

Efektywne wyszukiwanie informacji w Internecie

Internet rozrasta się w takim tempie, że znalezienie w nim czegokolwiek jest bardzo trudne. Obecnie w Internecie jest już kilkadziesiąt miliardów stron

Największa wyszukiwarka na świecie Google indeksuje kilka miliardów stron, a jest to ilość wstępnie już wyselekcjonowana przez ich algorytmy. Dochodzi do tego jeszcze niewidzialna część Internetu - niewidzialna dlatego że wyszukiwarki nie mogą jej zaindeksować.

Jest pełno katalogów i wyszukiwarek, do których możemy się udać i w których możemy szukać interesujących nas informacji - oczywiście wymaga to chociaż podstawowej wiedzy. Do skutecznego szukania informacji musimy wiedzieć jak używać wyszukiwarki. Obecnie wyszukiwarki indeksują tak dużo stron, że problemem nie jest brak informacji, ale to, że tych informacji jest za dużo. Dlatego należy się nauczyć jak szukać efektywnie.

Trzeba pamiętać, że **internet to nie tylko strony**. W internecie jest pełno wyszukiwarek, katalogów, specjalistycznych baz danych i innych serwisów, które pozwalają wyszukiwać specyficzne informacje. Pytanie brzmi: którego narzędzia użyć aby znaleźć to czego szukamy?

Grupy dyskusyjne w google

Możemy znaleźć wiele informacji na grupach dyskusyjnych, które obejmują swoimi tematami, prawie wszystko. Są to rozmowy luźne bądź specjalistyczne - zależnie od charakteru grupy.

Jeśli interesuje nas hobbystycznie bądź zawodowo jakiś temat i przydałaby się pomoc innych ludzi, to grupy dyskusyjne są najlepszym miejscem do znalezienia takiej pomocy. Wystarczy do tego czytnik grup dyskusyjnych wbudowany w wiele różnych programów pocztowych, dzięki któremu możesz czytać wybrane grupy dyskusyjne, pisać a także pomagać innym.

Grupy dyskusyjne dzięki **Google Groups Service** można przeszukiwać. Wyszukiwanie obejmuje także polskie grupy dyskusyjne. Na stronie google.pl w górnej zakładce znajdziesz przycisk Grupy dyskusyjne.

Każdy z nas używa narzędzia do wyszukiwania informacji w internecie. Istnieją dwa podstawowe narzędzia: katalogi i wyszukiwarki.

Katalogi

Katalogi są tworzone przez ludzi. Każda strona jest przeglądana przez redaktora i dopiero potem dodawana do katalogu. Katalogi są poukładane w kategorie, żeby łatwiej można

było dotrzeć do interesujących nas informacji. Katalog ma strukturę drzewa począwszy od bardzo ogólnych informacji. Wchodząc w głąb zaczyna się podział na coraz bardziej szczegółowe tematy.

Yahoo było pierwszym katalogiem internetowym i na bazie tego katalogu powstały kolejne katalogi.

DMOZ.org - Open directory project jest największym światowym projektem otwartego katalogu stron internetowych. Każdy może zostać redaktorem katalogu DMOZ - oczywiście nieodpłatnie - jest to największy ogólnoswiatowy katalog stron.

W kategorii <http://dmoz.org/World/Polska/> znajdziemy strony polskie skatalogowane w odpowiednie grupy. Google również korzysta z wyników katalogu DMOZ w katalogu na swojej stronie.

Polskie katalogi

Katalog Wirtualna Polska (wp.pl) - katalog ten zawiera mnóstwo stron z całej Polski. Główne wyniki wyszukiwania w pierwszej kolejności pobierane są z katalogu, a w drugiej - dopiero z wyszukiwarki Netsprint, jeżeli w katalogu nie można znaleźć żadnej odpowiedzi.

Katalog onet.pl - katalog portalu onet.pl zawierający bardzo dużo stron z całej Polski. Katalogi są poukładane w kategorie. W kategoriach mamy podkategorie. I tak od najogólniejszych w łatwy sposób można trafić do konkretnej kategorii, zawierającej poszukiwane informacje.

Jeżeli wyszukujesz coś w katalogu to wyniki wyszukiwania są ograniczone tylko do stron zawartych w katalogu, stron wyselekcjonowanych i opisanych przez redaktorów na podstawie sugestii osoby zgłaszającej. Wyszukiwarka wbudowana w katalog z reguły wyszukuje w tytułach i opisach ignorując treści na stronie. Brane za to pod uwagę są też nazwy kategorii.

Wyszukiwarki

Wyszukiwarki to zupełnie co innego. Ich wyniki zbierane są przez roboty, a oceniane przez algorytmy, które na podstawie zapytań szukającego starają się dawać jak najlepszą odpowiedź, czyli strony najbardziej pasujące do zapytania. Niektóre wyszukiwarki robią to lepiej inne gorzej.

