

Skuteczne zarządzanie treścią

*Przewodnik dla właścicieli
witryn internetowych*

Google™

Spis treści

Wprowadzenie	2
Krótkie omówienie metod wyszukiwania w sieci	3
Co nowego w wyszukiwarce Google?	4
Czy Google znajdzie Twoją witrynę?	5
Czy Google indeksuje Twoją witrynę?	6
<i>Kontrolowanie elementów indeksowanych przez Google</i>	7
<i>Robots.txt a tagi meta</i>	9
<i>Kontrolowanie kopii strony i fragmentów tekstu wyświetlanych w wynikach wyszukiwania</i>	10
Czy Twoja witryna zawiera oryginalne i użyteczne treści?	11
Zwiększanie widoczności: sprawdzone metody	12
Webmaster Central	13
<i>Sitemaps</i>	14
Najczęściej zadawane pytania	15
Słowniczek	19

Wprowadzenie

Jeśli chcesz żeby Cię zauważono, internet jest do tego najlepszym miejscem. Wystarczy porozmawiać z reklamodawcą, który zwiększył sprzedaż za pomocą usług online, autorem bloga, którego popularność doprowadziła do wydania książki lub wydawcą gazety, który dzięki wzmożonej liczbie odwiedzin rozszerzył zakres swojego pisma na skalę międzynarodową.

Użytkownicy Google bardzo często pytają, jak działa wyszukiwanie w sieci i jak wydawcy sieciowi zwiększają swoją widoczność w internecie.

Zdecydowaliśmy się opublikować tę niewielką książeczkę, aby pokazać, jak wyszukiwarka „widzi” publikowane treści i jak najlepiej dostosować swoją obecność w sieci tak, aby użytkownicy mogli znaleźć dokładnie to, na czym zależy wydawcy, a cała reszta pozowała w ukryciu.

Niniejsza książeczka zawiera porady dla webmasterów, omówienie narzędzi sieciowych, szczegółowe odpowiedzi na najczęściej zadawane pytania i jest przeznaczona tak dla małych wydawców sieciowych, jak i dla właścicieli dużych witryn.

Wraz z oszałamiającym rozwojem internetu w ostatniej dekadzie zmieniło się także podejście Google do wyszukiwania danych w sieci oraz relacje firmy z właścicielami witryn. Stworzyliśmy szereg narzędzi ułatwiających webmasterom zwiększenie widoczności ich treści, a także kontrolowanie sposobu indeksowania ich stron internetowych. Nie spoczywamy jednak na laurach. Mamy nadzieję, że, zachęceni niniejszą książeczką, zechcą Państwo podzielić się z nami swoimi opiniami i dadzą nam znać, co możemy zrobić, by uczynić sieć jeszcze przyjaźniejszym miejscem dla poszukiwaczy informacji i wydawców

- Zespół Google ds. webmasterów

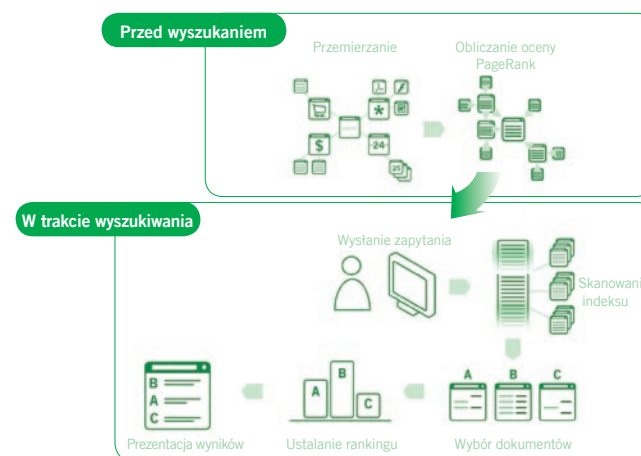
Krótkie omówienie metod wyszukiwania w sieci: Sposób działania

Ujmując rzecz najprościej jak się da, można postrzegać sieć jako ogromną książkę z imponującym indeksem, szczegółowo określającym lokalizację poszczególnych informacji.

Grupa specjalnych komputerów Google (tzw. „Googleboty”) stale „przemierza” (przegląda) miliardy stron sieci. Proces ten jest podporządkowany algorytmom, tj. specjalny program komputerowy określa, które witryny są przemierzane, jak często i jak głęboko. Nie akceptujemy opłat za częstsze przemierzanie witryn i wszelka działalność naszej firmy związana z wyszukiwaniem jest całkowicie oddzielona od dochodowej usługi AdWords.

Proces przemierzania witryn zaczyna się od listy adresów URL stron internetowych. Googlebot przegląda te witryny, wykrywając na poszczególnych stronach linki i dodając je do listy stron przeznaczonych do przemierzenia. Każda przemierzona strona jest kopiowana, w wyniku czego tworzony jest ogromny indeks wszystkich słów napotkanych przez Googlebota. W indeksie znajdują się także informacje na temat lokalizacji poszczególnych słów na poszczególnych stronach.

Gdy użytkownik zgłasza kwerendę, nasze komputery szukają w indeksie odpowiednich stron i zwracają mu najbardziej trafne wyniki. Stopień trafności jest określany na podstawie ponad 200 czynników, z których jednym jest ocena „PageRank” danej strony. Ocena ta jest miarą „ważności” strony w oparciu o linki z innych stron. Krótko mówiąc, każdy link prowadzący do strony XYZ poprawia jej PageRank.



