tylko mamy do czynienia z profilem, którego właściciel nie zadbał wystarczająco o ustawienia prywatności, będziemy mogli otwierać sobie wybrane „kawałki” analizowanego profilu. Jeśli klikniemy na przykład w ikonę *Pictures*, narzędzie przeniesie nas do zdjęć opublikowanych na analizowanym profilu, ikona *All* w sekcji *Places* pozwoli nam podejrzeć, w jakich miejscach sprawdzana osoba się oznaczyła.

<sup>8</sup> P. Karasek, *Analiza informacji z mediów społecznościowych...*, s. 201.

StalkScan – i inne podobne jej aplikacje – nie łamie ustawień prywatności, dlatego dobra konfiguracja tychże będzie zawsze stanowić barierę w pozyskiwaniu informacji z profili użytkowników.

Niektóre z tego rodzaju narzędzi ułatwiających pracę specjalisty OSINT-u będą wymagały podania tylko i wyłącznie numeru ID konta na Facebooku. W jaki sposób go pozyskać? Sposoby są dwa. Pierwszy sposób, to umiejętne analizowanie i przyglądanie się temu, co dzieje się w trakcie korzystania z narzędzi automatyzujących naszą pracę. Kiedy już korzystamy ze StalkScan i wygenerujemy sobie na przykład czyjeś *Pictures*, automatycznie w adresie URL łąduje numer ID konta, które jest poddawane analizie. Będzie to ciąg 15 cyfr jednoznacznie identyfikujący konkretne konto.

Drugi sposób to skorzystanie z narzędzia dedykowanego do odczytywania numeru ID, mianowicie aplikacji Find My Fb ID ([www.find-myfbid.com](http://www.find-myfbid.com)). Mechanizm postępowania jest podobny jak w przypadku SkalkScan: wklejamy link do profilu użytkownika, a jako rezultat otrzymujemy ID konta, który możemy wykorzystać z powodzeniem w dalszych analizach.

#### SEARCH IS BACK

Search Is Back ([searchisback.com](http://searchisback.com)) to jeszcze jedna aplikacja, której zadaniem jest ułatwianie nam wyszukiwania treści pochodzących z Facebooka. Praca na tym narzędziu jest bardzo intuicyjna, a ono samo wykorzystuje znane nam już metody tworzenia zapytań z wykorzystaniem Facebook Graph Search. Tyle tylko, że teraz możemy tę pracę zautomatyzować.

Na początek będziemy musieli zdefiniować, czego mają dotyczyć wyszukiwane przez nas informacje. Do wyboru mamy: osoby, zdjęcia, wydarzenia i posty. Jeśli zdecydujemy się na wyszukiwanie informacji o użytkownikach Facebooka, będziemy musieli w kolejnym kroku wybrać zbiorowość poddawaną analizie: wszyscy użytkownicy, tylko nasi znajomi, znajomi znajomych lub osoby spoza kręgu naszych facebookowych znajomych.

Na kolejnym etapie pozostanie nam – korzystając z dostępnych filtrów – definiować kryteria wyszukiwania osób. Do wyboru będziemy mieć parametry odwołujące się do deklarowanego na Facebooku miejsca pracy, miejsca zamieszkania, płci czy wreszcie imienia i nazwiska. Dodatkowo dla parametrów zmiennych, jak miejsce zamieszkania bądź zatrudnienie, mamy możliwość zdefiniowania, czy chodzi nam o stan obecny, o przeszłość, czy o oba te warianty.

Po zdefiniowaniu wymaganych parametrów wyszukiwania narzędzie wywoła nam już na Facebooku wyniki kwerendy. Podobny mechanizm działań będzie do wykorzystania podczas wyszukiwania treści innych niż profile użytkowników Facebooka.

#### RECRUITIN.NET

recruitin.net to portal typu *x-ray search*. To anglojęzyczne sformułowanie oznacza wyszukiwanie informacji z wykorzystaniem zaawansowanych komend wyszukiwawczych, o których opowiadamy w dalszej części niniejszego artykułu (*Google hacking*). Bardzo często o wyszukiwaniu *x-ray* mówi się w kontekście branży HR i wyszukiwania kandydatów do pracy oraz weryfikacji ich CV.

