

## Black Hat SEO

### Wprowadzenie

Jeśli spędziłeś dużo czasu w Internecie, najprawdopodobniej natkniesz się na termin. Black Hat w tym czy innym czasie. Termin ten jest zwykle związany z wieloma negatywnymi komentarzami. Ten tekst jest po to, by zająć się tymi komentarzami i zapewnić wgląd w prawdziwe życie zawodowego SEO Black Hat. Jak już wkrótce omówimy, nie możesz być świetnym Black Hatem bez zostania świetnym White Hatem.

### Co to jest SEO Black Hat?

Pytanie za milion dolarów, o którym wszyscy mają różne zdanie. Czym dokładnie jest SEO Black Hat? Odpowiedź tutaj zależy w dużej mierze od tego, kogo zapytasz. Zapytaj większość Białych Kapeluszy i od razu cytują Wskazówki Google dla webmasterów, takie jak grupa lemingów. Czy naprawdę zastanawiałeś się nad tym? Google publikuje te wytyczne, ponieważ wiedzą tak dobrze, jak Ty i ja, że nie mają możliwości wykrzyka lub powstrzymania tego, o czym tak głośno głoszą. Polegają oni na tłumach webmasterów, którzy na ślepo powtarzają wszystko, co mówią, ponieważ są internetowym motorem napędowym i każą wszystkim pracować mózgi, aby wierzyć w cokolwiek im mówią. Jest to jednak naprawdę dobra rzecz. Oznacza to, że zdecydowana większość specjalistów od marketingu internetowego i specjalistów SEO jest całkowicie ślepa na szeroki wachlarz narzędzi, które nie tylko zwiększają ruch na swoich stronach, ale także generują każdego roku miliony przychodów. Drugim argumentem, który najprawdopodobniej usłyszysz, jest wiek, "wyszukiwarki będą blokować twoje witryny, jeśli używasz technik Black Hat". Oczywiście, to prawda, jeśli nie rozumiesz podstawowych zasad i praktyk. Jeśli wskoczysz bez wiedzy, zawiedziesz. Dam ci tajemnicę. Gotowy? Nie używaj technik czarnego kapelusza w domenach White Hat. Nie bezpośrednio. Nie będziecie budować wejścia lub zamaskowanych stron na stronie z pieniędzmi, to byłoby idiotyczne. Zamiast tego kupujesz kilka wyrzucanych domen, budujesz na nich wejście i maskujesz / przekierowujesz ruch na strony z pieniędzmi. Tracisz domenę wejścia, kogo to obchodzi? Zbudujesz 10, aby je zastąpić. To nie jest rakietą, tylko zdrowy rozsądek. Wyszukiwarka nie może ukarać cię za zewnętrzne wpływy, które są poza twoją kontrolą. Nie mogą cię ukarać za linki przychodzące, ani też nie mogą cię ukarać za wysłanie ruchu do twojej domeny z innych stron przejściowych poza tą domeną. Gdyby mogli, po prostu wskazywałbym stronie wejście i linki spamowe na moich konkurentów, aby wytrącić je z SERPS. Zobacz ..... Zdrowy rozsądek. A zatem, czym jest SEO Black Hat? Moim zdaniem, SEO Black Hat i SEO White Hat prawie nie różnią się. Mistrzowie stron internetowych w białych kapeluszach poświęcają czas na uważne znajdowanie partnerów łączących, aby podnieść ranking swoich słów kluczowych. Czarne kapelusze robią to samo, ale piszemy zautomatyzowane skrypty, aby to zrobić podczas snu. Biały kapelusz SEO spędza miesiące doskonaląc na stronie pozycjonowanie swoich witryn w celu maksymalizacji rankingów, czerpiąc z SEO generatory zawartości stron internetowych, aby wypluć tysiące wygenerowanych stron, aby zobaczyć, która wersja działa najlepiej. Zaczynasz widzieć tutaj wzór? Powinieneś, Black Hat SEO i White Hat SEO są podobni w jednym z jedną kluczową różnicą. Czarne kapelusze są leniwe. Lubimy rzeczy zautomatyzowane. Czy kiedykolwiek słyszałeś zdanie "Pracuj mądrzej, nie mocniej?" Żyjemy według tych słów. Po co spędzać tygodnie lub miesiące tylko na budowaniu stron, aby Google zadał im trochę niejasnych kar. Jeśli spędziłeś jakiś czas na forach internetowych, poznałeś tę historię raz za razem. Mistrz stron internetowych gra według reguł, nie robi nic złego, ale ich strona jest zupełnie nieobecna z SERP (stron z wynikami wyszukiwania) pewnego ranka bez wyraźnego powodu. To frustrujące, wszyscy tam byliśmy. Miesiące pracy minęły i nic do pokazania. Zmęczyło mnie to, tak jak jestem pewien. Właśnie wtedy do mnie dostrzeli. Ktoś wybrał wyszukiwarki internetowe na "policję internetową"? Z pewnością nie, więc po co grać według ich

zasad? Na następnych stronach pokażę wam, dlaczego reguły wyszukiwarek nie mają sensu, a następnie omówię, w jaki sposób można wykorzystać te informacje na swoją korzyść.

## **Wyszukiwarka**

Jak mówiliśmy wcześniej, każdy dobry Czarny Kapelusz musi być solidnym Białym Kapeluszem. Zaczniemy więc od podstaw. Ta sekcja będzie miała charakter techniczny, ponieważ omawiamy działanie wyszukiwarek i zagłębiamy się w sposoby wykorzystania tych wewnętrznych mechanizmów. Zaczniemy, dobrze? Wyszukiwarki dopasowują zapytania do indeksu, który tworzą. Indeks składa się ze słów w każdym dokumencie oraz wskaźników do ich lokalizacji w dokumentach. Nazywa się to odwróconym plikiem. System wyszukiwania lub system IR (Information Retrieval) składa się z czterech podstawowych modułów:

- \* Procesor dokumentów
- \* Procesor zapytań
- \* Funkcja wyszukiwania i dopasowywania
- \* Możliwość rankingu

### **Procesor dokumentów**

Procesor dokumentów przygotowuje, przetwarza i wprowadza dokumenty, strony lub witryny, na które użytkownicy wyszukują. Procesor dokumentów wykonuje niektóre lub wszystkie z następujących kroków:

- \* Normalizuje strumień dokumentów do predefiniowanego formatu.
- \* Łamie strumień dokumentów w pożądane jednostki do odzyskania.
- \* Wyodrębnia elementy podtekstu i metatagi.
- \* Identyfikuje potencjalne elementy indeksowane w dokumentach.
- \* Usuwa słowa zatrzymania.
- \* Wytrąca terminy.
- \* Wyodrębnia wpisy indeksu.
- \* Oblicza wagi.
- \* Tworzy i aktualizuje główny odwrócony plik, na podstawie którego wyszukiwane są wyszukiwane hasła w celu dopasowania zapytań do dokumentów.