Co nowego w wyszukiwarce Google?

Choć fundamentalne zasady wyszukiwania danych w sieci zasadniczo się nie zmieniły, Google stale udoskonala wyniki swojej wyszukiwarki.

Czym zatem różni się wyszukiwanie dziś od tego co mieliśmy, powiedzmy, pięć lat temu? Jest przede wszystkim dużo szybsze.

Ponadto pięć lat temu nasze systemy przemierzania i indeksowania były znacznie mniej inteligentne. Na przykład dziś przeglądamy strony w sposób ciągły, a wizyty planujemy znacznie lepiej tak, aby zapewnić użytkownikom najświeższe wyniki. Wydajniejsze podejście uwzględnia fakt, że np. witryna gazety wymaga częstszego przemierzania niż statyczne witryny aktualizowane raz w miesiącu. Właściwie pozwalamy webmasterom kontrolować częstotliwość przemierzania ich witryn za pomocą specjalnych narzędzi. Ogólnie rzecz biorąc, wynikiem naszych starań jest świeższy i pełniejszy indeks.

Choć dzisiejsze wyszukiwanie danych odbywa się szybciej i wydajniej niż kiedykolwiek w przeszłości, to jednak od samego początku wyszukiwarki Google najważniejszym czynnikiem była widoczność w jej wynikach wyszukiwania:

[Czy Google znajdzie Twoją witrynę?](#) (strona 5)

[Czy Google indeksuje Twoją witrynę?](#) (strona 6)

[Czy Twoja witryna zawiera oryginalne i użyteczne treści?](#) (strona 11)

Czy Google znajdzie Twoją witrynę?

Uwzględnienie w wynikach wyszukiwania Google jest bezpłatne i łatwe; nie trzeba nawet zgłaszać witryny do Google. Właściwie ogromna większość witryn pojawiających się w naszych wynikach nie jest zgłaszana ręcznie, ale wyszukiwana i automatycznie dodawana podczas ich przemierzania przez Googleboty.

Choć Google przemierza miliardy stron, nie sposób przemierzyć wszystkie. Gdy Googlebot ominie jakąś witrynę, dzieje się to często z następujących powodów:

- witryna nie jest dobrze połączona wieloma linkami z innymi witrynami w internecie.
- witryna została udostępniona po zakończeniu ostatniego indeksowania internetu przez Googleboty.
- podczas próby przemierzania witryna była przejściowo niedostępna lub wystąpił błąd.

Za pomocą narzędzi Google dla webmasterów, np. Sitemaps, można określić, czy dana witryna znajduje się w indeksie Google i czy podczas próby jej przemierzenia wystąpił błąd (zob. str. 14). Za pomocą tych narzędzi można także ręcznie dodać swój adres URL do indeksu Google lub dostarczyć nam mapę witryny, która lepiej odzwierciedla jej zawartość. Dzięki temu będziemy mogli znaleźć nowe sekcje witryny i nową treść.

Czy Google indeksuje Twoją witrynę?

Od czasu do czasu zdarza się, że witryna nie pojawia się w wynikach wyszukiwania. Problem ten może wynikać z „indeksowalności” strony, tj. z możliwości jej skopiowania przez Googlebota w celu umieszczenia w wynikach wyszukiwania.

Struktura i treść

Częstym powodem nieobecności jakiejś strony internetowej w wynikach wyszukiwania jest jej struktura i treść. Google może na przykład nie indeksować stron wymagających od użytkownika wypełnienia formularza. To samo dotyczy stron zawierających „treści dynamiczne” (np. Flash, JavaScript, ramki czy dynamicznie generowane adresy URL). Jeśli nie ma pewności, co do przyczyny nieobecności witryny w wynikach wyszukiwania, warto wyświetlić ją w przeglądarce tekstowej, np. Lynx lub w normalnej przeglądarce z wyłączoną obsługą grafiki, Javascript i Flash. W ten sposób można szybko ustalić, czy wszystkie treści są dostępne.

Jeśli w witrynie jest dużo grafiki, warto upewnić się, że treść każdego zdjęcia jest także opisana. Dzięki temu wyszukiwarka może je prawidłowo indeksować, ale także udostępniać użytkownikom z upośledzeniem wzroku. Grafikę można także opisać tekstem alt i używać opisowych nazw plików, jak w przykładzie poniżej. Jest to logo firmy „Cukiernia u Jadzi”:

```

```

Adresy URL

Dodatkową przeszkodą może być sam adres URL. Jeśli zawiera on identyfikatory sesji lub inne parametry, albo kilkakrotnie przekierowuje użytkownika, taka strona może nie być indeksowana przez Google.

Serwer i sieć

Dostęp do niektórych stron danej witryny może być niemożliwy ze względu na problemy z serwerem lub siecią. Listę stron niedostępnych dla Googlebotów można wyświetlić za pomocą narzędzi dostępnych na stronie Webmaster Central firmy Google. Więcej informacji na temat Webmaster Central znajduje się na str. 13.

