

Indeksowanie stron internetowych przez Google

Konrad Synowiec
Michał Soboń

Czym jest i co ułatwia indeksowanie stron internetowych?

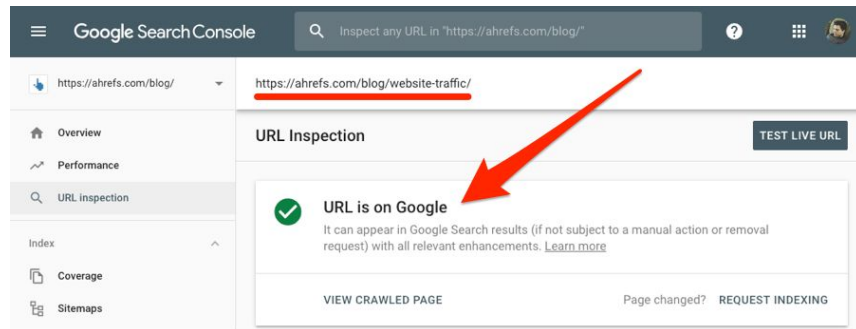
- Indeksowanie (ang. indexing) – Indeksowanie to zbieranie informacji o treści, frazach kluczowych, linkach i obrazach znajdujących się na stronie i zapisywanie ich w bazie danych wyszukiwarki.
- Korzystając z wyszukiwania, przeszukujemy właśnie tę bazę danych (indeks), a nie wszystkie witryny w internecie bezpośrednio.
- Za indeksowanie strony odpowiadają specjalne programy, czyli roboty sieciowe.

Indeksowanie ułatwia:

- Porządkowanie treści w Internecie
- Dopasowanie wyniku do wyszukiwania
- Zapewnienie jakości wyników wyszukiwania

Crawling

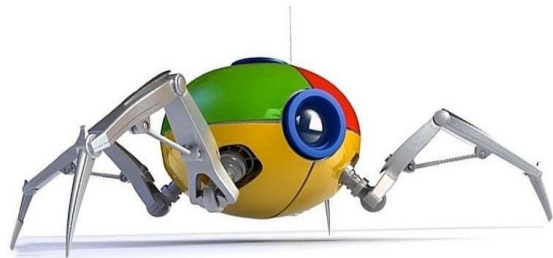
Aby wyszukiwarka mogła działać, najpierw musi pozyskać dane, na których będzie przeprowadzone wyszukiwanie. Pozyskiwaniem danych zajmuje się Googlebot, który co jakiś czas odwiedza strony internetowe pobierając z nich informacje. Jako, że nie istnieje żaden centralny rejestr wszystkich stron internetowych na świecie, bot musi analizować na przykład linki umieszczane na znanych mu stronach aby odkrywać nowe witryny. Twórca strony może też poprosić Googlebota, aby ten odwiedził jego stronę.



The screenshot displays the Google Search Console interface. The top navigation bar includes the Google Search Console logo and a search bar with the text "Inspect any URL in 'https://ahrefs.com/blog/'". The main content area shows the URL "https://ahrefs.com/blog/website-traffic/" under the "URL Inspection" section. A red arrow points to a green checkmark icon next to the text "URL is on Google". Below this, a message states: "It can appear in Google Search results (if not subject to a manual action or removal request) with all relevant enhancements. [Learn more](#)". At the bottom of the inspection card, there are three buttons: "VIEW CRAWLED PAGE", "Page changed?", and "REQUEST INDEXING". The left sidebar contains navigation options: Overview, Performance, URL inspection, Index, Coverage, and Sitemaps.