**Krok 4:** Wskaż elementy do indeksu. Rozpoznanie potencjalnych elementów do indeksowania w dokumentach radykalnie wpływa na charakter i jakość reprezentacji dokumentów, które silnik będzie wyszukiwał. Przy projektowaniu systemu musimy zdefiniować słowo "termin". Czy jest to znak alfanumeryczny między pustymi polami lub interpunkcją? Jeśli tak, to co z wyrażeniami nieskompozycyjnymi (wyrażenia, w których poszczególne słowa nie oddają znaczenia frazy, np. "Skunk works" lub "hot dog"), wieloznaczne nazwy własne lub symbole między wyrazami, takie jak łączniki lub apostrofy, które mogą oznaczać różnicę między "małymi biznesmenami" a małymi przedsiębiorcami. "Każda wyszukiwarka zależy od zestawu reguł, które musi wykonać procesor dokumentu, aby określić, jakie działanie ma podjąć" tokenizator ", tj. oprogramowanie używane do definiowania terminu odpowiedniego do indeksowania.

**Krok 5:** Usuwanie słów zatrzymania. Ten krok pomaga zaoszczędzić zasoby systemowe, eliminując z dalszego przetwarzania, a także potencjalnego dopasowywania, te terminy, które mają niewielką wartość w znajdowaniu przydatnych dokumentów w odpowiedzi na zapytanie klienta. Ten krok miał znacznie większe znaczenie niż teraz, gdy pamięć stała się o wiele tańsza, a systemy o wiele szybsze, ale ponieważ słowa zatrzymania mogą zawierać do 40 procent słów tekstu w dokumencie, nadal ma ono pewne znaczenie. Lista słów kończących się zwykle składa się z tych klas słów, które mają przekazywać niewielkie znaczenie merytoryczne, takie jak artykuły (a, the), spójniki (i, ale), wykrzykniki (o, ale), przyimki (w, nad,), zaimki (on, it) i formy czasownika "być" (is, are). Aby skasować słowa kończące, algorytm porównuje indeksowane terminy kandydatów w dokumentach z listą słów kończących i eliminuje pewne warunki z włączenia do indeksu do wyszukiwania

**Krok 6:** Wytrącanie terminu. Stemming usuwa sufiksy słów, być może rekurencyjnie w warstwie po warstwie przetwarzania. Proces ma dwa cele. Pod względem wydajności, zmniejszanie liczby zmniejsza liczbę unikalnych słów w indeksie, co z kolei zmniejsza przestrzeń dyskową wymaganą dla indeksu i przyspiesza proces wyszukiwania. Pod względem skuteczności, poprawa poprawia przypomnienie, redukując wszystkie formy słowa do formy bazowej lub łodygi. Na przykład, jeśli użytkownik prosi o analizę, może również chcieć dokumentów zawierających analizę, analizę, analizę, analizę i analizę. W związku z tym procesor dokumentu opiera się na pojęciach dokumentu w celu analizy, że dokumenty zawierające różne formy analizy będą miały takie samo prawdopodobieństwo odzyskania; nie wystąpiłoby to, gdyby silnik tylko indeksował wariant oddzielnie i wymagał od użytkownika wprowadzenia wszystkich. Oczywiście, wynik jest wadliwy. Może to negatywnie wpływać na precyzję, ponieważ wszystkie formy trzpienia będą pasować, gdy w rzeczywistości pomyślnie zapytanie dla użytkownika wynikałoby z dopasowywania wyłącznie formy słowa użytej w zapytaniu.

Systemy mogą implementować albo silny algorytm postępowania albo słaby algorytm postępowania. Silny algorytm rozstrzygający usunie oba sufiksy fleksyjne (-s, -es, -ed) i sufiksy pochodne (-able, -aciousness, -ability), podczas gdy słaby algorytm rozstrzygający usunie tylko sufiksy fleksyjne (-s, -es, -ed).

Krok 7: Wyodrębnić wpisy indeksowe. Po wykonaniu kroków od 1 do 6 procesor dokumentu wyodrębnia pozostałe wpisy z oryginalnego dokumentu. Na przykład poniższy akapit pokazuje pełny tekst wysłany do wyszukiwarki w celu przetworzenia:

Komentarze Milosevica, prowadzone przez oficjalną agencję prasową Tanjug, rzucają cień na rządy na rozmowach, które społeczność międzynarodowa wezwała, aby spróbować zapobiec totalnej wojnie w serbskiej prowincji. "Prezydent Milosevic powiedział, że dobrze wiadomo, że Serbia i Jugosławia były zdecydowanie zaangażowane w rozwiązywanie problemów w Kosowie, który jest integralną częścią Serbii, pokojowo w Serbii przy udziale przedstawicieli wszystkich społeczności etnicznych" - powiedział Tanjug. Milosevic przemawiał podczas spotkania z brytyjskim ministrem spraw zagranicznych Robinem Cookiem, który wygłosił ultimatum, aby w ciągu tygodnia wziąć udział w negocjacjach w sprawie propozycji autonomii dla Kosowa z etnicznymi albańskimi przywódcami z prowincji. Cook wcześniej powiedział na konferencji, że Milosevic zgodził się zbadać tę propozycję.

Kroki od 1 do 6 redukują ten tekst do wyszukiwania następujących elementów:

Milosevic comm carri offic nowy agen Tanjug cast wątpliwość rządzą rozmowa interna komunalne wezwanie spróbuj zapobiegać wszystkim-out wojna Serbska prowincja Prezydent Milosevic powiedział, że dobrze znana Serbii Jugosławia firma zdecydowanie rozwiązuje problem Kosowo integruje część Serbia pokój Serbia udział reprezentuje etniczną komunę Tanjug powiedział Milosevic speak meeti British Foreign Sekretarz Robin Cook zapewnia ultimatum udział w negocjacjach autonomia

czasu tygodniowego Propos Kosowo etniczna prowincja Alban przewodniczący konferencji Cook Earl powiedział Milosevic zgadza się z propozycją badania.

Wyjście z kroku 7 jest następnie wstawiane i przechowywane w odwróconym pliku, który zawiera wpisy indeksu oraz wskazanie ich pozycji i częstotliwości występowania. Specyfika wpisów indeksu będzie jednak różnić się w zależności od decyzji podjętej w kroku 4 dotyczącej tego, co stanowi "termin indeksowalny". Bardziej wyrafinowane procesory dokumentów będą miały identyfikatory, a także rozpoznawalność nazw i kategorie, aby upewnić się, że wpisy indeksu, takie jak Milosevic, są oznaczone jako osoba i wpisy takie jak Jugostawia i Serbia jako kraje.