Protokół wykluczenia robotów

Niektóre strony są blokowane przez Protokół wykluczenia robotów — normę techniczną, dzięki której wydawcy mogą blokować indeksowanie wybranych treści swoich witryn (zobacz str. 7). Jeśli witryna nie pojawia się w wynikach wyszukiwania Google, warto sprawdzić, czy plik robots.txt lub tag meta nie blokuje dostępu naszym robotom indeksującym.

Kontrolowanie elementów indeksowanych przez Google

Każdy wydawca internetowy stawia sobie inne cele. Na przykład niektórzy wydawcy gazet udostępniają najnowsze artykuły bezpłatnie, ale dostęp do archiwum jest już płatny. Niektórym zależy na widoczności w poszczególnych modułach wyszukiwarki (np. wyszukiwarce grafiki, Google Mobile, itp.), podczas gdy innym zależy na widoczności w ogólnych wynikach wyszukiwania.

Staramy się szanować życzenia wydawców — w końcu to ich materiały — ale nie jesteśmy jasnowidzami. Wydawcy muszą więc wyraźnie określać sposób indeksowania publikowanych treści. Służy do tego Protokół wykluczenia robotów — uznana specyfikacja techniczna informująca wyszukiwarki, które witryny lub ich fragmenty mają być pomijane, a które mają pojawiać się w wynikach wyszukiwania.

Robots.txt: kontrola w całej witrynie

Głównym elementem Protokołu wyłączenia robotów jest prosty plik tekstowy zwany robots.txt, od lat będący standardem w internecie. Plik robots.txt pozwala kontrolować dostęp na wielu poziomach, od całej witryny po poszczególne katalogi, strony określonego typu lub nawet indywidualne strony.

Nie chcę, aby niektóre strony mojej witryny znajdowały się w indeksie Google. W jaki sposób mogę zapobiec ich wyświetlaniu w wynikach wyszukiwania Google?

Generalnie większość właścicieli witryn nie blokuje dostępu do swoich stron dla Googlebotów, aby użytkownicy mieli jak największy dostęp do publikowanych treści. Są jednak strony, które nie powinny być uwzględniane w indeksie np. wewnętrzne dzienniki lub artykuły udostępniane odpłatnie.

Można wyłączyć strony z indeksu Google tworząc plik robots.txt i umieszczając go w katalogu głównym na swoim serwerze sieciowym. W pliku robots.txt znajdują się informacje na temat stron, które nie powinny być indeksowane przez wyszukiwarki. Tworzenie pliku robots.txt jest proste i daje wydawcom zaawansowany poziom kontroli dostępu wyszukiwarek do ich witryn.

Jeśli na przykład webmaster nie chce indeksować wewnętrznych dzienników, plik robots.txt powinien zawierać następujące wpisy:

User-Agent: Googlebot – wiersz User-Agent określa, że następna sekcja instrukcji jest przeznaczona tylko dla Googlebota.

Disallow: /dzienniki/ – wiersz Disallow informuje Googlebota, aby nie otwierał plików w podkatalogu dziennika witryny

Właściciel witryny określił, że żadna strona w katalogu dziennika nie powinna pojawiać się w wynikach wyszukiwania Google.

Informacje zawarte w pliku robots.txt będą zrozumiałe dla wszystkich głównych wyszukiwarek. Istnieje też możliwość definiowania osobnych zasad dla poszczególnych wyszukiwarek.

Tagi Meta: kontrola szczegółowa

Oprócz pliku robots.txt, który pozwala krótko formułować instrukcje dotyczące dużej liczby plików umieszczonych w witrynie, dostępne są także tagi meta umożliwiające bardziej szczegółową kontrolę dostępu do poszczególnych stron. Aby zastosować tę metodę, wystarczy dodać określone tagi meta do strony HTML definiujące sposób jej indeksowania. Obie metody — pliki robots.txt i tagi meta — zapewniają duży stopień elastyczności i pozwalają stosunkowo łatwo implementować zasady dostępu.

W mojej witrynie znajduje się artykuł, który jest dostępny tylko dla użytkowników zarejestrowanych. W jaki sposób mogę wykluczyć go z wyników wyszukiwania Google?

W tym celu wystarczy dodać tag meta NOINDEX w pierwszej sekcji <head> artykułu. Powinno to wyglądać mniej więcej tak:

```
<html>
<head>
<meta name="googlebot" content="noindex">
[...]
```

Google nie będzie indeksować tego pliku.

Warto jednak pamiętać, że w niektórych przypadkach indeksowanie tego typu stron może być pożądane — np. w przypadku płatnych artykułów archiwalnych dostępnych online. Choć treści tego typu nie będą wyświetlane w wynikach wyszukiwania Google, niektóre usługi Google, takie jak Wyszukiwanie w archiwum Google News, będą uwzględniać je w swoich indeksach z wyraźną informacją o opłatach za dostęp.

Robots.txt a tagi meta

Generalnie pliki robots.txt pozwalają kontrolować całą witrynę, a tagi meta poszczególne jej elementy. Tagi meta są więc szczególnie przydatne, gdy nie mamy dostępu do edycji całej witryny, a tylko do poszczególnych plików. Tagi meta pozwalają także określić złożone zasady kontroli dostępu do poszczególnych stron.