Recruitin to nic innego jak narzędzie, które stanowi swego rodzaju nakładkę na Google, ułatwiającą nam pisanie zaawansowanych kwerend w wyszukiwarce koncentrujących się na wybranym portalu: LinkedIn, Xing, Dribbble, Twitter, GitHub, StackOverflow.

Za pomocą Recruitin możemy zdefiniować kryteria wyszukiwawcze, które, na przykładzie LinkedIn, będą dotyczyły miejsca zatrudnienia, zajmowanego stanowiska, kraju, w którym ma pracować nasz kandydat. Jeśli szukamy konkretnej osoby, jej imię i nazwisko trzeba będzie potraktować po prostu jako słowa kluczowe wymuszane w kwerendzie. Dla równowagi Recruitin pozwala nam też na zdefiniowanie słów, które wykluczamy z naszych wyników wyszukiwania.

Być może oczekivalibyśmy, że od razu po uruchomieniu wyszukiwania Recruitin przeniesie nas do przeszukiwanego portalu, by wyświetlić nam wyniki wyszukiwania. Tymczasem jako rezultat kwerendy otrzymujemy tzw. *string* do Google, a więc kwerendę, którą możemy w Google uruchomić. Jako rezultat naszego przeszukiwania LinkedIn za pomocą Recruitin otrzymamy listę kont użytkowników LinkedIn spełniających nasze kryteria. Częstokroć profile te będziemy mogli zobaczyć w pełnej wersji, bez logowania do LinkedIn. Możliwości te zostaną ograniczone wtedy, gdy będziemy chcieli podejrzeć zbyt wiele kont lub wcześniej logowaliśmy się z tego samego komputera i IP do naszego konta w LinkedIn. Dostęp, przynajmniej częściowo anonimowy, do treści w Internecie to jednak osobny wątek i temat na odrębny artykuł.

#### ONEMILLIONTWEETMAP.COM

Onemilliontweetmap to aplikacja łącząca w sobie dwa intensywne trendy w podejściu do informacji: nieustannego monitoringu oraz wizualizacji. Na interaktywnej mapie otrzymujemy wizualizację intensywności

życia Twittera w czasie rzeczywistym, na bieżąco. Możemy prześledzić, w których częściach świata użytkownicy Twittera są najbardziej aktywni. Co więcej, narzędzie pozwala na analizowanie tego ruchu. Możemy wykorzystać w tym celu operatory, które zostały opisane we fragmencie dotyczącym wyszukiwania treści bezpośrednio na Twitterze.

Interesującą funkcjonalnością narzędzia jest możliwość wyfiltrowania tweetów opublikowanych w danej lokalizacji z możliwością ich czytania bezpośrednio na portalu, bez przechodzenia do Twittera i logowania się w portalu.

Możemy wykorzystać także mapę ciepła – popularnej w tego rodzaju narzędziach funkcjonalności, odzwierciedlającej zaangażowanie użytkowników Twittera.

Możliwość wykorzystania tego narzędzia do analizy ruchu na Twitterze są bardzo duże i tak naprawdę ograniczane tylko i wyłącznie wyobraźnią analityka. Nie jest to jednak jedyne warte uwagi narzędzie do analizy tego, co na Twitterze się dzieje, a dobór aplikacji będzie też finalnie zależał od subiektywnych upodobań i potrzeb poszukującego informacji.

## TWITONOMY

By pokazać możliwości różnego podejścia do analizowania i przetwarzania treści pochodzących z Twittera, poświęcimy nieco uwagi narzędziu Twitonomy. To aplikacja, która – w odróżnieniu od Onemillion-tweetmap – każdorazowo będzie wymagać od użytkownika zalogowania i „wpięcia” jej do Twittera. Po zalogowaniu będziemy mogli przystąpić do wykorzystania możliwości aplikacji. Jest to narzędzie, dzięki któremu będziemy mogli monitorować wybrane konta na Twitterze, monitorować hashtagi i odczytywać pojawiające się w poszczególnych wątkach tweety bezpośrednio z poziomu Twitonomy, a nie Twittera. Będzie to oczywiście monitoring dokonywany w czasie rzeczywistym, a więc archiwum może obejmować, w zależności od intensywności dyskusji w danym wątku, kilka do kilkudziesięciu godzin.