**Krok 8:** Termin przypisanie wagi. Wagi są przypisane do terminów w pliku indeksu. Najprostszy z wyszukiwarek po prostu przypisuje wagę binarną: 1 dla obecności i 0 dla nieobecności. Im bardziej zaawansowana jest wyszukiwarka, tym bardziej złożony jest schemat ważenia. Mierzenie częstotliwości występowania terminu w dokumencie tworzy bardziej wyrafinowaną wagę, z normalizacją długości częstotliwości jeszcze bardziej wyrafinowaną. Duże doświadczenie w badaniach wyszukiwania informacji przez wiele lat wyraźnie pokazało, że optymalna waga pochodzi z użycia "tf / idf". Algorytm ten mierzy częstotliwość występowania każdego terminu w dokumencie. Następnie porównuje tę częstotliwość z częstotliwością występowania w całej bazie danych.

Nie wszystkie terminy są dobrymi "dyskryminatorami" - to znaczy wszystkie terminy nie wyróżniają bardzo dobrze jednego dokumentu od drugiego. Prostym przykładem może być słowo "the". To słowo występuje w zbyt wielu dokumentach, aby pomóc odróżnić jeden od drugiego. Mniej oczywistym przykładem może być słowo "antybiotyk". W sportowej bazie danych, gdy porównamy każdy dokument z bazą danych jako całością, termin "antybiotyk" prawdopodobnie byłby dobrym wyróżnikiem wśród dokumentów, a zatem miałby przypisaną dużą wagę. I odwrotnie, w bazie danych poświęconej zdrowiu lub medycynie "antybiotyk" prawdopodobnie byłby słabym dyskryminatorem, ponieważ występuje bardzo często. Schemat ważenia TF / IDF przypisuje wyższe wagi tym warunkom, które naprawdę odróżniają jeden dokument od pozostałych.

### **Procesor zapytań**

Przetwarzanie zapytania ma siedem możliwych kroków, ale system może skrócić te kroki i przystąpić do dopasowania zapytania do odwróconego pliku w dowolnym miejscu podczas przetwarzania. Przetwarzanie dokumentów dzieli wiele kroków z przetwarzaniem zapytań. Więcej kroków i więcej dokumentów sprawia, że proces jest droższy w przetwarzaniu pod względem zasobów obliczeniowych i szybkości reakcji. Jednak im dłuższy czas oczekiwania na wyniki, tym wyższa jakość wyników. Dlatego projektanci systemów wyszukiwania muszą wybrać to, co jest najważniejsze dla ich użytkowników - czas lub jakość. Publicznie dostępne wyszukiwarki zazwyczaj wybierają czas na bardzo wysoką jakość, mając za dużo dokumentów do przeszukiwania.

Kroki w przetwarzaniu zapytań są następujące (z opcją zatrzymania przetwarzania i rozpoczęcia dopasowywania oznaczoną jako "Matcher"):

W tym momencie wyszukiwarka może pobrać listę terminów zapytania i przeszukać je względem odwróconego pliku. W rzeczywistości jest to punkt, w którym większość publicznie dostępnych wyszukiwarek wykonuje wyszukiwanie.

\* Warunki zapytania tokenize.

Rozpoznawanie terminów zapytań a operatory specjalne.

-----> Matcher

\* Usuń słowa kończące.

\* Słowa kluczowe.

\* Utwórz reprezentację zapytań.

-----> Matcher

\* Rozwiń warunki zapytania.

\* Oblicz masę.

-----> Matcher

**Krok 1:** Tokenizacja. Gdy tylko użytkownik wprowadza zapytanie, wyszukiwarka - system oparty na słowach kluczowych lub pełny system przetwarzania języka naturalnego (NLP) - musi tokenizować strumień zapytania, tj. Podzielić go na zrozumiałe segmenty. Zwykle token jest definiowany jako ciąg alfanumeryczny występujący między białą spacją i / lub interpunkcją.

**Krok 2:** Parsowanie. Ponieważ użytkownicy mogą wykorzystać specjalne operatory w swoich zapytaniach, w tym operatorów logicznych, sąsiadujących lub bliskich, system musi najpierw przetworzyć zapytanie na zapytania i operatory. Operatory te mogą występować w formie zastrzeżonej interpunkcji (np. cudzysłówów) lub zarezerwowanych terminów w specjalnym formacie (np. AND, OR). W przypadku systemu NLP, procesor zapytań rozpoznaje operatorów pośrednio w języku używanym bez względu na to, jak operatorzy mogą być wyrażani (np. Przyimki, spójniki, kolejność).

**Kroki 3 i 4:** Zatrzymaj listę i wynik. Niektóre wyszukiwarki posuwają się dalej i zatrzymują listę oraz wstrzymują zapytanie, podobnie do procesów opisanych powyżej w sekcji Procesor dokumentów. Lista zatrzymań może zawierać również słowa z często występujących zapytań, np. "Chciałbym uzyskać informacje na temat". Ponieważ jednak większość publicznie dostępnych wyszukiwarek zachęca do bardzo krótkich zapytań, o czym świadczy rozmiar okna zapytania, silniki mogą opuścić te dwa kroki.

**Krok 5:** Tworzenie zapytania. Sposób, w jaki każda wyszukiwarka tworzy reprezentację zapytania, zależy od tego, jak system dopasowuje się do niego. Jeśli używany jest statystyczny matcher, zapytanie musi odpowiadać reprezentacjom statystycznym dokumentów w systemie. Dobre zapytania statystyczne powinny zawierać wiele synonimów i innych terminów, aby stworzyć pełną reprezentację. Jeśli używany jest układ logiczny Boolean, system musi utworzyć logiczne zbiory terminów powiązanych z AND, OR lub NOT.

System NLP rozpozna pojedyncze terminy, wyrażenia i nazwy jednostek. Jeśli używa jakiegokolwiek logiki Boolowskiej, rozpozna także operatory logiczne z kroku 2 i utworzy reprezentację zawierającą logiczne zestawy terminów: AND, OR, lub NOT. W tym momencie wyszukiwarka może przyjąć reprezentację zapytania i wykonać wyszukiwanie względem odwróconego pliku. Bardziej zaawansowane wyszukiwarki mogą wykonać dwa kolejne kroki.

**Krok 6:** Rozszerzenie zapytania. Ponieważ użytkownicy wyszukiwarek zwykle zawierają tylko jedno wyrażenie swoich potrzeb informacyjnych w zapytaniu, wysoce prawdopodobne jest, że informacje, których potrzebują, mogą być wyrażane za pomocą synonimów, a nie dokładnych terminów zapytań w dokumentach wyszukiwanych przez wyszukiwarkę. Dlatego bardziej wyrafinowane systemy mogą rozszerzyć zapytanie na wszystkie możliwe terminy synonimiczne, a być może nawet na szersze i węższe terminy. Proces ten zbliża się do tego, co pośrednicy wyszukiwania zrobili dla użytkowników końcowych we wcześniejszych czasach komercyjnych systemów wyszukiwania. Wówczas pośrednicy mogli używać tego samego kontrolowanego słownika lub tezauryusa używanego przez indeksatorów,

którzy przypisali deskryptory tematyczne do dokumentów. Obecnie zasoby takie jak WordNet są ogólnie dostępne, lub wyspecjalizowane obiekty ekspansji mogą wziąć początkowe zapytanie i powiększyć je poprzez dodanie powiązanego słownika.