Niekiedy ten sam problem można rozwiązać za pomocą obu narzędzi:

Jak mogę zapewnić, że indeksowany jest tylko tekst strony, a nie jej grafika

Można to osiągnąć, blokując dostęp do określonego typu plików graficznych w całej witrynie za pomocą pliku robots.txt. Następujące wiersze w pliku robots.txt informują Google, aby nie indeksować żadnych plików z rozszerzeniem *.jpg ani *.jpeg:

```
User-agent:   Googlebot
Disallow:     /*.jpg$
Disallow:     /*.jpeg$
```

Jeśli system zarządzania treścią (CMS) przechowuje grafikę w oddzielnym katalogu, można też wykluczyć cały katalog. Jeśli obrazy znajdują się w katalogu /obrazy/, można go wykluczyć w następujący sposób:

```
User-agent:   *
Disallow:     /obrazy/
```

Innym sposobem jest dodanie tagu NOINDEX do każdego pliku zawierającego grafikę. Wszystkie opisane wyżej metody pozwalają zablokować indeksowanie grafiki; jedyną różnicą jest tu zakres wykluczenia.

Kontrolowanie kopii strony i fragmentów tekstu wyświetlanych w wynikach wyszukiwania

Wyniki wyszukiwania zwykle zawierają link do kopii strony oraz fragment zawartego na niej tekstu. Oto przykładowy pierwszy wynik wyszukiwania hasła "Kaczka".

Kaczka krzyżówka - Wikipedia, wolna encyklopedia

Krzyżówka (**kaczka** krzyżówka) (*Anas platyrhynchos*) - średniej wielkości wędrowny ptak wodny z rodziny kaczkowatych, zamieszkujący Europę, Amerykę Północną i ...
pl.wikipedia.org/wiki/Kaczka_krzyżówka - 27k - [Kopia](#) - [Podobne strony](#)

Fragment tekstu – tekst występujący na stronie internetowej.

Kopia w pamięci podręcznej – link umożliwiający wyświetlenie kopii strony przechowywanej na serwerach Google.

Do czego służy fragment tekstu? Użytkownicy chętniej odwiedzą jakąś witrynę, jeśli w wynikach wyszukiwania będą mogli przeczytać fragment jej tekstu. Dzięki temu użytkownik może znacznie łatwiej określić trafność wyświetlonych wyników. Jeśli użytkownik nie jest w stanie szybko tego określić, zwykle rozpoczyna nowe wyszukiwanie.

Do czego służy link do kopii strony? Link do kopii strony przydaje się w wielu sytuacjach, np. gdy witryna jest tymczasowo niedostępna; gdy nastąpi przeładowanie witryny z wiadomościami w wyniku jakiegoś zdarzenia o dużym znaczeniu; lub gdy witryna zostanie przypadkowo usunięta. Kolejną zaletą kopii strony jest podkreślanie wyszukiwanych słów, dzięki czemu użytkownik może szybko ocenić trafność wyświetlonej strony pod względem wyszukiwanego hasła.

Większość wydawców internetowych chce, by w wynikach wyszukiwania wyświetlany był zarówno fragment tekstu, jak i link do kopii strony. W niektórych jednak przypadkach może to być niepożądane:

Moja gazeta jest aktualizowana kilka razy dziennie. Częstotliwość przemierzania nie nadąża za aktualizacjami, przez co link do kopii strony w wynikach wyszukiwania prowadzi do nieaktualnych informacji. Jak mogę zapobiec wyświetlaniu linku do kopii strony w tym przypadku?

Właściciel witryny publikującej wiadomości może zapobiec wyświetlaniu linku do kopii strony w wynikach wyszukiwania poprzez dodanie do strony tagu NOARCHIVE:

```
<META NAME="GOOGLEBOT" CONTENT="NOARCHIVE">
```

Podobnie można zrezygnować z wyświetlania fragmentu tekstu strony, za pomocą tagu NOSNIPPET:

```
<META NAME="GOOGLEBOT" CONTENT="NOSNIPPET">
```

Uwaga: dodanie tagu NOSNIPPET uniemożliwia także wyświetlanie linku do kopii strony.

Czy Twoja witryna zawiera oryginalne i użyteczne treści?

Gdy możliwe jest wykrycie i zindeksowanie witryny, trzeba sobie jeszcze odpowiedzieć na pytanie, czy treść jej stron jest oryginalna i użyteczna.

Najpierw przeanalizujmy tekst jako całość. Czy tytuł i linki tekstowe są wystarczająco opisowe? Czy tekst jest płynny, zrozumiały i intuicyjny?

Tak jak rozdział w książce jest uporządkowany wokół określonych tematów i motywów, tak samo każda strona internetowa powinna koncentrować się na określonym obszarze lub temacie. Słowa i wyrażenia kluczowe wynikają z takiego tekstu w sposób naturalny, a użytkownicy chętniej pozostaną dłużej na stronie zawierającej odpowiednią treść i linki.

Trzeba jednak pamiętać, aby wyrażenia zawarte na stronie pokrywały się z wyrażeniami wyszukiwanymi przez użytkowników. Na przykład jeśli witryna jest przeznaczona dla entuzjastów Forda, oprócz wyrażenia typu „pojazdy amerykańskie” tekst witryny powinien także zawierać słowa „Ford” i „samochód”.

