

WYSZUKIWANIE BOOLOWSKIE

Wyszukiwanie boolowskie jest rozszerzeniem wyszukiwania prostego (opartego o słowa kluczowe) o operatory logiczne: AND, OR, NOT oraz ich kombinację. Większość modeli wyszukiwania oparta jest na podstawach logiki boolowskiej, a dokładniej na relacjach między poszczególnymi termami.

Występują trzy logiczne operatory:

- AND
- OR
- NOT

Operator OR

Wstawienie operatora OR pomiędzy słowa kluczowe powoduje wyświetlenie wyników zawierających przynajmniej jedno z dwóch szukanych słów. Operatora OR używamy wtedy, gdy szukamy terminu dającego się opisać przez kilka słów. Używając OR rozszerzamy wyniki wyszukiwania. Wykorzystujemy szukając bliskoznaczne terminy.



Wyszukane będą dokumenty zawierające przynajmniej jedno ze słów – baton, mars

Poszukiwane terminy	Rezultaty
baton	39,200,000
mars	241,000,000
baton OR mars	282,000,000

Im więcej terminów połączymy w wyszukiwaniu za pomocą OR logicznego, tym więcej otrzymamy rekordów:

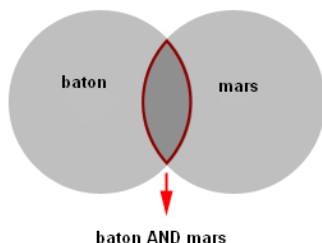
Poszukiwane terminy	Rezultaty
baton	39,200,000
mars	241,000,000
baton OR mars	282,000,000
baton OR mars OR snickers	283,000,000

- baton OR mars
- baton OR mars OR snickers

Operator AND

Operator AND wstawiony pomiędzy słowa kluczowe oznacza ich koniunkcję. Używamy operatora AND jeśli chcemy uzyskać wyniki zawierające obydwa podane słowa. Użycie AND zawęży wyniki wyszukiwania. Jeśli podstawowe słowa kluczowe nie są oddzielane, domyślnym operatorem jest AND.

Wyszukane zostaną dokumenty zawierające słowa baton oraz mars



Poszukiwane terminy	Rezultaty
baton	39,200,000
mars	241,000,000
baton AND mars	1,430,000

Im więcej terminów połączymy w wyszukiwaniu za pomocą AND logicznego, tym mniej otrzymamy rekordów:

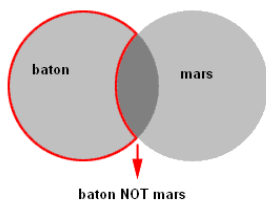
Poszukiwane terminy	Rezultaty
baton	39,200,000
mars	241,000,000
baton AND mars	1,430,000
baton AND mars AND snickers	1,610

- baton AND mars

baton AND mars AND snickers

Operator NOT

Wstawienie słowa NOT pomiędzy słowa kluczowe powoduje wykluczenie następującego po nim słowa lub wyrażenia kluczowego. Używamy operatora NOT, aby wyeliminować szczególne oferty z wyników wyszukiwania.



Wyszukane zostaną takie dokumenty zawierające słowo baton, w których słowo mars nie występuje

Należy być ostrożnym przy tym wyszukiwaniu, żeby mieć pewność, iż w dokumencie którego szukamy na pewno nie występuje term który wykluczaliśmy.

Poszukiwane terminy	Rezultaty
baton	39.200.000
mars	241.000.000
baton NOT mars	37,600,000

Należy być ostrożnym przy tym wyszukiwaniu, żeby mieć pewność, iż w dokumencie którego szukamy na pewno nie występuje term który wykluczaliśmy.

Wyszukujemy dokumenty zawierające term „piła” ale takie, w których nie ma słów „miasto”, „miasta”, „miasteczko”, itd. (* zastępuje tutaj dowolną końcówkę wyrazu). Jest to często przydatne, np. gdy pierwszy wyraz ma wiele znaczeń, a my szukamy konkretnego jego znaczenia w danym dokumencie.