**Krok 7:** Ważenie terminu zapytania (zakładając więcej niż jedno zapytanie). Ostatni etap przetwarzania zapytań obejmuje obliczanie wag dla terminów w zapytaniu. Czasami użytkownik kontroluje ten krok, wskazując, ile wagę ma dany termin lub po prostu, który termin lub pojęcie w zapytaniu ma największe znaczenie i musi pojawić się w każdym pobranym dokumencie, aby zapewnić trafność. Pozostawienie wagi użytkownikowi nie jest powszechne, ponieważ badania wykazały, że użytkownicy nie są szczególnie dobrzy w określaniu względnego znaczenia terminów w swoich zapytaniach. Nie mogą dokonać tego ustalenia z kilku powodów. Po pierwsze, nie wiedzą, co jeszcze istnieje w bazie danych, a terminy dokumentów są ważone przez porównanie z bazą danych jako całością. Po drugie, większość użytkowników szuka informacji o nieznanym przedmiocie, więc może nie znać poprawnej terminologii. Niewiele wyszukiwarek wdraża systemową wagę zapytań, ale niektóre dokonują niejawnego ważenia, traktując pierwszy termin (terminy) w zapytaniu jako mający większe znaczenie. Silniki wykorzystują te informacje w celu dostarczenia użytkownikowi listy dokumentów / stron. Po tym ostatnim kroku, rozszerzona, ważona kwerenda jest przeszukiwana względem odwróconego pliku dokumentów.

### **Funkcja wyszukiwania i dopasowywania**

Sposób, w jaki systemy realizują swoje funkcje wyszukiwania i dopasowywania, różni się w zależności od tego, który teoretyczny model wyszukiwania informacji leży u podstaw filozofii projektowania systemu. Ponieważ odróżnianie tych modeli wykracza daleko poza cele tego artykułu, dokonamy tylko kilku ogólnych uogólnień w poniższym opisie funkcji wyszukiwania i dopasowywania. Przeszukiwanie odwróconego pliku dokumentów spełniających wymagania zapytania, określane po prostu jako "dopasowywanie", jest zwykle standardowym wyszukiwaniem binarnym, bez względu na to, czy wyszukiwanie kończy się po pierwszych dwóch, pięciu, czy wszystkich siedmiu krokach przetwarzania zapytania. Podczas gdy przetwarzanie obliczeniowe wymagane do prostego, nieważonego, nie-boolewskiego dopasowywania zapytań jest znacznie prostsze niż wtedy, gdy model jest zapytaniem opartym na NLP w ważonym modelu Boolean, wynika również, że im prostsza jest reprezentacja dokumentu, reprezentacja zapytania i algorytm dopasowujący, przy mniej trafnych wynikach, z wyjątkiem bardzo prostych zapytań, takich jak jednowyrazowe, niejednoznaczne zapytania szukające najbardziej ogólnie znanych informacji. Po ustaleniu, który podzbiór dokumentów lub stron w pewnym stopniu odpowiada wymaganiom zapytania, wynik podobieństwa jest obliczany między zapytaniem a każdym dokumentem / stroną w oparciu o algorytm oceny stosowany przez system. Rankingi algorytmów scoringowych są oparte na obecności / nieobecności terminu (terminów) zapytania, częstotliwości terminów,  $tf / idf$ , spełnieniu logiki Boole'a lub wagach terminów zapytania. Niektóre wyszukiwarki używają algorytmów scoringowych nie opartych na treści dokumentu, ale raczej na relacjach między dokumentami lub historii pobierania dokumentów / stron. Po obliczeniu podobieństwa każdego dokumentu w podziorze dokumentów, system prezentuje uporządkowaną listę dla użytkownika. Wyrafinowanie kolejnej kolejności dokumentów zależy od modelu używanego przez system, a także od bogactwa mechanizmu ważenia dokumentów i zapytań. Na przykład wyszukiwarki wymagające tylko obecności dowolnego ciągu alfanumerycznego z kwerendy pojawiającej się w dowolnym miejscu, w dowolnej kolejności, w dokumencie, wytworzyłyby zupełnie inny ranking niż wyszukiwarka, która wykonała poprawne frazowanie językowe zarówno dla dokumentu, jak i dla dokumentu. reprezentacja zapytań i wykorzystana sprawdzony schemat ważenia  $tf / idf$ . Jednak wyszukiwarka określa rangę, lista wyników rankingowych trafia do użytkownika, który może następnie kliknąć i śledzić wewnętrzne wskaźniki systemu do wybranego dokumentu / strony. informacje o trafności lub zmodyfikować ich zapytanie na podstawie wyników, które widzieli. Jeśli

któryś z nich jest dostępny, system dostosuje swoją reprezentację zapytań, aby odzwierciedlić tę wartość dodaną, i ponownie uruchomi wyszukiwanie za pomocą ulepszonego zapytania, aby wygenerować nowy zestaw dokumentów lub prosty ponowny ranking dokumentów od wstępne wyszukiwanie.

### **Jakie funkcje dokumentu dobrze pasują do zapytania**

Omówiliśmy działanie wyszukiwarek, ale jakie cechy zapytania sprawiają, dobre dopasowania? Przyjrzyjmy się kluczowym cechom i rozważmy niektóre zalety i wady ich użyteczności, pomagając uzyskać dobrą reprezentację dokumentów / stron.

**Częstotliwość terminów:** Jak często termin zapytania pojawia się w dokumencie, jest jednym z najbardziej oczywistych sposobów określania trafności dokumentu dla zapytania. Choć najczęściej prawdziwe, kilka sytuacji może podważyć to założenie. Po pierwsze, wiele słów ma wiele znaczeń - są one polisemiczne. Pomyśl o słowach takich jak "basen" lub "ogień". Wiele nieistotnych dokumentów przedstawianych użytkownikom wynika z dopasowania właściwego słowa, ale ma niewłaściwe znaczenie. Ponadto w zbiorze dokumentów w określonej domenie, takim jak edukacja, powszechne terminy zapytań, takie jak "edukacja" lub "nauczanie", są tak powszechne i występują tak często, że zdolność silnika do rozróżniania istotnych z nieistotnych w kolekcja gwałtownie spada. Wyszukiwarki, które nie używają algorytmu ważenia tf / idf, nie odpowiednio obniżają wagi nazbyt częstych terminów, ani nie przypisują wyższych wag odpowiednim rozróżnieniu (i rzadziej występującym) terminom, np. "Wczesne dzieciństwo". Lokalizacja terminów: wiele wyszukiwarek preferuje słowa znalezione w tytule lub akapicie wiodącym lub w metadanych dokumentu. Niektóre badania pokazują, że lokalizacja - w której termin występuje w dokumencie lub na stronie - wskazuje jego znaczenie dla dokumentu. Warunki występujące w tytule dokumentu lub strony, które pasują do terminu zapytania, są dlatego często ważone bardziej niż terminy występujące w treści dokumentu. Podobnie, warunki zapytania występujące w nagłówkach sekcji lub w pierwszym akapicie dokumentu mogą być bardziej odpowiednie. te, o których wspomina wiele innych stron, lub mają dużą liczbę "odnośników"

**Popularność:** Google i kilka innych wyszukiwarek dodaje popularność do analizy linków, aby pomóc określić trafność lub wartość stron. Popularność wykorzystuje dane dotyczące częstotliwości, z jaką strona jest wybierana przez wszystkich użytkowników jako sposób przewidywania trafności. Chociaż popularność jest czasem dobrym wskaźnikiem, zakłada, że podstawowe potrzeby informacyjne pozostają takie same.