Zwiększanie widoczności: sprawdzone metody

Właściciele witryn często pytają nas o najlepsze sposoby na poprawienie widoczności i oceny swoich witryn w naszych wynikach wyszukiwania. W takich sytuacjach najczęściej odpowiadamy: „Myśl jak użytkownik, ponieważ właśnie w ten sposób staramy się myśleć my.”

Co to oznacza w praktyce? Przede wszystkim trzeba dać użytkownikom to, czego szukają, ponieważ trafność jest czynnikiem napędzającym i pozwalającym utrzymać duży ruch w witrynie.

Wielu właścicieli witryn nadmiernie koncentruje się na ocenie poszczególnych stron. Tymczasem poza oceną PageRank pozycja w rankingu zależy od ponad 200 innych czynników. Warto więc poświęcić więcej czasu na poprawę jakości treści i jej dostępności, niż na próby przechytrzenia algorytmów wyszukiwarki. Jeśli witryna nie spełnia naszych wymogów jakości, może zostać zablokowana w indeksie.

Na czym się koncentrować:

1. Stwórz trafną, atrakcyjną treść: użytkownicy trafiają na witrynę za pomocą linków, więc warto popracować nad wizualną atrakcyjnością każdej strony.
2. Staraj się angażować użytkowników: jeśli można, warto umieszczać w witrynie opcje komentarzy. Tworzenie społeczności zachęca do regularnego odwiedzania witryny. Angażowanie użytkowników pomaga poprawić widoczność i zwiększa wierność użytkowników.
3. Monitoruj swoją witrynę: za pomocą Centrum dla webmasterów (zobacz str. 13) można sprawdzić jakie hasła wyszukują użytkownicy trafiający na witrynę lub śledzić zmiany rankingu w wynikach wyszukiwania po wprowadzeniu większych zmian w witrynie.
4. Staraj się powiązać witrynę z linkami o jak najwyższej jakości.
5. Stosuj przejrzyste linki tekstowe i pamiętaj, by zawierały opis treści docelowej.

Czego unikać:

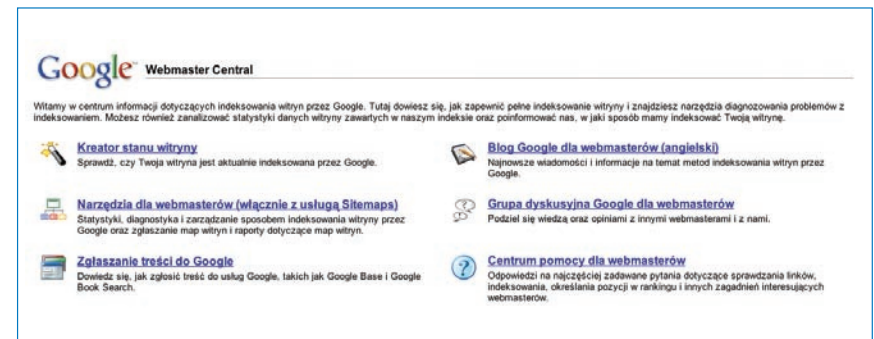
1. Nie wypełniać strony listami słów kluczowych.
2. Nie próbować „maskować” stron, umieszczając na nich tekst widoczny tylko dla wyszukiwarek.
3. Nie umieszczać stron lub linków przeznaczonych tylko dla robotów indeksujących, których jedynym celem jest oszukanie wyszukiwarek.
4. Nie umieszczać ważnych nazw, treści lub linków w postaci grafiki — grafika jest „nieczytelna” dla wyszukiwarek.
5. Nie umieszczać wielu kopii tej samej strony pod różnymi adresami URL w celu oszukania wyszukiwarek.

W razie wątpliwości warto przejrzeć wytyczne dla webmasterów, dostępne pod adresem: google.pl/webmasters/guidelines.html

Webmaster Central

Jesteśmy firmą, której celem jest zapewnianie najbardziej trafnych i użytecznych wyników wyszukiwania informacji w internecie i staramy się oferować skalowalną i sprawiedliwą pomoc techniczną dla wszystkich webmasterów i wszystkich witryn, bez względu na ich wielkość. Właśnie dlatego stworzyliśmy stronę Webmaster Central i umieściliśmy ją pod adresem google.pl/webmasters.

Strona Webmaster Central jest bogatym źródłem informacji dla wszystkich wydawców internetowych. Znajdują się tam wyczerpujące odpowiedzi na temat przemierzenia, indeksowania czy rankingu; można tam podzielić się z nami swoimi opiniami lub zgłosić problem; są tam dostępne narzędzia diagnostyczne dla webmasterów pomagające usuwać problemy związane z przemierzaniem stron.



Oto co można znaleźć w centrali Webmaster Central.