Operator NEAR

Wstawienie zwrotu NEAR (operator sąsiedztwa) pomiędzy słowa kluczowe umożliwia wyszukiwanie terminów znajdujących się w określonej odległości od siebie w dowolnej kolejności. Im są bardziej do siebie zbliżone, tym wyżej dokument zostanie wyświetlony na liście wyników wyszukiwania. NEAR to taki bardziej „restrykcyjny” AND. Bliskość poszukiwanych terminów i jej wzięcie pod uwagę w wyszukiwaniu jest zależne od poszczególnych wyszukiwarek.

Poszukiwane terminy	Rezultaty
baton	39,200,000
mars	241,000,000
baton AND mars	1,430,000
baton NEAR mars	2,150

Ćwiczenia do wykonania. Ćw. I

Wykonaj następujące ćwiczenia za pomocą jednej z wyszukiwarek Google, AltaVista itp:
Wyszukiwanie boolowskie w Internecie:

Query: szukaj informacji o uczelniach wyższych
Query In Boolean: uniwersytet OR politechnika
1. Uniwersytet – wyników:
2. politechnika – wyników:
3. Uniwersytet OR politechnika – wyników:

4. Uzupełnij zdanie:

Im.....(mniej/więcej) termów połączonych operatorem OR w zapytaniu, tym(mniej/więcej) odpowiedzi uzyskamy w odpowiedzi.

Query: szukaj informacji o uniwersytetach w Londynie
Query In Boolean: university AND London
5. University – wyników:
6. London – wyników:
7. University AND London – wyników:

8. Uzupełnij zdanie:

Im.....(mniej/więcej) termów połączonych operatorem AND w zapytaniu, tym(mniej/więcej) odpowiedzi uzyskamy w odpowiedzi.

Query: szukaj informacji o Mazurach ale niech Cię nie interesują strony,
które mówią o wakacjach na mazurach
Query In Boolean: mazury NOT wakacje
9. mazury – wyników:
10. wakacje – wyników:
11. mazury NOT wakacje – wyników:

12. Uzupełnij zdanie:

Im.....(mniej/więcej) termów połączonych operatorem NOT w zapytaniu, tym(mniej/więcej) odpowiedzi uzyskamy w odpowiedzi.

13. mazury – wyników:
14. wakacje – wyników:
15. Mazury NEAR wakacje – **wyników:**
16. Mazury AND wakacje – **wyników:**

LOGIKA BOOLOWSKA W INTERNECIE

Podczas wyszukiwania informacji w Internecie logika boolowska może zostać użyta na trzy różne sposoby, w postaci:

- pełnej logiki boolowskiej z użyciem operatorów logicznych
- implikowanej logiki boolowskiej z kluczowymi słowami wyszukiwania
- ustalonego języka w fill-in template użytkownika

PEŁNA LOGIKA BOOLOWSKA

Używamy operatorów logicznych. Obsługiwana przez wiele przeglądarek. Możemy używać nawiasów w celu wymuszania odpowiedniej kolejności wykonywania operacji logicznych AND, OR, NOT i NEAR

IMPLIKOWANA LOGIKA BOOLOWSKA

Wpisujemy termy posiadające ogólne znaczenie nie używając operatorów logicznych pomiędzy nimi. Zamiast nich stosujemy symbole „+”(AND), „-”(NOT), „|” (OR) (nie wszędzie działa). Przestrzeń pomiędzy słowami kluczowymi reprezentuje OR lub AND logiczne. Większość wyszukiwarek wypełniała ją za pomocą operatora OR, ale z biegiem czasu coraz więcej z nich zaczęło używać jako domyślnego operatora AND logiczne. Aby dowiedzieć się którego operatora dana wyszukiwarka używa jako domyślnego należy zajrzeć do strony pomocy tej wyszukiwarki. Obecnie tylko parę wyszukiwarek używa jako domyślnego OR logicznego.