**Data publikacji:** Niektóre wyszukiwarki zakładają, że im nowsze są te informacje, tym bardziej prawdopodobne jest, że będą użyteczne lub istotne dla użytkownika. Dlatego silniki prezentują wyniki począwszy od najnowszych do mniej aktualnych.

**Długość:** chociaż długość per se niekoniecznie przewiduje trafność, to jest czynnikiem, gdy jest używana do obliczenia względnej wartości podobnych stron. Tak więc, przy wyborze między dwoma dokumentami, które zawierają te same warunki zapytania, zakłada się, że dokument, który zawiera proporcjonalnie wyższe wystąpienie terminu w stosunku do długości dokumentu, jest bardziej odpowiedni.

**Bliskość terminów zapytań:** gdy terminy w zapytaniu występują blisko siebie w dokumencie, jest bardziej prawdopodobne, że dokument jest odpowiedni dla zapytania, niż gdy terminy występują z większej odległości. Podczas gdy niektóre wyszukiwarki nie rozpoznają zwrotów per se w zapytaniach, niektóre wyszukiwarki wyraźnie rankingują dokumenty w wynikach wyższych, jeśli terminy zapytania występują obok siebie lub w pobliżu, w porównaniu do dokumentów, w których terminy występują na odległość. Rzeczowniki właściwe mają czasami większą wagę, ponieważ wiele wyszukiwań odbywa się

na ludziach, miejscach lub rzeczach. Chociaż może to być przydatne, jeśli wyszukiwarka zakłada, że szukasz nazwy zamiast tego samego słowa, co zwykłe hasło dnia codziennego, wyniki wyszukiwania mogą być szczególnie przekręcone. Wyobraź sobie, że uzyskałeś informacje na temat "Madonny", gwiazdy rocka, kiedy szukałeś zdjęć Madonny dla klasy historii sztuki.

## **Podsumowanie**

Skoro już omówiliśmy działanie wyszukiwarki, możemy omówić metody, aby z nich skorzystać. Zaczniemy od treści. Jak widać na powyższych stronach, wyszukiwarki to proste parsery testowe. Biorą szereg słów i starają się zredukować je do ich podstawowego znaczenia. Nie potrafią zrozumieć tekstu, ani nie mają zdolności rozróżniania tekstu poprawnego gramatycznie od pełnego bełkotu. To oczywiście zmieni się wraz z rozwojem wyszukiwarek, a koszt sprzętu spadnie, ale my, czarni kapelusze, będziemy ewoluować, zawsze dążąc do tego, aby pozostać przynajmniej o krok do przodu. Porozmawiajmy o podstawach generowania treści, a także o oprogramowaniu do tego używanym, ale najpierw musimy zrozumieć powieloną treść. Powszechnie podawanym mitem na forach internetowych jest to, że zduplikowana zawartość jest przeglądana przez wyszukiwarki jako procent. Dopóki pozostaniesz poniżej progu, możesz przejść bez kary. To miła myśl, po prostu szkoda, że jest całkowicie błędna.

## **Zduplikowana treść**

Czytałem na setki postów na forum omawiających duplikaty treści, z których żadna nie dawała pełnego obrazu, pozostawiając mi więcej pytań niż odpowiedzi. Postanowiłem poświęcić trochę czasu na badania, aby dowiedzieć się dokładnie, co dzieje się za kulisami. Oto, co odkryłem. Większość ludzi przyjmuje założenie, że zduplikowana treść jest analizowana na poziomie strony, podczas gdy w rzeczywistości jest znacznie bardziej złożona. Po prostu mówiąc, że "zmiana 25 procent tekstu na stronie nie jest już duplikatem treści" nie jest prawdziwym ani dokładnym stwierdzeniem. Pozwala zbadać, dlaczego tak jest. Aby uzyskać pewne zrozumienie, musimy przyjrzeć się algorytmowi k-shingle, który może, ale nie musi być używany przez główne wyszukiwarki. Widziałem następujące przykłady jako przykład, więc użyjmy go również tutaj. Załóżmy, że masz stronę zawierającą następujący tekst:

Szybki, brązowy lis przeskoczył nad leniwym psem.

Zanim przejdziemy do tego punktu, wyszukiwarka już usunęła wszystkie znaczniki i HTML ze strony, pozostawiając tylko ten zwykły tekst, abyśmy mogli się przyjrzeć. Algorytm shinglinga znajduje grupy słów w treści tekstu w celu określenia wyjątkowości tekstu. Pierwszą rzeczą, którą robią, jest pozbycie się wszystkich słów kończących, takich jak i,,,,,. Usuwają również wszystkie słowa wypełniające, pozostawiając nam tylko słowa akcji, które są uważane za rdzeń treści. Po wykonaniu tej czynności z powyższego tekstu tworzone są następujące "stwierdzenia". (Mam zamiar dołączyć słowa stop dla uproszczenia)

Szybki brązowy lis

podskoczył szybki brązowy lis

brązowy lis podskoczył

lis skoczył nad

przeskoczył przez leniwych

nad leniwym psem



Są one zasadniczo jak unikalne odciski palców, które identyfikują ten blok tekstu. Wyszukiwarka może teraz porównać ten "odcisk palca" z innymi stronami w celu znalezienia zduplikowanej treści. Po znalezieniu duplikatów strona "otrzymuje" wynik "zduplikowanej zawartości". Jeśli zbyt wiele "odcisków palców" pasuje do innych dokumentów, wynik jest na tyle wysoki, że wyszukiwarki zaznaczają stronę jako duplikat, wysyłając ją do dodatkowego piekła lub, co gorsza, całkowicie usuwając ją z indeksu.

Moja stara pani przysięga, że widziała, jak leniwy pies przeskakuje przez szybkiego lisa.