- Diagnostyka potencjalnych problemów w dostępie do stron i propozycje rozwiązań
- Możliwość usuwania stron z naszego indeksu
- Sprawdzanie, czy plik robots.txt dopuszcza i blokuje odpowiednie strony
- Wyświetlanie statystyk stron i zapytań użytkowników odwiedzających witrynę:
- Statystyki zapytań: określanie haseł najczęściej wyszukiwanych przez użytkowników odwiedzających witrynę oraz możliwości rozszerzenia witryny o potencjalnie interesujące tematy w celu zwiększenia liczby odsłon
- Analiza strony: możliwość wyświetlania witryn tak, jak widzi je Google. Można wyświetlić najpopularniejsze słowa w witrynie, linki prowadzące do witryny i sposób jej opisu w innych witrynach połączonych z nią linkami
- Częstotliwość przemierzania: możliwość sprawdzenia częstotliwości przemierzania witryny przez Googleboty i powiadomienia Google, czy proces ten ma się odbywać częściej czy rzadziej

Sitemaps

Strona Webmaster Central oferuje wydawcom usługę Sitemaps do analizy standardowych wyników wyszukiwania, a także wyników Google Mobile oraz wyszukiwania wiadomości Google News.

Sitemaps jest protokołem, który obsługujemy wraz z innymi wyszukiwarkami, dzięki czemu webmasterzy mogą udzielać nam więcej informacji o swoich stronach internetowych. Usługa Sitemaps stanowi uzupełnianie istniejących standardowych mechanizmów przemierzania internetu i pozwala webmasterom informować Google o poszczególnych stronach w witrynach, poprawiając ich widoczność w wynikach wyszukiwania Google.

Poza usługą Sitemaps dla wyszukiwania w internecie oferujemy także usługę Sitemaps dla wyszukiwania Google Mobile, dzięki której wydawcy mogą zgłaszać adresy URL oferujące treści dla urządzeń komórkowych za pomocą naszego indeksu witryn komórkowych.

W przypadku wydawców wiadomości sieciowych w Google News usługa Sitemaps daje dostęp do statystyk związanych z publikowanymi artykułami, od haseł wyszukiwania po częstotliwość odsłon. W połączeniu z narzędziami diagnostycznymi dostępnymi na stronie Webmaster Central usługa Sitemaps dla wiadomości Google News pozwala także wyświetlić błędy uniemożliwiające Google przemierzanie lub ekstrakowanie artykułów informacyjnych z witryny wydawcy. Ponadto wydawca może zgłosić protokół Sitemaps dla wiadomości zawierający adresy URL, które chciałby uwzględnić w Wyszukiwaniu wiadomości Google. Usługa Sitemaps dla wiadomości Google News, w przeciwieństwie do tej samej usługi dla standardowej wyszukiwarki i wyszukiwarki Google Mobile, jest aktualnie dostępna tylko w języku angielskim, choć chcemy wkrótce udostępnić ją także w innych językach.

Najczęściej zadawane pytania

Dlaczego nie mogę otrzymać indywidualnej pomocy technicznej dla swojej witryny?

Według niektórych wyliczeń w internecie znajduje się ok. 100 milionów witryn. Każda z nich jest dla nas ważna, ponieważ bez nich, bez względu na ich wielkość, nasz indeks byłby mniej dokładny i w efekcie mniej użyteczny dla naszych użytkowników.

Strona Webmaster Central jest znakomitym źródłem pomocy technicznej dla wszystkich typów witryn. Publikujemy tam pytania wydawców wraz z odpowiedziami, tak by informacje te były dostępne dla każdego. Poprzez stronę Webmaster Central można także poznać innych, doświadczonych webmasterów, którzy mogą udzielić wielu cennych rad i pomóc rozwiązać wiele problemów.

Czy wasze reklamy mają wpływ na ranking? Czy wasza lista reklam i wyniki wyszukiwania są od siebie zupełnie niezależne?

Rankingi reklam i wyszukiwania nie są ze sobą w żaden sposób powiązane; pracują nad nimi oddzielne zespoły, tak aby nie powstawały żadne kolizje. Uważamy, że obiektywność naszych wyników wyszukiwania jest najważniejszym czynnikiem, który pozwala nam dotrzeć do największej liczby użytkowników.

Jak dodać witrynę do indeksu wyszukiwania Google?

Obecność w wynikach wyszukiwania Google jest bezpłatna i nie wymaga ręcznego zgłaszania witryny w Google. Google jest całkowicie zautomatyzowaną wyszukiwarką; specjalne roboty regularnie przemierzają sieć w poszukiwaniu witryn, które można dodać do naszego indeksu. Właściwie ogromna większość witryn figurujących na naszym indeksie nie jest zgłaszana ręcznie, ale wyszukiwana i dodawana automatycznie w procesie przemierzania sieci.

Ponadto narzędzia dla webmasterów Google (np. strona Webmaster Central) pozwalają w prosty sposób zgłaszać mapę witryny lub adresy URL do indeksu Google i korzystać ze szczegółowych raportów dotyczących widoczności stron w Google. Dzięki narzędziom dla webmasterów właściciele witryn mogą automatycznie informować Google o wszystkich aktualnych bieżących stronach i wszelkich aktualizacjach.

Ile czasu potrzebuje średnio Google, aby znaleźć nowo utworzoną witrynę i jaka jest ogólna częstotliwość przemierzania sieci przez Google?

Nie ma ustalonego czasu potrzebnego do wyszukania nowej witryny przez Google.

Googlebot regularnie przemierza sieć, aktualizując nasz indeks. Za pomocą strony Webmaster Central webmaster może sprawdzić częstotliwość przemierzania swojej witryny i poinformować Google, czy proces ten ma odbywać się częściej czy rzadziej.

