

# Internet Tutorials

## Boolean Searching on the Internet

### A Primer in Boolean Logic

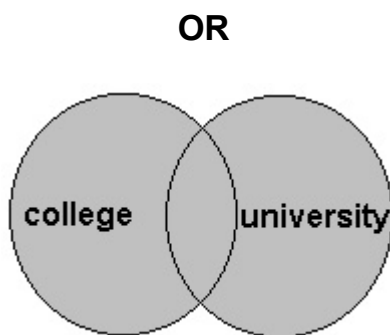
The Internet is a vast computer database. As such, its contents must be searched according to the rules of computer database searching. Much database searching is based on the principles of Boolean logic. Boolean logic refers to the logical relationship among search terms, and is named for the British-born Irish mathematician George Boole.

On Internet search engines, the options for constructing logical relationships among search terms extend beyond the traditional practice of Boolean searching. This will be covered in the section below, [Boolean Searching on the Internet](#).

Boolean logic consists of three logical operators:

- OR
- AND
- NOT

Each operator can be visually described by using Venn diagrams, as shown below.



**college OR university**

**Query:** I would like information about college.

- In this search, we will retrieve records in which AT LEAST ONE of the search terms is present. We are searching on the terms **college** and also **university** since documents containing either of these words might be relevant.

- This is illustrated by:
  - the shaded circle with the word **college** representing all the records that contain the word "college"
  - the shaded circle with the word **university** representing all the records that contain the word "university"
  - the shaded overlap area representing all the records that contain both "college" and "university"

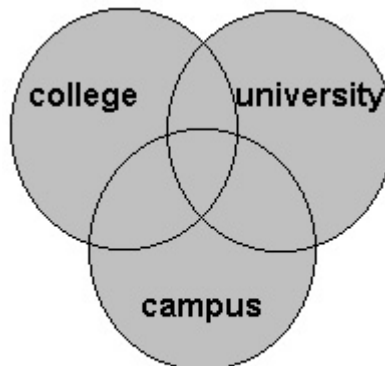
OR logic is most commonly used to search for synonymous terms or concepts.

Here is an example of how OR logic works:

Search terms	Results
college	17,320,770
university	33,685,205
college OR university	33,702,660

OR logic collates the results to retrieve all the unique records containing one term, the other, or both.

The *more* terms or concepts we combine in a search with OR logic, the *more* records we will retrieve.

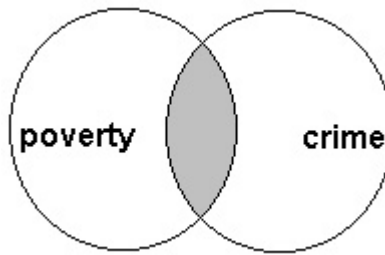


For example:

Search terms	Results
college	17,320,770
university	33,685,205
college OR university	33,702,660

college OR university OR campus

33,703,082

**AND****poverty AND crime**

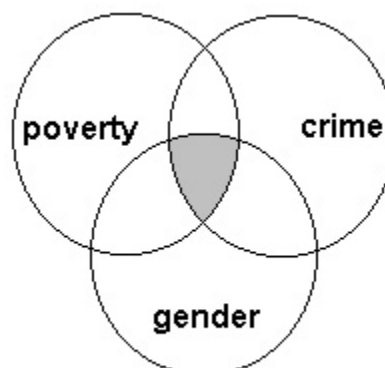
**Query:** I'm interested in the relationship between poverty and crime.

- In this search, we retrieve records in which BOTH of the search terms are present
- This is illustrated by the shaded area overlapping the two circles representing all the records that contain both the word "poverty" and the word "crime"
- Notice how we do not retrieve any records with only "poverty" or only "crime"

Here is an example of how AND logic works:

Search terms	Results
poverty	783,447
crime	2,962,165
poverty AND crime	1,677

The *more* terms or concepts we combine in a search with AND logic, the *fewer* records we will retrieve.

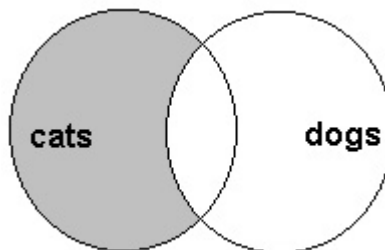


For example:

Search terms	Results
poverty	783,447
crime	2,962,165
poverty AND crime	1,677
poverty AND crime AND gender	76

A few Internet search engines make use of the *proximity operator* NEAR. A proximity operator determines the closeness of terms within the text of a source document. NEAR is a restrictive AND. The closeness of the search terms is determined by the particular search engine. Google defaults to proximity searching by default.

### NOT



**cats NOT dogs**

**Query:** I want information about cats, but I want to avoid anything about dogs.

- In this search, we retrieve records in which ONLY ONE of the terms is present
- This is illustrated by the shaded area with the word **cats** representing all the records containing the word "cats"
- No records are retrieved in which the word "dogs" appears, even if the word "cats" appears there too

Here is an example of how NOT logic works:

Search terms	Results
cats	3,651,252
dogs	4,556,515
cats NOT dogs	81,497

NOT logic excludes records from your search results. Be careful when you use NOT:

the term you do want may be present in an important way in documents that also contain the word you wish to avoid.