Indeksowanie

Gdy robot znajdzie nową stronę zostaje poddana ona analizie. Zapisywane są jej słowa kluczowe, to jak bardzo jest aktualna, czy jest dostosowana do urządzeń mobilnych, jak długo się ładuje i tak dalej. Indeks wyszukiwarki Google zawiera setki miliardów stron internetowych. Internet cały czas się zmienia, Googlebot musi więc regularnie odświeżać swoją bazę danych, cały czas się uczy, dobierając odpowiedni interwał skanowania dla każdej z indeksowanych stron.



```
66.249.79.55 - - [02/Dec/2022:22:33:32 +0000] "GET /robots.txt HTTP/1.1" 404 207 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.79.55 - - [02/Dec/2022:22:33:32 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/107.0.5304.110 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.79.45 - - [02/Dec/2022:22:37:38 +0000] "GET /static/css/bootstrap.min.css HTTP/1.1" 200 159515 "https://aioner.xyz/" "Mozilla/5.0 (Linux; Android 6.0.1;
Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/107.0.5304.115 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.75.45 - - [03/Dec/2022:13:22:15 +0000] "GET /robots.txt HTTP/1.1" 404 207 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.75.45 - - [03/Dec/2022:13:22:15 +0000] "GET / HTTP/1.1" 200 9619 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.47 - - [05/Dec/2022:01:22:41 +0000] "GET /robots.txt HTTP/1.1" 404 207 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.49 - - [05/Dec/2022:01:22:42 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.45 - - [05/Dec/2022:20:10:07 +0000] "GET /robots.txt HTTP/1.1" 404 207 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.49 - - [05/Dec/2022:20:10:07 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/107.0.5304.110 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.45 - - [05/Dec/2022:20:11:07 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.76.168 - - [05/Dec/2022:20:11:07 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/99.0.4844.84 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.47 - - [05/Dec/2022:20:11:08 +0000] "GET /favicon.ico HTTP/1.1" 404 207 "-" "Googlebot-Image/1.0"
66.249.66.47 - - [05/Dec/2022:20:11:11 +0000] "GET /favicon.ico HTTP/1.1" 404 207 "-" "Googlebot-Image/1.0"
66.249.66.45 - - [06/Dec/2022:04:17:01 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/107.0.5304.110 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.47 - - [06/Dec/2022:05:17:01 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/107.0.5304.110 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.49 - - [06/Dec/2022:13:07:31 +0000] "GET /robots.txt HTTP/1.1" 404 207 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.47 - - [06/Dec/2022:13:07:31 +0000] "GET /commands/ HTTP/1.1" 200 13198 "-" "Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36
(KHTML, like Gecko) Chrome/107.0.5304.110 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.49 - - [06/Dec/2022:13:09:01 +0000] "GET /static/css/commands.css HTTP/1.1" 200 1169 "https://aioner.xyz/commands/" "Mozilla/5.0 (Linux; Android 6.0.1;
Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/107.0.5304.115 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.49 - - [06/Dec/2022:13:17:00 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/107.0.5304.110 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.66.45 - - [06/Dec/2022:14:17:01 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/107.0.5304.110 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.64.241 - - [08/Dec/2022:21:55:06 +0000] "GET /robots.txt HTTP/1.1" 404 207 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.64.241 - - [08/Dec/2022:21:55:06 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (Linux; Android 6.0.1; Nexus 5X Build/MMB29P) AppleWebKit/537.36 (KHTML,
like Gecko) Chrome/107.0.5304.115 Mobile Safari/537.36 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.64.245 - - [09/Dec/2022:01:28:39 +0000] "GET / HTTP/1.1" 200 9620 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
66.249.64.245 - - [09/Dec/2022:01:30:32 +0000] "GET /static/css/bootstrap.min.css HTTP/1.1" 200 159515 "https://aioner.xyz/" "Mozilla/5.0 AppleWebKit/537.36 (KHTML,
like Gecko; compatible; Googlebot/2.1; +http://www.google.com/bot.html) Chrome/107.0.5304.115 Safari/537.36"
66.249.64.241 - - [09/Dec/2022:01:30:32 +0000] "GET /static/js/popper.min.js HTTP/1.1" 200 21257 "https://aioner.xyz/" "Mozilla/5.0 AppleWebKit/537.36 (KHTML, like
Gecko; compatible; Googlebot/2.1; +http://www.google.com/bot.html) Chrome/107.0.5304.115 Safari/537.36"
```