Co mam zrobić, jeśli chcę, by moja witryna, poza zwykłymi wynikami wyszukiwania, pojawiała się także w wynikach wyszukiwania wiadomości lub grafiki?

Google zawsze pozwala wydawcom internetowym zrezygnować z poszczególnych usług. W tym celu trzeba się skontaktować z odpowiednim zespołem pomocy technicznej.

Jak pisaliśmy wcześniej, za pomocą Protokołu wykluczenia robotów można blokować indeksowanie zarówno grafiki, jak i stron internetowych. Do tego celu można także wykorzystać funkcję „usuwania adresu URL”, dostępną na stronie Webmaster Central. Obejmuje ona wyniki wyszukiwania w internecie i wyniki wyszukiwania grafiki.

Ponadto ponieważ na funkcje Googlebota składa się praca kilku innych robotów, można precyzyjnie określić blokowane elementy:

- Googlebot: przemierzanie stron uwzględnionych w indeksie witryn internetowych i Google News.
- Googlebot-Mobile: przemierzanie stron uwzględnionych w indeksie Google Mobile.
- Googlebot-Image: przemierzanie stron uwzględnionych w indeksie grafiki.
- Mediapartners-Google: przemierzanie stron w celu ustalenia treści AdSense. Ten robot jest stosowany tylko do przemierzania witryn, w których wyświetlane są reklamy AdSense.
- Adsbot-Google: przemierzanie stron w celu oceny jakości strony docelowej AdWords. Ten robot jest stosowany tylko do przemierzania witryn reklamowanych w programie Google AdWords.

Aby na przykład całkowicie zablokować Googlebota, można użyć następującej składni:

User-agent: Googlebot

Disallow:

Czy mogę zdecydować, który fragment tekstu mojej witryny będzie wyświetlany w wynikach wyszukiwania?

Nie. To nie jest dobry pomysł, zarówno z perspektywy użytkowników, jak i samego autora treści. Wybieramy fragment tekstu z witryny i wyświetlamy go w kontekście wyszukiwanego hasła, co z kolei pozwala określić trafność wyniku.

Według badań użytkownicy częściej odwiedzają witryny, które w wynikach wyszukiwania wyświetlane są wraz z fragmentem tekstu. Przyczyną takiego zachowania jest łatwość oceny trafności takiego wyniku w kontekście wyszukiwanego hasła. Jeśli użytkownik nie może szybko podjąć takiej decyzji, przechodzi do kolejnych wyników.

Wydawcy internetowi mogą umieszczać na swoich stronach tagi meta będące dodatkową informacją dla Google, gdy z treści strony nie można algorytmicznie wygenerować użytecznego fragmentu tekstu. W tym celu wystarczy dodać następujące tagi przed sekcją <head> strony:

```
<meta name="description" content="Dlaczego Ania lubi króliki? Zaraz się dowiemy.">
```

Każdy wydawca internetowy, który nie chce, aby w wynikach wyszukiwania był wyświetlany fragment tekstu strony może użyć tagu NOSNIPPET w następujący sposób:

```
<meta name="robots" content="nosnippet">
```

Wreszcie niekiedy, aby wyświetlić fragment tekstu w wynikach wyszukiwania korzystamy z opisu strony w ramach projektu Open Directory Project. Aby uniemożliwić korzystanie z tego opisu, wystarczy dodać następujący tag meta:

```
<meta name="robots" content="noodp">
```

Artykuły z informacjami z ostatniej chwili na mojej witrynie są umieszczane tylko na kilka godzin, po czym są aktualizowane i przenoszone do sekcji standardowych artykułów. Chcę, aby w wynikach wyszukiwania Google dostępne były pełne artykuły, a nie artykuły z ostatniej chwili.

Jedną z możliwości jest umieszczenie wszystkich informacji z ostatniej chwili w jednym katalogu i zablokowanie dostępu do niego za pomocą pliku robots.txt.

Innym sposobem jest dodanie tagu NOFOLLOW do sekcji <HEAD> kodu html sekcji informacji z ostatniej chwili. Dzięki temu Googlebot nie będzie podążał za linkami umieszczonymi na tej stronie. Trzeba jednak pamiętać, że tag NOFOLLOW uniemożliwia tylko podążanie za linkami między stronami. Jeśli do artykułu prowadzi link zewnętrzny, Google może go indeksować.

Jeśli mam wiele nazw domen i we wszystkich mam tę samą treść, czy moje strony zostaną usunięte z wyników wyszukiwania?

Choć niektórzy wydawcy starają się oszukiwać wyszukiwarki, duplikując treść lub uruchamiając witryny lustrzane, treść można także duplikować całkiem legalnie. Google nie chce karać tych witryn. Na przykład nie traktujemy jednakowo tych samych treści publikowanych w różnych językach (powiedzmy po angielsku w jednej witrynie i po francusku w innej).

Publikowanie tych samych treści w wielu witrynach (np. rozpowszechnianie artykułów prasowych) nie zawsze oznacza konieczność całkowitego usunięcia tych witryn z wyników wyszukiwania. Trzeba jednak pamiętać, że każda kolejna kopia artykułu będzie wyświetlana niżej w wynikach wyszukiwania ponieważ w porównaniu z pojedynczą kopią prowadzi do niej tylko niewielka część linków. Zatem pojedyncza kopia artykułu jest z reguły wyświetlana wyżej w wynikach wyszukiwania i może dotrzeć do większej liczby odbiorców.