Przez specjalistów marketingu narzędzie może być wykorzystywane także do analizy skuteczności działań na Twitterze, ale to oczywiście osobny wątek. Warto jednak zwrócić uwagę na fakt, że wielokrotnie, zajmując się OSINT-em, docieramy do miejsc, w których wykorzystujemy narzędzia pierwotnie przygotowane z myślą o grupach zawodowych zajmujących się innymi obszarami. W tym przypadku mówimy o marketingu, który jak najbardziej korzysta z dobrodziejstw OSINT-u, nie zawsze nazywając to wprost.

Podsumowując, przedstawiliśmy kilka propozycji narzędzi, które dedykowane są przeszukiwaniu treści z platform społecznościowych. Jest to ledwie sygnalizacja możliwości, a temat do całościowego opisania jest tym trudniejszy, że same narzędzia, o których mowa, często się zmieniają. Pojawiają się, działają, po czym znikają i zastępowane są przez kolejne aplikacje. Wszystko to prowadzi do tego, że chcąc być na bieżąco z tematem, będziemy musieli testować różne narzędzia i uważnie się przyglądać ich możliwościom i ograniczeniom. Niewątpliwie w pewnym momencie natknemy się też na barierę związaną z koniecznością uiszczenia opłaty za korzystanie z poszczególnych aplikacji. Coraz więcej narzędzi jest udostępnianych w modelu tzw. *freemium*, pozwalającym na korzystanie bez opłat do pewnego momentu, ograniczonego na przykład czasem lub liczbą zrealizowanych kwerend. Po wykorzystaniu limitu konieczne jest jednak opłacenie subskrypcji. Niektóre z narzędzi, w szczególności te najbardziej zaawansowane, w ogóle nie są dostępne do wytestowania bez ponoszenia opłat. Przykład to choćby Wynyard Social Media Analyzer – aplikacja stworzona przede wszystkim z myślą o prewencji incydentów kryminalnych i nacjonalistycznych. Nie jest to też narzędzie, na którego zakup będzie mogło pozwolić sobie mikroprzedsiębiorstwo z niewielkim budżetem.

## **Wyszukiwanie z wykorzystaniem technik *Google hacking***

Znajomość narzędzi do wyszukiwania treści w mediach społecznościowych niewątpliwie może ułatwiać życie. Warto jednak nie tylko dobrze rozumieć mechanizmy działania tych narzędzi, ich wady i zalety, możliwości i ograniczenia, ale też umieć niekiedy obejść się po prostu bez nich. Dobrą strategią będzie znajomość wspomnianych już technik *x-ray search*, a więc budowania kwerend wyszukiwawczych z wykorzystaniem komend, na przykład w wyszukiwarce Google.

W przypadku wyszukiwania informacji w Google przydatna jest znajomość składni adresu URL strony, którą chcemy przeszukać. Komend wyszukiwania zaawansowanego w wyszukiwarkach jest kilkanaście, jednak przy mediach społecznościowych najbardziej przydadzą nam się cztery z nich:

1. komenda `site:` na przykład „anna kowalska” `site:facebook.com` – pozwoli na wyszukiwanie osób o imieniu i nazwisku „Anna Kowalska” na portalu Facebook;
2. komenda `inurl:` na przykład `inurl:polityka site:facebook.com/pages` – pozwoli na wyszukanie stron na Facebooku, w których adresie `www`

pojawia się słowo polityka (bo tylko te są indeksowane w Google), postów na Facebooku mających w swojej treści słowo polityka, przykładowo: <https://www.facebook.com/pages/biz/Polityka-w-Europie-1239066012883016/>;

3. filetype: na przykład biotechnologia filetype:pdf site:researchgate.net – wymusi wyszukiwanie plików pdf ze słowem kluczowym „biotechnologia” w serwisie ResearchGate;
4. intitle: na przykład intitle: „Jan Kowalski” site:linkedin.com – Google wyszuka strony, które zawierają w tytule słowo „Jan Kowalski” w obrębie portalu LinkedIn.