Czynniki wpływające na końcowe wyświetlone wyniki

- Znaczenie zapytania
- Trafność treści
- Jakość treści
- Użyteczność treści
- Kontekst i ustawienia

Znaczenie zapytania

Aby wyświetlić odpowiednie wyniki wyszukiwania, trzeba najpierw ustalić, jakie informacje są szukane, czyli co się kryje za danym zapytaniem. W tym celu tworzone są modele językowe, który pomagają określić, jak tych kilka słów wpisanych w polu wyszukiwania ma się do najbardziej przydatnych dostępnych treści.

Pomimo wieloznaczności niektórych słów w konkretnych językach, algorytm potrafi rozpoznać ich zastosowanie w zapytaniu.

Replace

Q how to **change** a light bulb|



Exchange

Q does post office **change** foreign currency|



Adjust

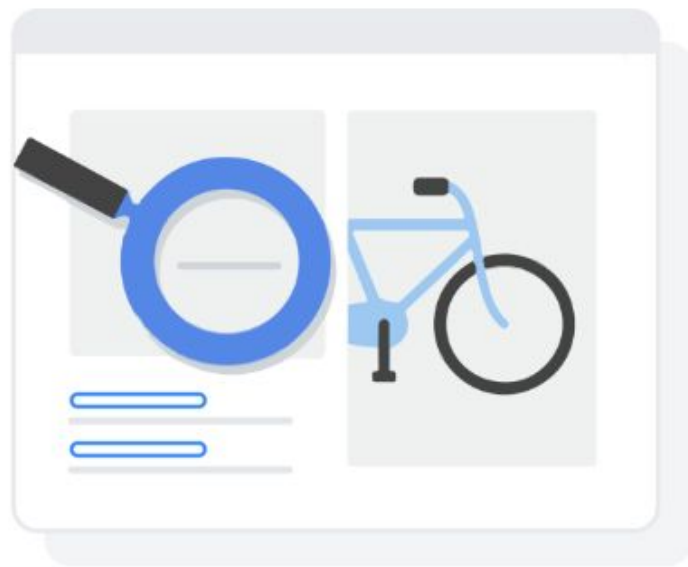
Q how to **change** brightness on a laptop|



Trafność treści

Podstawowym sygnałem trafności informacji jest obecność na stronie internetowej tych samych słów kluczowych co w wyszukiwanym haśle. Jeśli te słowa kluczowe pojawiają się na stronie, w nagłówkach lub w treści, informacje najprawdopodobniej są trafne.

Gdy wyszukiwane jest hasło „psy”, prawdopodobnie nie chodzi o strony, na której słowo to jest powtarzane setki razy. Algorytmy wyszukiwarki analizują więc, czy dana strona zawiera inne trafne zasoby oprócz słowa kluczowego „psy”, np. zdjęcia psów, filmy czy listę ras.



Jakość treści

Po rozpoznaniu trafności, systemy Google starają się ocenić jakość prezentowanych informacji, tak, by zaprezentować użytkownikowi jak najbardziej rzetelne i wiarygodne materiały.

Jednym z czynników branych pod uwagę podczas oceny są na przykład linki do innych witryn. Jeżeli na stronie znajdują się odniesienia do stron, które zostały wcześniej uznane za wiarygodne, to istnieje duże prawdopodobieństwo, że prezentowane treści mają podobny poziom wiarygodności.

