

## Becoming an Advanced Searcher

### Sandeep Krishnamurthy

#### *Implicit AND*

Say, you enter these words in your search box-

laptop computer
-----------------

Google will provide with you a list of pages that have both words- laptop and computer. Pages that have one of these two words (i.e., either laptop or computer) will not be returned. This does not mean that the words have to be in the same order in the documents- the word computer may appear on line 1 and the word laptop may appear on line 7.

If you wish to search on an exact phrase, your search will have to look like one of the two search boxes shown below-

"laptop computer"
-------------------

laptop+computer
-----------------

Suppose you are interested in documents that contain either laptop or computer, you could explicitly tell Google to do that by searching in either of these two ways-

laptop OR computer
--------------------

laptop   computer
-------------------

#### *What Google Does Not Include*

Google returns pages that matches the words you enter exactly. If you search for children, it will not provide you with any records that have the word kids in it<sup>1</sup>.

Google ignores some common words called "stop words," e.g., the, on, be, where, and how. Since these words are exceedingly common, they simply slow down the search process.

If your search looks like this-

to be or not to be
--------------------

Google will ignore to, be and or. The results will simply be a search on the word "not". If you really want the links related to Shakespeare, you must use quotes-

"to be or not to be"
----------------------

Alternatively, you could use the \* character- also known as the wildcard. If you search for-

\* purpose \* \* life

you will find records that include information such as "the purpose of your life".

Another way of doing this is to use the "+" term in this way-

+to +be +or not +to +be

Using the + sign automatically over-rides the stop word exclusion and asks Google to explicitly include it.

Want to try something fun? Search for-

to be

Since these are stop words, Google should yield no results, right? Is that really what happens? Is there any case, where Google yields no results?

Google ignores most punctuation and special characters including , . ; ? [ ] ( ) @ / \*. A search that looks like this-

to be, or, not to be

will yield the same results as-

to be or not to be

Google is case insensitive. A search that looks like this-

LAPTOP COMPUTER

will yield the same results as-

laptop computer

### *Related Words*

Google returns pages that match variants of your search terms by using a technique called Automatic Stemming [jargon alert]. Google automatically includes words that are closely related to the one that you entered. If you are searching for computer,

Google will include "computers", "computer's", for instance. You need to explicitly tell Google not to do this.

You are an expert Google user, if you know how to use -[called tilde]. If you include this in your search, it tells Google to look at synonyms for the word. If your search looks like-

~cheap laptop computer

Google will include synonyms for the term cheap- inexpensive, low price etc.

Google gives greater importance to pages that have search terms near each other. Let us say there is a page where the word laptop and computer show up in the first paragraph it is more likely to be served rather than a page where laptop appears on line 1 and computer on line 10.

### *10-Word Limit*

Google limits search queries to 10 words. You cannot enter entire paragraphs in the search box. What this means is that you must identify meaningful keywords and use those to search rather than searching on everything.

### *Order of terms matter*

First run this search-

laptop computer

Then, run this one-

computer laptop

What do you find? The results will be quite different. Google gives higher priority to pages that have search terms in the same order as the search.

### *Excluding terms*

You can ask Google to provide search results that do NOT include a word. Let us say you want to find out information about laptop computers and do not want to include pages that have the word "Dell" or "Toshiba", you can do this by running this search-

laptop computer -Dell

Interestingly, this does not affect the advertised results! So, a company can still make sure its information is displayed by simply purchasing text-ads.

You can experiment with this in this way. First, run a search on any word. Second, run the search again but omit what appears as the first result. You will clearly see how this works. This is how I implemented this-

software

software -APACHE

---

<sup>i</sup> Thank you, Google Guide.