Znajomość i umiejętność wykorzystania w praktyce operatorów wyszukiwania zaawansowanego daje dużo elastyczności w budowaniu kwerend. Pozwala też zrozumieć mechanizmy stojące za wyszukiwaniem treści, a finalnie także przygotować własne narzędzia automatyzujące kwerendy.

## Podsumowanie

Media społecznościowe wywierają realny wpływ na życie polityczne i społeczne. Zbyt duża liczba informacji pojawiająca się w nich sprawia, że coraz więcej osób ma problem z ich weryfikacją. Stąd pojawiające się nierzadko doniesienia dotyczące manipulacji, prowokacji lub celowych działań dezinformacyjnych. Najnowsze dane pokazują, że prawie połowa kont w social mediach może być fałszywa. Zatem umiejętność skutecznego wyszukiwania informacji w tych zasobach staje się coraz istotniejsza. Opisane metody i narzędzia pomocne będą zarówno dla osób zajmujących się researchem, jak i dla analityków bezpieczeństwa, marketerów, handlowców czy managerów. Research w mediach społecznościowych pozwala ustrzec się przed pułapką *fake news* i innych nieprawdziwych informacji. Rośnie liczba osób, dla których opisywany kanał komunikacji staje się jedynym sposobem na pozyskiwanie informacji bieżącej o świecie. Jednak nie idzie za tym rosnąca świadomość zjawiska, jakim jest *bubble filters* (po polsku bańka informacyjna). O personalizacji wyników przez wyszukiwarkę wie już prawie każdy użytkownik Google. Jednak mało kto wie, że to, co ukazuje nam się w mediach społecznościowych, nie jest informacją obiektywną, ale dostosowaną do naszych preferencji, jakie ustalił algorytm na podstawie zgromadzonych o nas danych. W celu wyjścia z bańki informacyjnej konieczne jest weryfikowanie informacji, a tego nie da się zrobić bez rozwijania kompetencji cyfrowych w zakresie wyszukiwania informacji.

## STRESZCZENIE

Media społecznościowe stały się popularnym kanałem informacji dla dużej części społeczeństwa. Liczba informacji umieszczanych tam każdego dnia pozostaje w większości poza zasięgiem wyszukiwarek. Celem tego artykułu jest pokazanie technik i narzędzi, które można wykorzystać do przeszukiwania zasobów social mediów. Zaprezentowano tylko te narzędzia, które działają na rynku polskim i są bezpłatne przynajmniej w wersji podstawowej.

*Patrycja Hrabiec-Hojda, Justyna Trzeciakowska*

## INFORMATION RETRIEVAL TECHNIQUES IN SOCIAL MEDIA FOR OPEN SOURCE INTELLIGENCE PURPOSE

Social media has become a popular information channel for a large part of society. The amount of information that is placed there every day remains mostly beyond the reach of search engines. The purpose of this article is to show the techniques and tools that can be used to search and explore social media. Only those tools that operate on the Polish market and are free at least in the basic version are presented.

**KEY WORDS:** *social media, SOCMINT, Social Media Intelligence, OSINT, information retrieval*

## Bibliografia

- Bartosik-Purgat M., *Media społecznościowe jako źródło informacji o produktach w świetle badań międzykulturowych – przykład Facebooka*, „Handel Wewnętrzny” 2016, nr 6.
- Dzikowski J., *Wyszukiwanie danych osobowych w internecie dla celów informatyki śledczej*, „Studia Oeconomica Posnaniensia” 2013, nr 1.
- Karasek P., *Analiza informacji z mediów społecznościowych jako narzędzie wspierające kontrolę bezpieczeństwa w procedurach migracyjnych*, „Przegląd Bezpieczeństwa Wewnętrznego” 2018, nr 19.
- Merrell K., *The History of Social Media: Social Networking Evolution!*, <https://historycooperative.org/the-history-of-social-media/> (dostęp: 27.01.2019).
- Omad D., Bartlett J., Miller C., *Introducing Social Media Intelligence (SOCMINT)*, „Intelligence and National Security” 2012, nr 27 (6).