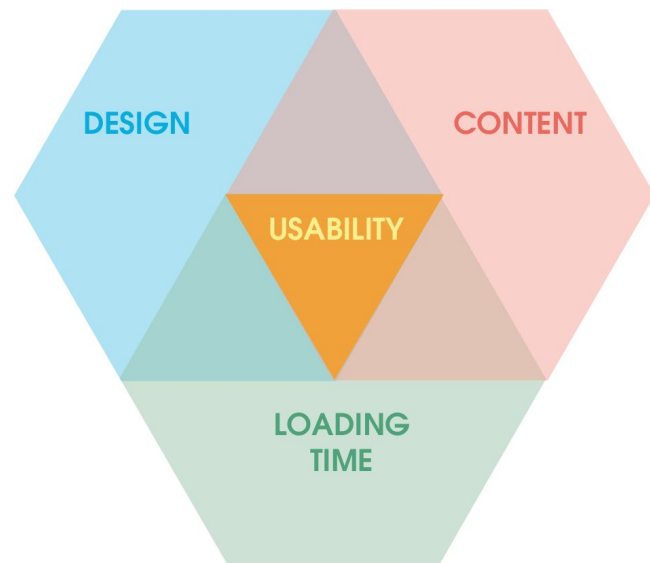
Internet nieustannie się zmienia, dlatego witryny muszą być poddawane ciągłej ewaluacji.



Użyteczność treści

Branża jest również pod uwagę użyteczność treści. Gdy wszystko inne jest na relatywnie podobnym poziomie, treści łatwiej dostępne mogą okazać się bardziej atrakcyjne.

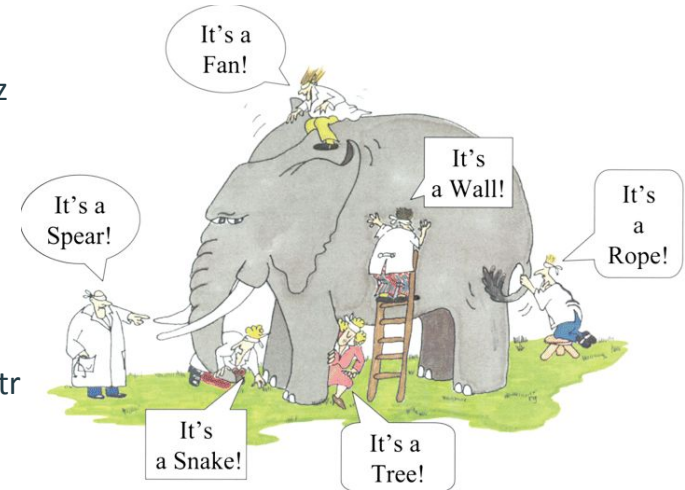
Zwraca się uwagę na szybkość wczytywania treści oraz dostosowanie do urządzeń mobilnych w przypadku korzystania z telefonu



Kontekst i ustawienia

Podczas wyświetlania wyników pod uwagę brana jest historia wyszukiwań użytkownika, a także informacje jakiego języka używa oraz w jakim miejscu się znajduje. Przykładowo gdy jesteśmy w Krakowie i wyszukamy słowo “koncert” otrzymamy informacje o biletach i wydarzeniach w okolicy Krakowa, natomiast gdy tę samą frazę wyszukamy będąc w Paryżu, to dostaniemy informacje na temat koncertów w tym mieście.

Ważne są też ustawienia wyszukiwania, na przykład jeżeli włączymy filtr SafeSearch, Google zwróci nam wyniki, w których będą odfiltrowane nieodpowiednie treści.



Optymalizacja dla wyszukiwarek internetowych

SEO (ang. search engine optimization) - to proces optymalizacji stron internetowych, którego celem jest podniesienie pozycji witryny w wynikach wyszukiwania, aby zwiększyć jej popularność, wiarygodność i rozpoznawalność.



Źródła

- [Porządkowanie informacji – jak działa wyszukiwarka Google](#)
- [Wyszukiwarka Google – odkryj, jak działa wyszukiwarka Google](#)
- [Indeksowanie strony: definicja i podstawy - Ster Na SEO](#)

Dziękujemy za uwagę