Dodatkowo aby zapewnić odpowiednią jakość wyszukiwania, Google nie umieszcza w wynikach wielu kopii jednej strony. Zamiast tego wybieramy jedną jej wersję. Webmasterzy mogą wskazać preferowaną wersję za pomocą pliku robots.txt lub tagu meta, blokując wszystkie niechciane kopie.

Dlaczego moja witryna jest blokowana i nie może znaleźć się w indeksie Google?

Po pierwsze witryna wcale nie musi być blokowana. Istnieje wiele przyczyn, dla których może ona nie pojawiać się w wynikach wyszukiwania (zobacz str. 5-11).

Jeśli Twoja witryna nie zawiera żadnych przeszkód utrudniających jej wykrycie lub indeksację, wtedy może rzeczywiście być blokowana. Jedną z przyczyn blokowania witryny może być jej niezgodność ze standardami jakości określonymi we Wskazówkach dla webmasterów (dostępnych na stronie Webmaster Central). Najczęściej ma to miejsce, gdy witryna stosuje nieuczciwe metody uzyskania wyższego miejsca w wynikach wyszukiwania. Powszechne naruszenia wytycznych obejmują maskowanie (umieszczenie w witrynie tekstu widocznego tylko dla wyszukiwarek) lub tworzenie stron/linków, których jedynym celem jest wprowadzanie wyszukiwarek w błąd i manipulacja wynikami wyszukiwania.

W przypadku naruszenia zasad jakości można wprowadzić w witrynie odpowiednie zmiany i kliknąć link „Żądanie ponownego uwzględnienia” w interfejsie Narzędzi webmastera, aby poprosić o jej ponowną ocenę.

Słowniczek

Adres URL (Uniform Resource Locator)

Adres witryny w internecie składający się z protokołu dostępu (http), nazwy domeny (www.google.pl), i w niektórych przypadkach lokalizacji pliku (www.google.pl/webmaster).

HTML (Hypertext Markup Language)

Język znaczników używany w internecie do definiowania struktury tekstu.

Indeksowanie

Proces dodawania treści witryny do wyszukiwarki.

Katalog główny

Nadrzędny lub główny katalog w systemie plików komputerowych.

Link do kopii strony w pamięci podręcznej

Migawka wyglądu strony podczas ostatniego cyklu indeksowania Google. Kopia pozwala wyświetlić stronę nawet po usunięciu jej oryginału, choć jej treść może być trochę inna. Aby wyświetlić kopię strony, kliknij link „Kopia” pod wynikiem wyszukiwania.

Maskowanie

Udostępnianie innych treści wyszukiwarce, a innych treści użytkownikom.

Ocena Page Rank

Technologia Google pozwalająca określić pozycję witryny w wynikach wyszukiwania. Ocena PageRank wynika z unikatowego demokratycznego charakteru sieci i ogromnej struktury linków będącej wskaźnikiem wartości poszczególnych stron. Ważne strony o wysokiej jakości mają wyższą ocenę PageRank, która jest zapamiętywana przez Google przy każdym wyszukiwaniu. Google łączy ocenę PageRank z zaawansowanymi technikami dopasowania tekstu, aby lepiej wyszukiwać odpowiednie strony.

Protokół wykluczenia robotów

Specyfikacja techniczna informująca wyszukiwarki, które witryny lub ich fragmenty mają pozostawać niewidoczne dla wyszukiwania, a które mają pojawiać się w wynikach wyszukiwania.

Przemierzanie

Proces stosowany przez wyszukiwarki do gromadzenia stron internetowych.

Robot indeksujący

Oprogramowanie służące do wykrywania i indeksowania adresów URL w sieci wewnętrznej lub internecie.

Robots.txt

Plik tekstowy umożliwiający wydawcom internetowym kontrolę dostępu do ich witryn na wielu poziomach, od całej strony po poszczególne katalogi, strony określonego typu czy nawet indywidualne dokumenty. Ten plik informuje roboty indeksujące, które katalogi są dostępne do przemierzenia.

Rozszerzenie pliku

Nazwa pliku komputerowego (.doc, .txt, .pdf itd.) często używana do identyfikacji typu danych przechowywanych w pliku.

Słowo kluczowe

Słowo wprowadzane w pole wyszukiwania wyszukiwarki. Wyszukiwarka przeszukuje strony zawierające dane słowo lub wyrażenie.

System zarządzania treścią (CMS)

Oprogramowanie służące do zarządzania treścią od plików komputerowych, graficznych i dźwiękowych, po treści internetowe.

Tagi meta

Tag w języku HTML opisujący treść strony internetowej. Tagi meta mogą służyć do kontroli indeksowania poszczególnych stron w witrynie.

Treść dynamiczna

Np. grafika, animacje lub wideo oparte o Flash, JavaScript, ramki lub dynamiczne adresy URL.

Witryna lustrzana

Replika witryny; czasami służąca do oszukiwania wyszukiwarek i optymalizacji wyników indeksowania i pozycji witryny w wynikach wyszukiwania.

Więcej informacji na temat strony Webmaster Central znajduje się pod adresem:

google.pl/webmasters/

