

Indeksowanie stron internetowych przez Google

Piotr Mularczyk

Aleksander Nosal

Jakub Ryś

Michał Mirek





1 803 847 241



Wyszukiwanie

- Szukając dowolnej frazy w Google, wyszukiwarka nie przeszukuje całej sieci, tylko przegląda indeksy, w których są wszystkie strony, do których udało dotrzeć się Google.
- Wyszukiwarka posiada tylko te informacje, które zostały im udostępnione
- Wyniki są szeregowane w kolejności, która ma jak najlepiej odpowiadać użytkownikowi

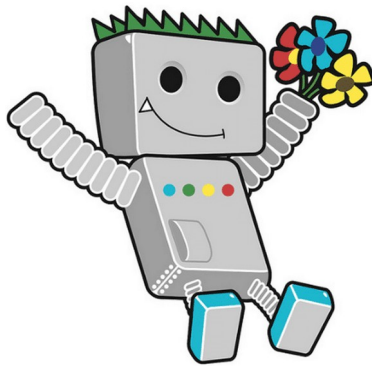


Czym jest indeksacja?

- Tytuł – w tytule najważniejsze jest pierwsze 65 znaków. Oprócz słów kluczowych tytuł warto rozwinąć o coś więcej. Ważne są także podstrony – powinny być unikalne
- Treść strony – jednym z najważniejszych elementów są treści na stronie, w tym słowa kluczowe, na ich podstawie algorytm przypisuje witrynę do odpowiedniej strony
- Linkowanie wewnętrzne – boty Google zwracają uwagę na linki do innych podstron zamieszczone na naszej stronie.

Czym jest indeksacja?

- Obrazki i atrybuty alt – Boty Google podczas przeglądania strony nie widzą, co znajduje się na zdjęciu, ale czytają opis obrazka, na podstawie którego określają co się na nim znajduje i przypisują go do odpowiedniej kategorii.
- Poprawność kodu, title, description – to również jedne z najważniejszych elementów, na jakie zwracają uwagę boty.





Dlaczego linki są ważne

Wszelkie napotkane odnośniki przegląda GoogleBot, który jest crawlerem - programem pobierającym informacje ze stron internetowych.



Jak działa wyszukiwanie

Oprogramowanie google przeszukuje indeks i znajduje wszystkie strony zawierające te hasła. W tym przypadku mamy tysiące możliwych wyników. Jak Google wybiera dokumenty, które naprawdę zawierają przydatne informacje?

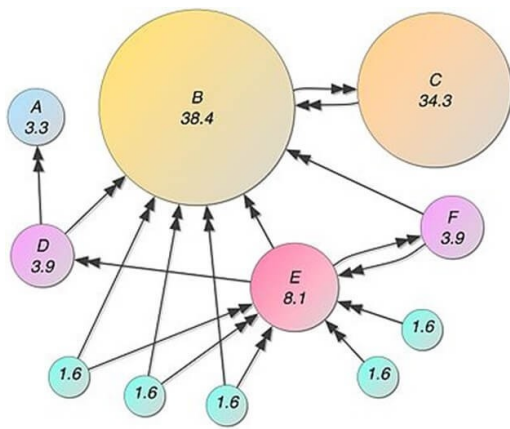


Zadaje pytania. Jest ich ponad 200:

- Ile razy słowa kluczowe pojawiają się na tej stronie?
- Czy te słowa pojawiają się w tytule, URL-u, obok siebie?
- Czy na stronie są synonimy tych słów?
- Czy strona znajduje się w cenionej witrynie?
- A może witryna jest niskiej jakości albo nawet spamuje?
- Jaki jest PageRank tej strony?

Czym jest PageRank?

Metoda nadawania indeksowanym stronom internetowym określonej wartości liczbowej, oznaczającej ich jakość.



$$PR_x = \frac{1 - d}{N} + d \left(\frac{PR_y}{L_y} + \frac{PR_z}{L_z} \dots \right),$$

gdzie:

- PR – PageRank danej strony
- d – współczynnik tłumienia, liczba pomiędzy 0 i 1. Dla obliczeń przyjmuje się zazwyczaj wartość 0,85
- N – liczba stron internetowych
- L – liczba linków do których odsyła dana strona internetowa



- Ocenia ważność strony internetowej na podstawie liczby prowadzących do niej linków i ważności samych linków
- Google łączy wszystkie zebrane informacje, ustala ogólną pozycję strony i pokazuje wyniki wyszukiwania mniej więcej pół sekundy po wpisaniu zapytania.



Search Console

- Google oferuje dostęp do Search Console, by właściciele witryn mogli szczegółowo wskazać, jak wyszukiwarka ma indeksować ich strony internetowe.
- Google nigdy nie pobiera opłat za częstsze indeksowanie witryny – udostępnia te same narzędzia wszystkim stronom internetowym, by zapewnić naszym użytkownikom jak najlepsze wyniki wyszukiwania.



Proces indeksacji

- Czas indeksowania wynosi od kilku dni do kilku tygodni.
- Proces indeksacji strony nie oznacza, że roboty Google będą umieszczać naszą stronę na wysokich pozycjach w wynikach wyszukiwania.
- Zaindeksowana witryna może znajdować się zarówno na 1. pozycji w Google, jak i na 50.



Indeksacja, a ranking

- Indeksowanie polega tylko na dostarczeniu informacji do silnika.
- Następnie wystawiana jest ocena i szeregowanie wyników.
- Bardzo ważne jest pozycjonowanie stron.

GoogleBot Desktop i GoogleBot Mobile





GoogleBot Desktop i GoogleBot Mobile

- Rynek smartfonów i innych urządzeń mobilnych odpowiada za 56% ruchu w sieci.
- Google stworzyło dwa boty które odwiedzają strony:
 - GoogleBot Desktop symuluje zachowanie użytkownika na komputerze
 - GoogleBot Mobile symuluje przeglądanie strony przez urządzenie mobilne



GoogleBot Desktop i GoogleBot Mobile

- Google obecnie prowadzi dwa indeksy. Jeden dla wersji komputerowej i drugi dla wersji mobilnej.
- Podczas rankingowania analizowane są nieco inne parametry
 - Rozmiar grafiki ma zdecydowanie większe znaczenie na stronach mobilnych niż na wariantach desktopowych.



Dziękujemy za uwagę!