

Wykład: Wyszukiwarki internetowe

Wyszukiwarki - domeny, z których trafiają na polskie witryny internauci łączący się z obszaru **Polski** (na podstawie wyników oglądalności witryn uczestniczących w badaniu site-centric gemiusTraffic)

Wyszukiwarki-domeny - trzy ostatnie tygodnie

	17.X.2006 - 23.X.2006	24.X.2006 - 30.X.2006	31.X.2006 - 6.XI.2006
google.com	79.60%	79.70%	80.70%
onet.pl	9.70%	9.40%	9.20%
wp.pl	6.90%	6.60%	6.40%
interia.pl	1.20%	1.20%	1.10%
msn.com	0.90%	1.30%	1.00%
szukacz.pl	0.40%	0.40%	0.50%
yahoo.com	0.60%	0.50%	0.30%
netsprint.pl	0.30%	0.40%	0.30%
o2.pl	0.20%	0.20%	0.20%

źródło: www ranking.pl.



<http://www.google.pl>

Od 1998 roku

Baza danych oparta o superkomputer, który składa się z setek tanich komputerów połączonych w sieć. Google zawiera odnośniki do około 8 miliardów adresów URL. Googol" to termin matematyczny oznaczający 1 ze 100 zerami.

<http://www.google.pl/intl/pl/press/facts.html>

Operatory wyszukiwania:

AND	toyota skoda	Domyślny operator
OR	Jedno LUB drugie	toyota OR opel
+	Wymusza obecność słowa	+jak +to zrobić
-	wyklucza dane słowo	-zbrodnia
*	nie uwzględnia odmian	
duże litery	nie ma znaczenia	
„”	Wyszukiwane wyrażenia obok siebie	„ochrona środowiska”
..	Liczby z podanego przedziału	1990..2006
~	podaje również synonimy wyrazu	~food

Pojedyncze litery są pomijane, ale również i przyimki!

Typy zapytań

- **Link:**wz.utp.edu.pl
Wyświetlenie stron z odnośnikami do wskazanej witryny
- **Related:**www.gazeta.pl
Lista stron podobnych stron do podanej
- **Define:**google
Wyszukanie definicji słowa „google”, w języku angielskim
- **Info:**wz.utp.edu.pl
Krótkie informacje nt. podanej strony

Modyfikatory zapytań:

- uniwersytet **site:**wz.utp.edu.pl
Ograniczenie wyszukiwania słowa „uniwersytet” do wskazanej strony
- sas **filetype:**pdf **site:**wz.utp.edu.pl czy też budżet **filetype:**xls
Wyszukiwanie plików z rozszerzeniem „pdf” na stronie wz. utp.edu.pl
- **intext:**zabawa
Ograniczenie wyszukiwania wyłącznie do widocznego tekstu na stronie internetowej
- **intitle:**ochrona środowiska
Ograniczenie przeszukiwania słowa „ochrona” do tytułu strony internetowej, „środowisko” może wystąpić gdziekolwiek
- **allintitle:**ochrona środowiska
Obydwa wyrazy muszą wystąpić w tytule strony
- **inurl:**stat sas
Wyszukiwanie wyrazu „stat” w adresach URL, natomiast wyraz „sas” może wystąpić w dowolnym miejscu strony
- **allinurl:**stat sas
Obydwa poszukiwane słowa muszą wystąpić w adresie URL

URL (ang. Uniform Resource Locator) zuniifikowany format adresowania zasobów (informacji, danych, usług), stosowany w Internecie i w sieciach lokalnych. Pozwala adresować wszelkie zasoby dostępne w Internecie.