Powyższe daje nam następujące gonty:

moja stara pani przysięga

stara dama to przysięga

Pani przysięga, że ona

przysięga, że widziała

że widziała

widziała leniwych

zobaczyłem leniwego psa

leniwy skok psa

Leniwy pies przeskoczył

pies przeskoczyć

przeskocz nad szybkim

nad szybkim brązem

szybki brązowy lis

Porównując te dwa zestawy stwierdzeń widzimy, że tylko jeden pasuje ("szybki brązowy lis"). Zatem jest mało prawdopodobne, aby te dwa dokumenty były duplikatami. Nikt prócz Google nie wie, jaki procent musi być zgodność tych dwóch dokumentów, aby uznać je za duplikaty, ale niektóre dokładne testy z pewnością go zawęziły). Więc co możemy zabrać z powyższych przykładów? Przede wszystkim szybko zaczynamy zdawać sobie sprawę, że powielanie treści jest o wiele trudniejsze niż powiedzenie "dokument A i dokument B są w 50 procentach podobne". Po drugie, widzimy, że ludzie dodający słowa "zatrzymaj słowa" i "słowa wypełniające", aby uniknąć duplikowania treści, w dużej mierze marnują swój czas. To słowa "działania" powinny być przedmiotem zainteresowania. Zmiana słów akcji bez zmiany znaczenia tekstu może okazać się wystarczająca, aby ominąć te algorytmy. Z drugiej strony mogą istnieć inne mechanizmy w działaniu, których nie możemy jeszcze zobaczyć, co również jest niemożliwe. Sugeruję eksperymentowanie i znajdowanie tego, co działa dla ciebie w twojej sytuacji. Ostatni akapit tutaj jest naprawdę ważną częścią podczas generowania treści. Nie można po prostu dodawać ogólnych słów stop tu i tam i spodziewać się oszukać każdego. Pamiętajmy, że mamy tu do czynienia z algorytmem komputerowym, a nie z pewną nadprzyrodzoną mocą. Wszystko, co robisz, powinno być z punktu widzenia naukowca. Przemyśl każdą decyzję za pomocą logiki i rozumowania. W SEO nie ma magii, tylko surowe dane i liczby. Zawsze dziel test i przeprowadzaj kontrolowane eksperymenty.

**Co tworzy dobry generator treści?**

Teraz rozumiemy, jak wyszukiwarka analizuje dokumenty w Internecie, rozumiemy również zawiłości zduplikowanych treści i ich usuwanie. Nadszedł czas, aby sprawdzić niektóre podstawowe techniki generowania treści. Jednym z częściej używanych spinnerów tekstowych jest ten znany jako Markov. Markov właściwie nie jest przeznaczony do generowania treści, to tak naprawdę coś, co nazywa się łańcuchem Markowa, opracowanym przez matematyka Andreja Markowa. Algorytm przyjmuje każde słowo w treści i zmienia kolejność na podstawie algorytmu. W ten sposób powstaje w dużej mierze unikatowy tekst, ale jest on również BARDZO NIEWYKONANY. Jakość wyjścia zależy od jakości danych wejściowych. Drugą kwestią związaną z Markovem jest fakt, że najprawdopodobniej nigdy nie przejdzie on przez ludzką recenzję dla czytelności. Jeśli nie przetasujesz wystarczająco dużo łańcuchów Markowa, również natrafisz na zduplikowane problemy z treścią ze względu na naturę shinglingu, o czym była mowa wcześniej. Niektóre osoby mogą sobie z tym poradzić, zastępując słowa w treści synonimami. Niektóre popularne oprogramowanie, które używa łańcuchów Markowa to RSSGM i YAGC oba są dość stare i przestarzałe w tym momencie. Warto przyjrzeć się tylko po to, by zrozumieć podstawy, ale jest tam o wiele lepsze pakiety. Tak więc rozmawialiśmy o starych metodach robienia rzeczy, ale nie można oszukiwać wyszukiwarek po prostu powtarzając słowo kluczowe w ciele swoich stron (szkoda, że to jeszcze nie było łatwe). Co więc dzisiaj działa? Teraz i w przyszłości LSI staje się coraz ważniejsza. LSI oznacza Latent Semantic Indexing. Brzmi to skomplikowanie, ale tak naprawdę nie jest. LSI to po prostu proces, za pomocą którego wyszukiwarka może wywnioskować znaczenie strony na podstawie zawartości tej strony. Na przykład powiedzmy, że indeksują stronę i znajdują słowa takie jak bomba atomowa, Projekt Manhattan, Niemcy i Teoria względności. Chodzi o to, że wyszukiwarka może przetworzyć te słowa, znaleźć dane relacyjne i ustalić, że strona dotyczy Alberta Einsteina. Tak więc ranking słów kluczowych nie jest już tak prosty jak treść, o której mowa i powtarza docelowe wyrażenie kluczowe w kółko jak za starych dobrych czasów. Teraz musimy upewnić się, że mamy inne kluczowe frazy, które według wyszukiwarki są powiązane z główną frazą kluczową. Więc jeśli Markov jest łatwy do wykrycia, a LSI zaczyna nabierać znaczenia, które oprogramowanie działa, a które nie?

### **Oprogramowanie**

Fantomaster Shadowmaker: Jest to prawdopodobnie jeden z najstarszych i najbardziej znanych pakietów maskowania z wyższej półki. Jest to również jeden z najbardziej nieaktualnych. Za 3 000,00 \$ otrzymujesz po prostu przestarzały interfejs do powolnego budowania stron HTML. Wiem, jestem ostry, ale naprawdę byłem zawiedziony tym oprogramowaniem. Silnik treści nie robi nic, aby adresować LSI. Po prostu splota niespokrewnione zdania razem z losowych źródeł podczas losowania losowego słowa kluczowego. O ile nie zmieni się drastycznie, uniknęłbym tego.

SEC (Search Engine Cloaker): Kolejny dobrze znany płatny skrypt. Ten jest dobrej jakości, a praca zapewnia wyniki. Silnik treści jest głównie ręczny, dzięki czemu możesz budować zdania, które następnie są mieszane razem dla twoich treści. Jeśli rozumiesz SEO i masz czas poświęcić na tworzenie treści, strony budowane przez długi czas. Mam dwie uwagi. Oprogramowanie jest WOLNE. Zajmuje kilka dni, aby ustawić kilka przyzwoitych stron. To samo w sobie nie jest bardzo czarnym kapeluszem. Pamiętaj, że jesteśmy leniwi! Innym problemem jest maskowanie IP. Ich lista adresów IP jest strasznie nieaktualna i zawiera tylko kilka tysięcy adresów IP.

SSEC lub Simplified Search Engine: Jest to jeden z najlepszych systemów dostarczania IP na rynku. Ich lista adresów IP jest aktualizowana codziennie i zawiera blisko 30 000 adresów IP. Fora tylko dla członków są najlepsze w branży. Subskrypcja jest warta tylko informacji tam zawartych. Silnik zawartości jest również na najwyższym poziomie. Jest elastyczny, więc możesz użyć własnego zastrzeżonego systemu treści, który automatycznie scrapuje wyszukiwarki pod kątem twoich treści, lub możesz użyć niestandardowej zawartości podobnej do SEC, ale szybciej. Możesz także mieszać i

dopasowywać źródła treści, co zapewnia najwyższą kontrolę. Jest to jedyne oprogramowanie, które bierze pod uwagę LSI bezpośrednio z silnika treści. Jest to także najszybszy twórca stron, z którym się zetknąłem. Możesz łatwo połączyć kilka tysięcy stron, z których każda zawiera setki stron treści w ciągu zaledwie kilku godzin. Wsparcie jest na najwyższym poziomie, a kompetentny personel naprawdę wie, o czym mówią. Ten dostaje ode mnie złotą gwiazdę.