Niekwestionowanym **liderem rynku jest** wyszukiwarka google.pl. Prawie 50% stron jest znajdowane właśnie dzięki tej wyszukiwarce. Swój sukces zawdzięcza unikalnemu algorytmowi i systemowi oceny stron. **Jeżeli mamy bardziej skomplikowane zapytanie to właśnie tam powinniśmy się udać.**

Na polskim rynku istnieje szereg różnych wyszukiwarek, najpopularniejsze z nich oprócz **google.pl** to:

- onet.pl
- wp.pl

Onet.pl i wp.pl w sumie kontrolują ok 40% rynku polskiego.

Oprócz tego istnieją jeszcze wyszukiwarki:

- netsprint.pl - również wyszukiwarka działająca na własnym oprogramowaniu. Wspiera wyniki wyszukiwania wp.pl.
- szukacz.pl - również wyszukiwarka działająca na własnym oprogramowaniu, włącznie z odmianą fleksyjną.
- gooru.pl - bardzo podobna z wyglądu wyszukiwarka do google tylko działająca na polskim rynku.
- emulti.pl - wyszukuje wyniki w kilku wyszukiwarkach i na podstawie tych informacji generuje własną listę wyników.

Oczywiście na Zachodzie istnieje o wiele więcej wyszukiwarek, do których warto zajrzeć, jeśli interesują nas zachodnie wyniki.

- google.com
- yahoo.com - nie to samo co katalog
- altavista.com
- aol.com
- ask.com
- teoma.com
- clusty.com

Zaawansowane wyszukiwanie

Potrzebujemy wiersz Adama Mickiewicza pt. *"Czatyrdah"*

Wpisujemy w wyszukiwarce (np. www.google.pl)

Mickiewicz

i dostajemy tysiące, albo dziesiątki tysięcy wyników. Wyszukanie interesującego nas tekstu jest bardzo utrudnione, gdyż akurat wiersz ten może być dopiero na 30 czy 40 stronie wyników.

Musimy skonkretyzować swoje zapytanie.

Mickiewiczów poza naszym wieszczem jest dużo, dlatego lepszym zapytaniem będzie

Adam Mickiewicz

ale nadal to zapytanie nie określa o jaki wiersz nam chodzi.

Kolejnym dodatkiem do zapytania jest

Adam Mickiewicz Czatyrdach

- teraz można szukać, google w wyniku takiego zapytania zwróci konkretne informacje, często linki prowadzące bezpośrednio do wiersza.

Wpisując jeden wyraz z reguły tracisz czas.

Wpisuj dwu-, trzywyrazowe frazy, żeby skonkretyzować swoje zapytanie, na pewno wyniki będą o wiele lepsze niż pod samym słowem Mickiewicz.

Frazy

Jeżeli potrzebujesz wyszukać jakieś frazy np. "Adam Mickiewicz" i poszukujesz dokładnie tego człowieka, nie innego, to wpisując to w wyszukiwarkę należy wpisać

"Adam Mickiewicz"

Czyli imię i nazwisko zamknąć w cudzysłowy.

Jeżeli zadasz pytanie Adam Mickiewicz do onet.pl bez cudzysłówów to w odpowiedzi otrzymasz strony zawierające samo Adam i strony zawierające samo Mickiewicz oraz strony zawierające i Adam i Mickiewicz.

Jeżeli zadasz pytanie do google.pl to otrzymasz lepsze wyniki bo na stronie musi wystąpić i Adam i Mickiewicz, ale nie koniecznie obok siebie.

Dopiero zadanie pytania z użyciem cudzysłowu (tworzysz w ten sposób frazę) "Adam Mickiewicz" w wyniku otrzymasz tylko strony na których dokładnie występuje taka fraza, czyli pierwsze Adam a zaraz po spacji Mickiewicz.

Dokładnie to czego szukałeś.

Jeżeli szukasz jakieś frazy, która ma być zapisana na stronie dokładnie tak jak ją wpiszesz w wyszukiwarkę. np. Tusz HP 610 (tzn na stronie ma być to dokładnie tak samo zapisane bez żadnych dodatkowych wyrazów pomiędzy), to wpisując do wyszukiwarki musisz ująć tę frazę w cudzysłów.

"tusz HP 610"

Wyszukiwarka zwróci tylko te wyniki, które będą zawierać dokładnie taką frazę.

Jeżeli nie użyjesz cudzysłowu to znaczy, że szukasz stron, na których te trzy wyrazy występują w dowolnym miejscu, niekoniecznie obok siebie lub jak np. w onet.pl występuje chociaż jeden z wyrazów.

Operatory wyszukiwania AND, NOT

Wyszukiwarki posiadają wiele pomocniczych operatorów, które pozwalają nam dokładnie określić czego szukamy. Wpisywanie pojedynczego słowa do pola wyszukiwarki jest najprostszą funkcją wyszukiwarki i obecnie ze względu na ogrom zwracanych wyników - mało przydatną. Z reguły podczas wyszukiwania należy używać fraz, które pozwalają dokładniej określić czego szukamy.