Pomoc wyszukiwarki google: <http://www.google.pl/intl/pl/help/basics.html#and>

PLB	ILB
baton AND snickers	+baton +snickers
baton NOT snickers	baton -snickers
baton OR snickers	baton snickers

USTALONY JĘZYK W FILL-IN TEMPLATE UŻYTKOWNIKA

Większość wyszukiwarek internetowych oferuje szablon (template), który pozwala użytkownikowi wybrać operatory z menu. Zazwyczaj operatory logiczne wyrażone są za pomocą języka zastępczego raczej niż za pomocą samego operatora.

Przykład google (wyszukiwanie zaawansowane) oraz altavista :

Szukaj	ze wszystkimi słowami	<input type="text"/>
	z wrażeniem	<input type="text"/>
	z którymkolwiek ze słów	<input type="text"/>
	bez słów	<input type="text"/>

Advanced Web Search		Help
Build a query with...		
all of these words:	<input type="text"/>	FIND
this exact phrase:	<input type="text"/>	
any of these words:	<input type="text"/>	
and none of these words	<input type="text"/>	

Porównanie 3 sposobów korzystania z logiki boolowskiej

	Full Boolean	Implied Boolean	Template Terminology
OR	college or university	college university *	any of these words can contain the words should contain the words
AND	poverty and crime	+poverty +crime	all of these words must contain the words
NOT	cats not dogs	cats -dogs	must not contain the words should not contain the words
NEAR,etc	cats near dogs	N/A	near

* w zależności od tego jak przeglądarka interpretuje przestrzeń między wyrazami będzie OR lub AND logiczne(gdy pozostawimy pustą przestrzeń między wyrazami).

Implikowana logika boolowska:

Zamiast „uniwersytet OR politechnika” dajemy po prostu „uniwersytet politechnika”

Zamiast „University AND London” dajemy po prostu „+University +London”


Zamiast “mazury NOT wakacje” dajemy po prostu “mazury -wakacje”

Nie ma odpowiednika NEAR dla implikowanej logiki boolowskiej

Ćw. 17

Uzupełnij tabelę

	Pełna logika Boolowska	Implikowana logika boolowska
OR	college or university	
AND	poverty and crime	
NOT	cats not dogs	
NEAR, etc.	cats near dogs	

ID	Piotr 	Magda 	Sylwia 	Jan 	Bartek 	Szymon 
smak	czekolada	wanilia	wanilia	wiśnia	czekolada	Wanilia
potomstwo	brak	1 syn	Syn i córka	2 córki	Brak	2 synów, 1 córka
region	Środkowy zachód	południe	wschód	południe	zachód	Środkowy zachód
zwierzaki	2 koty	1 ptak	1 ptak, 2 psy	brak	5 gerbils	brak
praca	Bez pracy	Pół etatu	Pełny etat	Bez pracy	Pełny etat	Pół etatu
wiek	20 lat	30 lat	40 lat	30 lat	25 lat	30 lat

18. Kto lubi czekoladę **AND** ma pracę?

- Bartek I Piotr
- Sylwia I Magda
- Bartek
- Sylwia i Magda
- Piotr i Sylwia

19. (**Mieszka na południu OR** środkowym zachodzie) **AND NOT** posiadają zwierzaki **AND** nie pracują

- Piotr I Jan
- Szymon
- Sylwia I Piotr
- Jan
- nikt

20. (**30 lat AND** lubi czekoladę) **OR** (pracuje **AND** mieszka na południu)

- Jan, Piotr I Szymon
- Bartek
- Magda
- Sylwia
- Sylwia I Bartek

21. (**Lubi czekoladę OR** Lubi wiśnie) **AND NOT** posiada zwierzaki **AND** mieszka na Wschodzie

- Piotr i Sylwia
- Szymon I Jan
- Jan
- Magda
- pusty zbiór

22. (**30 lat OR** mieszka na zachodzie) **AND** (lubi czekoladę **OR** lubi wanilię)

- Magda I Bartek
- Szymon I Bartek
- Magda I Szymon I Sylwia
- Magda I Bartek I Szymon
- żadne z powyższych