BlogSolution: sprzedawany jako automatyczny program do tworzenia blogów, BlogSolution nie działa w prawie każdym ważnym obszarze. Utworzone blogi nie są blogami typu wordpress, ale raczej autorskim oprogramowaniem blogowym specjalnie napisanym dla BlogSolution. Ta "cecha" oznacza, że Twoje blogi wyróżniają się niczym brzydki kciuk w oczach wyszukiwarek. Nie mieszają się w ogóle pozostawiając ślady stóp w każdym miejscu. Licencja ogranicza cię do 100 blogów, co oznacza, że nie możesz zbudować wystarczająco dużo, by zarobić przyzwoitą kwotę. Silnik treści to także żart, który wykorzystuje kanały rss i daje mnóstwo łatwych do wykrycia duplikatów blogów, które nie mają sobie równych.

Blog Cloaker: Kolejna solidna oferta od chłopaków, którzy opracowali SSEC. Jest to naturalna ewolucja tego oprogramowania. Ten masowy builder witryn oparty jest na blogach wordpress. To oprogramowanie jest najlepsze w branży. Interfejs ma system stworzony przez prawdziwych profesjonalistów. Dostępne są te same opcje zawartości, które są widoczne w SSEC, ale z kilkoma różnymi typami przekierowań, w tym przekierowywaniem nagłówków, JavaScriptem, odświeżaniem meta, a nawet iframe. To znowu jest rozwiązanie do maskowania IP z tą samą wiodącą na rynku listą adresów IP jak SSEC. Miesięczny abonament może wydawać się początkowo zniechęcający, ale cena przyjęcia jest warta każdego grosza, jeśli poważnie myślisz o zarabianiu w tej branży. Dosłownie nie ma nic lepszego niż to.

## **Maskowanie**

Czym więc jest maskowanie? Maskowanie pokazuje po prostu różne treści dla różnych osób na podstawie różnych kryteriów. Maskowanie automatycznie powoduje złą reputację, ale opiera się głównie na nieznajomości tego, jak to działa. Istnieje wiele uzasadnionych powodów do ukrycia stron. W rzeczywistości nawet maskuje Google'a. Czy kiedykolwiek odwiedzałeś stronę internetową za pomocą telefonu komórkowego i automatycznie przekierowywałeś ją do mobilnej wersji strony? Zgadnij co, to jest maskowanie. Co powiesz na strony internetowe, które automatycznie wyświetlają informacje na podstawie Twojej lokalizacji? Zgadnij co, to jest maskowanie. Na tej podstawie możemy podzielić maskowanie na dwie główne kategorie: maskowanie agentów użytkownika i maskowanie oparte na protokole IP. Maskowanie agentów użytkownika jest po prostu metodą pokazywania innym stronom lub innej zawartości użytkownikom w oparciu o ciąg agenta użytkownika, do którego odwiedzają witrynę. Agent użytkownika jest po prostu identyfikatorem, który każda przeglądarka internetowa i pajak wyszukiwarki wysyła do serwera WWW, gdy łączy się ze stroną. Powyżej użyliśmy przykładu telefonu komórkowego. Na przykład telefon komórkowy Nokia będzie miał agenta użytkownika podobnego do: User-Agent: Mozilla / 5.0 (SymbianOS / 9.1; U; [en]; Series60 / 3.0 NokiaE60 / 4.06.0) AppleWebKit / 413 (KHTML, podobnie jak Gecko) Safari / 413. Wiedząc o tym, możemy zauważyć różnicę między telefonem komórkowym odwiedzającym naszą stronę a zwykłym użytkownikiem przeglądającym naszą stronę, na przykład z przeglądarki Internet Explorer lub Firefox. Możemy wtedy napisać skrypt, który będzie wyświetlał różne informacje dla tych użytkowników na podstawie ich klienta użytkownika. Brzmi nieźle, prawda? Cóż, działa na podstawowe rzeczy, takie jak mobilne i niemobilne wersje stron, ale jest również bardzo łatwy do wykrycia, oszukania i obejścia. Na przykład Firefox ma poręczną wtyczkę, która pozwala zmienić ciąg znaków użytkownika użytkownika w dowolne miejsce. Korzystając z tej wtyczki, mogę sprawić, że skrypt będzie uważany za bota

wyszukiwarki Google, co spowoduje, że twoje maskowanie będzie kompletnie bezużyteczne. Co jeszcze możemy zrobić, jeśli oprogramowanie użytkownika jest tak łatwe do sfalszowania?

## **Maskowanie IP**

Każdy odwiedzający stronę internetową musi najpierw nawiązać połączenie z adresem IP. Te adresy IP rozwiązują się na serwerach DNS, które z kolei identyfikują pochodzenie tego gościa. Każdy robot wyszukiwarki musi się identyfikować za pomocą unikalnego podpisu, który można wyświetlić za pomocą wyszukiwania wstecznego dns. Oznacza to, że mamy pewną metodę fire do identyfikacji i maskowania na podstawie adresu IP. Oznacza to również, że w ogóle nie polegamy na kliencie użytkownika, więc nie ma sposobu na obejście maskowania przy użyciu protokołu IP (choć należy zachować ostrożność, ponieważ będziemy dyskutować). Najtrudniejszą częścią maskowania IP jest kompilacja listy znanych ip wyszukiwarki. Na szczęście oprogramowanie takie jak Blog Cloaker i SSEC już to robi dla nas. Po uzyskaniu tych informacji możemy następnie wyświetlać różne strony różnym użytkownikom na podstawie adresu IP, w którym odwiedzają naszą stronę. Na przykład mogę pokazać bota wyszukiwarki słowo kluczowe skierowana strona pełna kluczowych zwrotów związanych z tym, co chcę pozycjonować. Kiedy człowiek odwiedza tę samą stronę, mogę pokazać reklamę lub produkt partnerski, aby zarobić trochę pieniędzy. Zobacz moc i potencjał tutaj.

Jak więc możemy wykryć maskowanie IP? Każda większa wyszukiwarka obsługuje pamięć podręczną indeksowanych stron. Ta pamięć podręczna będzie zawierać stronę, ponieważ bot wyszukiwarki zobaczył ją w czasie indeksowania. Oznacza to, że twoja konkurencja może zobaczyć twoją zamaskowaną stronę, klikając pamięć podręczną w SERPIE. W porządku, łatwo to obejść. Użycie metadag noarchive na twoich stronach wymusza na wyszukiwarkach, aby nie wyświetlały kopii strony z pamięci podręcznej w wynikach wyszukiwania, dzięki czemu unikniesz szpiegujących webmasterów. Jedyną inną metodą wykrywania jest podszywanie się pod ip, ale jest to bardzo trudne i czasochłonne. Zasadniczo konfigurujesz komputer tak, aby działał tak, jakby korzystał z jednego z adresów IP Google podczas odwiedzania strony. To pozwoliłoby ci połączyć się tak, jakbyś był botem wyszukiwarki, ale problem polega na tym, że dane do strony były wysyłane do ip, które podszywasz, co nie jest na twoim komputerze, więc wciąż masz pecha . Ta lekcja tutaj? Jeśli poważnie myślisz o tym, użyj maskowania IP. Jest to bardzo trudne do wykrycia i zdecydowanie najmocniejsza opcja.