Założmy, że chcesz znaleźć wiersz, Czatyrdah Adama Mickiewicza, ale nie interesuję cię kupienie go w księgarni.

Czatyrdah AND Mickiewicz NOT księgarnia

W ten sposób mówisz dokładnie wyszukiwarce co ma zrobić. Wyszukiwarka wyszukuje wyniki zawierające słowa: Czatyrdah i Mickiewicz i z tego zbioru odrzuca strony zawierające słowo księgarnia.

Zbiór stron, które nie zostały odrzucone zostaną wyświetlone jako wynik zapytania.

Co się stanie jeśli usuniesz operatory?

Czatyrdah Mickiewicz księgarnia

Przed wszystkim **brak operatora NOT** przed słowem księgarnia spowoduje zwrócenie stron, które zawierają to słowo, czyli dokładnie nie to co szukasz.

Zapytanie zwróci różne wyniki i jest to zależne od ustawień domyślnych konkretnej wyszukiwarki.

Niektóre wyszukiwarki zwrócą tylko strony zawierające wszystkie słowa, czyli każde z podanych słów musi wystąpić na jednej stronie, a inne zwrócą strony, które zawierają przynajmniej jedno ze słów.

[google.pl](#) na zadane pytanie Czatyrdah Mickiewicz księgarnia - zwróci konkretne odpowiedzi strony zawierające wszystkie słowa.

[onet.pl](#) - zwróci strony które zawierają chociaż jedno z wymienionych słów. Nie jest to niestety dobre rozwiązanie bo często na górze wyników są strony, które zawierają tylko słowo Czatyrdah. Dlatego szczególnie w onet.pl używaj operatorów AND żeby otrzymać konkretne wyniki.

Operatory AND i NOT obsługuje praktycznie każda z liczących się wyszukiwarek.

Uprozczone wyszukiwanie booleanowe

Wpisywanie pełnych nazw operatorów czyli **AND** lub **NOT** jest czasochłonne, dlatego istnieje skrócony zapis. Operatory są po prostu zastąpione znakami matematycznymi.

Nie musisz wpisywać zapytania do wyszukiwarki.

*Czterydah **AND** Mickiewicz **NOT** księgarnia*

To samo można osiągnąć wpisując

Czterydah+Mickiewicz-księgarnia

Jak widać o wiele krócej i wygodniej, a zapytanie zwróci te same wyniki co znacznie dłuższy poprzedni zapis.

AND zastąpione jest znakiem + (plus)

NOT zastąpione jest znakiem - (minus)

W każdej wyszukiwarce działa to identycznie.

Operator OR

Załóżmy, że szukasz czegoś co może się różnie nazywać lub jest różnie zapisywane. Imię i nazwisko naszego wieszka można zapisać na różne sposoby.

- A.[**spacja**]Mickiewicz
- Adam Mickiewicz
- A.[**bez spacji**]Mickiewicz

Wyszukiwarki udostępniają operator OR, czyli: lub, który pozwala szukać np. tego samego tylko w różnym zapisie.

Zapis tego pytania będzie następujący:

*A. Mickiewicz **OR** "A. Mickiewicz" **OR** "Adam Mickiewicz"*

Frazy połączone z operatorem OR mówią wyszukiwarce, że ma szukać wszystkich 3 kombinacji, ale wystarczy, że jedna z kombinacji wystąpi na stronie a zostanie zwrócona w wynikach wyszukiwania.

Małe i duże litery

Większość wyszukiwarek interpretuje wyniki wg zasady: najpierw szuka wyrazu zapisanego dokładnie tak jak wpisałeś, a potem wyrazu tego samego zapisanego jednak małymi literami.

Google nie zwraca uwagi na wielkość liter, tzn. nie ważne jest jak wpiszesz, wyniki otrzymasz te same dla wyrazu pisanego z małej bądź dużej litery (całe zapytanie zmieniane jest po prostu na małe litery).

Wyszukiwanie w określonych polach

Każdy robot zbiera i zapisuje informacje o stronach do bazy wartości pól, które dana wyszukiwarka następnie analizuje. Może to być cała strona lub tylko kilka wybranych pól jak np. tytuł, opis i słowa kluczowe.

Każda wyszukiwarka ma przycisk: **zaawansowane szukanie**. Tam dostaniesz kilka opcji, dzięki którym możesz określić dokładniej w jakich polach ma wyszukiwarka szukać.

url: lub site: - możesz ograniczyć swoje wyniki wyszukiwania do konkretnej domeny podając url:domena lub site:domena - w ten sposób ograniczasz wyniki wyszukiwania tylko do podanej domeny.

Google pozwala szukać również wyników tylko w określonych typach plików filetype. Np. marketing +Internet filetype:pdf, (znajduje dokumenty pdf)
marketing +Internet filetype:ppt, (znajduje prezentacje)
marketing +Internet filetype:doc (znajduje dokumenty word)

Opcji tych jest o wiele więcej i zależą one od konkretnej wyszukiwarki.