## Boolean Searching on the Internet

When you use an Internet search engine, the use of Boolean logic may be manifested in three distinct ways:

1. Full Boolean logic with the use of the logical operators
2. Implied Boolean logic with keyword searching
3. Predetermined language in a user fill-in template

### 1. Full Boolean logic with the use of the logical operators

Many search engines offer the option to do full Boolean searching requiring the use of the Boolean logical operators.

#### Examples:

**Query:** I need information about cats.

**Boolean logic:** OR

**Search:** cats OR felines

**Query:** I'm interested in dyslexia in adults.

**Boolean logic:** AND

**Search:** dyslexia AND adults

**Query:** I'm interested in radiation, but not nuclear radiation.

**Boolean logic:** NOT

**Search:** radiation NOT nuclear

**Query:** I want to learn about cat behavior.

**Boolean logic:** OR, AND

**Search:** (cats OR felines) AND behavior

Note: Use of parentheses in this search is known as *forcing the order of processing*. In this case, we surround the OR words with parentheses so that the search engine will process the two related terms first. Next, the search engine will combine this result with the last part of the search that involves the second concept. Using this method, we are assured that the

semantically-related OR terms are kept together as a logical unit.

## 2. Implied Boolean logic with keyword searching

*Keyword searching* refers to a search type in which you enter terms representing the concepts you wish to retrieve. Boolean operators are not used.

*Implied Boolean logic* refers to a search in which symbols are used to represent Boolean logical operators. In this type of search on the Internet, the *absence* of a symbol is also significant, as the space between keywords defaults to either OR logic or AND logic. Nowadays, most search engines default to AND.

Implied Boolean logic has become so common in Web searching that it may be considered a de facto standard.

### Examples:

**Query:** I need information about cats.

**Boolean logic:** OR

**Search:** [None]

It is extremely rare for a search engine to interpret the space between keywords as the Boolean OR. Rather, the space between keywords is interpreted as AND. To do an OR search, choose either option #1 above (full Boolean logic) or option #3 below (user fill-in template).

**Query:** I'm interested in dyslexia in adults.

**Boolean logic:** AND

**Search:** +dyslexia +adults

**Query:** I'm interested in radiation, but not nuclear radiation.

**Boolean logic:** NOT

**Search:** radiation -nuclear

**Query:** I want to learn about cat behavior.

**Boolean logic:** OR, AND

**Search:** [none]

Since this query involves an OR search, it cannot be done with keyword searching. To conduct this type of search, choose either option #1 above (full Boolean logic) or option #3 below (user fill-in template).

### 3. Predetermined language in a user fill-in template

Some search engines offer a search template which allows the user to choose the Boolean operator from a menu. Usually the logical operator is expressed with substitute language rather than with the operator itself.

#### Examples:

**Query:** I need information about cats

**Boolean logic:** OR

**Search:** Any of these words/Can contain the words/Should contain the words

**Query:** I'm interested in dyslexia in adults.

**Boolean logic:** AND

**Search:** All of these words/Must contain the words

**Query:** I'm interested in radiation, but not nuclear radiation.

**Boolean logic:** NOT

**Search:** Must not contain the words/Should not contain the words

**Query:** I want to learn about cat behavior.

**Boolean logic:** OR, AND

**Search:** Combine options as above if the template allows multiple search statements

### Quick Comparison Chart: Full Boolean vs. Implied Boolean vs. Templates

	Full Boolean	Implied Boolean	Template Terminology
<b>OR</b>	college or university	[rarely available] <i>*see note below</i>	any of these words can contain the words should contain the words
<b>AND</b>	poverty and crime	+poverty +crime	all of these words must contain the words
<b>NOT</b>	cats not dogs	cats -dogs	must not contain the words should not contain the words
<b>NEAR, etc.</b>	cats near dogs	N/A	near

\* Most multi-term search statements will resolve to AND logic at search engines that

use AND as the default. Nowadays most search engines default to AND. Always play it safe, however, and consult the Help files at each site to find out which logic is the default.

### Where to Search: A Selected List

Feature	Search Engine
Boolean operators	<a href="#"><u>AltaVista Advanced Web Search</u></a>   <a href="#"><u>AllTheWeb Advanced Search</u></a>   <a href="#"><u>Dogpile</u></a>   <a href="#"><u>Google</u></a> [OR only]   <a href="#"><u>Ixquick</u></a>
Full Boolean logic with parentheses, e.g., <i>behavior and (cats or felines)</i>	<a href="#"><u>AlltheWeb Advanced Search</u></a>   <a href="#"><u>AltaVista Advanced Web Search</u></a>   <a href="#"><u>Ixquick</u></a>   <a href="#"><u>MSN Search</u></a>
Implied Boolean +/-	Most search engines offer this option
Boolean logic using search form terminology	Most advanced search options offer this, including: <a href="#"><u>AllTheWeb Advanced Search</u></a>   <a href="#"><u>AltaVista Advanced Web Search</u></a>   <a href="#"><u>AOL Advanced Search</u></a>   <a href="#"><u>Ask.com Advanced Search</u></a>   <a href="#"><u>Google Advanced Search</u></a>   <a href="#"><u>MSN Search Search Builder</u></a>   <a href="#"><u>Yahoo Advanced Web Search</u></a>
Proximity operators	<a href="#"><u>Exalead</u></a>   <a href="#"><u>Google</u></a> [by default]   <a href="#"><u>Ixquick</u></a>

A more complete list may be found on  
[How to Choose a Search Engine or Directory](#)

Updated: May 2006

Send comments to [Laura Cohen](#)