## **Budowanie łączy**

Jak już wcześniej mówiliśmy, czarne kapelusze to w zasadzie białe kapelusze, tylko leniwe! Kiedy budujemy strony, potrzebujemy też linków, aby uzyskać te strony w rankingu. Porozmawiajmy o typowych i nie tak powszechnych metodach.

Ping blog: ten jest dość stary, ale wciąż szeroko stosowany. Usługi indeksowania blogów konfiguruje protokół, w którym strona internetowa może wysłać polecenie ping, gdy tylko nowe strony zostaną dodane do bloga. Następnie mogą wysłać bota przechwytyjącego zawartość strony w celu indeksowania i wyszukiwania lub po prostu dodać jako link w swoim katalogu blogów. Black Hats wykorzystują to, pisząc skrypty, które wysyłają ogromną liczbę pingów do różnych serwisów, aby zachęcić roboty do przeszukiwania twoich stron. Ta metoda z pewnością napędza

Trackback: Inna metoda komunikacji używana przez blogi, trackbacks to w zasadzie metoda, w której jeden blog może powiedzieć innemu blogowi, że opublikował coś związanego z lub w odpowiedzi na istniejący blog. Jako czarny kapelusz widzimy, że jest to okazja do wprowadzenia linków do tysięcy naszych własnych stron poprzez zautomatyzowanie procesu i wysłanie trackbacków do tak wielu blogów, jak tylko możemy. Większość blogów w dzisiejszych czasach posiada oprogramowanie, które znacznie ogranicza, a nawet eliminuje spam trackback, ale wciąż jest realnym narzędziem.

Linki do EDU: Kilka lat temu Czarne Kapelusze zauważyły dziwny trend. Uniwersytety i agencje rządowe z bardzo wysokimi rankingami witryn często mają bardzo stare tablice ogłoszeń, o których dawno zapomniały, ale które wciąż mają publiczny dostęp. Skorzystaliśmy z tego, publikując miliony linków do naszych stron na tych opuszczonych stronach. To dało WIELKIE podwyższenie rankingów i sprawiło, że niektórzy sporej wielkości spamerzy Viagry zarobili miliony dolarów. Skuteczność tego podejścia zmniejszyła się z czasem.

Fora i księgi gości: Internet zawiera miliony forów i księgi gości, które są gotowe do pobrania. Podczas gdy większość forów jest mocno moderowana (przynajmniej te aktywne), to nadal pozostawia ci tysiące, w których możesz upuścić linki, których nikt nie zauważy, a nawet nie zechce. Rozmawiamy o porzuconych forach, starych księgach gości, itp. Teraz możesz także dodawać linki na aktywnych forach, ale wymaga to trochę więcej kreatywności. Umieszczanie posta związanego z tematem na forum i rzucanie linku Na przykład w kodzie BB do buźki. Pakiety oprogramowania takie jak Xrumer sprawiają, że jest to bardzo popularny sposób na zbieranie odnośników. Tak bardzo, że większość forów ma metody wykrywania i odrzucania tego typu linków. Niektóre osoby nadal z nich korzystają i nadal odnoszą sukces.

Link Networks: znane również jako farmy linków, są popularne od lat. Większość z nich ma bardzo uproszczony charakter. Strona Odsyłacze do strony B, strona B łączy do strony C, a następnie z powrotem do A. Są one dość łatwe do wykrycia z powodu ograniczonego zakresu IP. Nie trzeba wiele przetwarzania, aby dowiedzieć się, że tylko kilka osób jest zaangażowanych we wszystkie linki. Kluczem jest tutaj bardzo zróżnicowana grupa linków. Zobacz na przykład link Exchange. Mają ponad 300 serwerów na całym świecie z tysiącami adresów IP, więc byłoby to prawie niemożliwe do wykrycia. Wyszukiwarka musiałaby całkowicie zdyskontować linki w celu odfiltrowania tych linków.

### **Strategie tworzenia pieniędzy**

Mamy teraz solidne zrozumienie maskowania, jak działa wyszukiwarka, generowanie treści, oprogramowanie, którego należy unikać, oprogramowanie, które jest czystym złotem, a nawet strategie budowania linków. Więc jak połączyć to wszystko razem, aby zarobić trochę pieniędzy? Ruch, który wysyłasz. Załadujesz listę słów kluczowych związanych z pieniędzmi, utworzysz szablon z reklamami lub ofertami, a następnie wyślesz cały ruch związany z wejściami / zamaskowanymi na stronę indeksu. Narzędzie do tworzenia strony docelowej wyświetla najlepszą możliwą stronę z reklamami w oparciu o to, czego szukał użytkownik przychodzący. Nie mogą być łatwiejsze i zautomatyzowane trudne zadania, których wszyscy nienawidzimy.

Affiliate Marketing: wszyscy wiemy, czym jest program partnerski. Istnieją dosłownie dziesiątki tysięcy programów partnerskich z milionami produktów do sprzedaży. Najtrudniejszą częścią affiliate marketingu jest uzyskanie dobrze wykwalifikowanego kierowanego ruchu. To znowu jest dobre oprogramowanie i maskowanie. Niektóre sieci i podmioty stowarzyszone umożliwiają bezpośrednie łączenie. Bezpośrednie linkowanie polega na skonfigurowaniu zamaskowanych stron wraz ze wszystkimi słowami kluczowymi dotyczącymi produktów, a następnie przekierowaniu bezpośrednio do strony sprzedaży sprzedawcy lub oddziałów. Często skutkuje to najwyższymi współczynnikami konwersji, ale, jak powiedziałem, niektóre firmy stowarzyszone nie zezwalają na bezpośrednie linkowanie. To znowu miejsce, w którym wchodzi strona docelowa. Albo budujemy własne (do czego jesteśmy zbyt leniwi), albo używając czegoś takiego, jak narzędzie do tworzenia listy miejsc docelowych, które automatyzuje wszystko dla nas. Strony docelowe dają nam miejsce do wysyłania i czyszczenia naszego ruchu, a także wstępną kwalifikację kupującego i upewnienie się, że jakość ruchu przesyłanego do partnera jest jak najwyższa. W końcu chcemy zarabiać pieniądze, ale chcemy też utrzymywać silne relacje z partnerem, abyśmy mogli zarabiać.

## **Wniosek**

Jak widać, Black Hat Marketing nie różni się niczym od marketingu w White Hat. Automatyzujemy trudne i czasochłonne zadania, dzięki czemu możemy skupić się na ważnych zadaniach.