

Google - pomoc dla tłumacza (podstawy)

Google WWW pozwala na wyszukiwanie w setkach milionów dokumentów tekstowych z całego świata. Dzięki temu można wykorzystać je jako pomoc w procesie tłumaczenia tekstów.

Ten artykuł jest przeznaczony głównie dla tłumaczy poznających wyszukiwarkę Google i jej wielki potencjał.

Tłumacz i Google - podstawy

Google to najpopularniejsza i najefektywniejsza wyszukiwarka internetowa na świecie. Projekt ciągle się rozrasta o nowe moduły. Nas, jako tłumaczy, najbardziej interesują moduły: Google WWW, Images i Groups.

O jednej, niezmiernie ważnej rzeczy należy zawsze pamiętać: **Google nie jest wyrocznią**. Google może pomóc, ale może i zaszkodzić. Dlatego zawsze, o ile to możliwe, **należy zweryfikować pozyskane informacje**.

Google WWW

Esencja Google. Google WWW pozwala na wyszukiwanie w setkach milionów dokumentów tekstowych z całego świata. Ze względu na prostotę interfejsu zwykle, codzienne wyszukiwanie informacji jest intuicyjne i nie nastrocza większych problemów. Wyszukiwanie informacji wyspecjalizowanych może się okazać bardziej skomplikowane.

Wpisane nieprecyzyjnej frazy w wyszukiwarce spowoduje zalanie nas tysiącami stron, które mogą nie nieść żadnej interesującej nas treści. Pytanie brzmi, jak formułować odpowiednie zapytania?

Podstawy

Zacznę od przybliżenia najbardziej podstawowych technik wyszukiwania za pomocą najprostszych narzędzi oferowanych przez Google.

Cudzysłów

Najprostszym sposobem na doprecyzowanie jest użycie cudzysłowu. Fraza: [statek kosmiczny] i ["statek kosmiczny"] (*), to dwa zupełnie różne zapytania. Wpisanie frazy bez cudzysłowu da nam wynik nieprecyzyjny, z ogromną liczbą stron, na których słowa statek oraz kosmiczny owszem występują, jednak nie obok siebie, a o to nam przecież chodziło. Naprawimy to, ujmując frazę w cudzysłów. Wynik wyszukiwania będzie mniej chaotyczny, a mamy pewność, że oba słowa występują obok siebie.

Oczywiście w cudzysłów można ująć więcej słów. Można również używać więcej niż jednej frazy zamkniętej w cudzysłowie. Przykład: ["statek kosmiczny" mars] lub też: ["statek kosmiczny" "lot na marsa"].

W powyższym przykładzie napisałem mars z małej litery - jest to dozwolone (w znaczeniu - nie zawęży wyników wyszukiwania), gdyż wyszukiwarka nie rozróżnia małych i wielkich liter, traktując je jako te same znaki.

Minus -

Znak [-] pozwala na zmniejszenie liczby trafień poprzez wyszukiwanie stron niezawierających danego słowa. Czyli jeśli chcemy wyszukać informacje o Marsie, ale dostajemy dużo stron dotyczących batoniku lub papierosów wpisujemy: [mars -batonik]. W rezultacie otrzymamy wszystkie strony, na których występuje słowo mars i nie występuje słowo batonik. Można oczywiście wstawić znak [-] przed całą frazą. Gdy chcemy wyszukać informacje o marsie, jednak denerwuje nas stwierdzenie "czerwona planeta", do wyszukiwarki wpisujemy następujące zapytanie: [mars -"czerwona planeta"].

Filetype:

Wiele użytecznych publikacji dostępnych jest tylko w formacie .PDF. Jeśli wiemy, że interesujący nas tekst prawdopodobnie może być dostępny w takim właśnie formacie, gdyż jest to np. raport, zestawienie, umowa, instrukcja obsługi, praca naukowa, wpisujemy: [mars filetype:pdf]. Niekiedy pomimo tego, że znamy nawet tytuł interesującej nas publikacji, nie jesteśmy w stanie jej sprawnie wyszukać w natłoku dokumentów. Sam tytuł może się przewijać na wielu stronach: w bibliografiach, przypisach, cytatach itp. Użycie narzędzia filetype: pozwala na sprawnie wyszukanie takiej właśnie zagubionej pracy. Oczywiście filetype: nie ogranicza się tylko do plików PDF. Bez najmniejszych problemów można wyszukać pliki .DOC, .XLS, .JPG (choć to trochę mija się z celem, gdyż Google oferuje osobny moduł - Images, który opisałem w artykule [Google - Zaawansowane](#)), .PPS i inne, w sumie 12 typów plików.

Site:

Wiadomo, że "Google might not be right, but it certainly is quick". Narzędzie site: pozwala, w miarę, połączyć szybkość Google'a, ze względną poprawnością otrzymanych wyników. Dzięki niemu możemy wyszukiwać informacje w obrębie jednej strony: [mars site:onet.pl] lub, co o wiele bardziej przydatne, jednej domeny. Dlaczego bardziej przydatne? Ponieważ w Internecie publikować może każdy, a nie zawsze piszą osoby władające biegle danym językiem. Nietrudno więc sobie teraz wyobrazić wpadkę przy wyszukiwaniu kolokacji, gdy wyświetlane są strony w języku angielskim pisane przez mieszkańców Bangladeszu, Chin czy Paragwaju (bez urazy, oczywiście to tylko przykłady). Taka informacja może nas tylko zmylić. Wprowadzające w błąd wyniki można ograniczyć w ustawieniach samej wyszukiwarki: http://www.Google.pl/advanced_search?hl=pl, wybierając opcję "Pokaż strony w języku [angielskim]". Jednak to za mało. Nadal będziemy dostawać wyniki ze stron z domeny np. .com i to właśnie ją chcemy odfiltrować. Można zrobić to na dwa sposoby:

- wpisujemy: [mars -site:com] i otrzymujemy wyniki ze słowem mars, ze stron z domen innych niż .com,
- wpisujemy: [mars site:edu] lub [mars site:gov] i otrzymujemy wyniki ze słowem mars, ze stron z domen . edu lub .gov

Moim zdaniem wyszukiwanie z site:edu, zwłaszcza gdy szukamy zagadnień naukowych, jest najpewniejsze. Mamy duże prawdopodobieństwo, że strona została napisana przez osobę biegle władającą językiem, a do tego zajmującą się nauką lub nauczaniem. Site:gov natomiast przydatne jest niezmiernie przy wyszukiwaniu wszelkiego typu dokumentów urzędowych i rządowych.

Sposobów na Google jest mnóstwo i z pewnością nie wyczerpałem wszystkich, ani w tym artykule ani w kolejnych. Podałem tylko te, których sam używam i które dają najlepsze efekty. Jeśli ty drogi czytelniku znasz jakieś sprawdzone patenty na Google, które pomagają tłumaczowi, zapraszam do podzielenia się informacją. W razie pytań, sugestii, nagan, pochwał, błędów, nieścisłości również proszę o [wiadomość](#).

* We wszystkich przykładach należy wpisywać zapytania bez znaków [oraz]. Używam ich tylko do rozróżnienia zapytania od reszty zdania.

Przedruk dokonany za zgodą autora

Autor: Michał Tosza

Przedruk ze strony: <http://translators.org.pl/blog/index.php/2007/05/11/google-pomoc-dla-tlumacza-podstawy/>

Artykuł pobrano ze strony eioba.pl