

Chapter 3

Building a Query

3.1 Query Basics

One of the advantages of Google's effective method of finding and ordering pages for you is that even a simple search, such as typing in a couple of words, can produce excellent results. For a basic query, you don't need any fancy logical operators. Just type your keywords into the Google search box and press Enter or click on Google Search. As you consider those keywords, keep these facts in mind:

- *Choose between two and six keywords per search.* Google allows a maximum of 10 words for the search box. Words in excess of 10 will be ignored. Very specific or distinctive words can be searched for individually (such as *monozygote*), but usually at least two words should be considered the minimum. For example: *forensic odontology*. Many general subject searches can be formulated in a phrase of two or perhaps three words. If you want to find more specific information rather than an overview of the subject, you will probably need to add some refining terms. For example, *fingerprint analysis* is a general subject, while *fingerprint analysis reliability* is a more specific aspect of that subject.
- *Include more words to get fewer matches.* Google puts an AND operator between all the keywords you type in, so to narrow a broad search, you can include more words. To find more results from a narrow concept, use fewer words. It's important to remember the automatic AND when considering forms of keywords. Typing the keywords *war warfare warmonger warring* tells Google to look for pages that contain all of those words, not just any of them. Entering various forms of a keyword can be helpful because pages that use several forms of the keywords are likely to be highly relevant. On the other hand, too many variants may result in your missing important pages or in few or no hits at all.
- *Don't worry about capitalization.* Google is not case sensitive. *Congress* and *congress*, *Thomas Edison* and *thomas edison* look the same to Google. On the other hand, feel free to capitalize any words you want.
- *Spell it right.* Google is spelling sensitive. Even though Google will attempt to suggest correct spellings, if you spell your query

words right the first time, your searching will be more efficient and accurate.

- *Use different sets of keywords.* The previous chapter offered some ideas for generating keywords based on forms, related terms, and synonyms. Put those into phrases for individual searches. For example: *flame retardant plastic*, *flame resistant plastic*, and *fire retardant polymer* may bring up useful pages.

Tip 3.1 Repeated Keywords

You can sometimes cause Google to rearrange its results by repeating a keyword. In such a case, Google looks in its index for the keywords closer together or even repeated next to each other and puts those pages higher up in the rankings. For example, try a search on *arson* and then do another search on *arson arson*.

3.2 Order of Appearance

Some search engines assume that keywords you type in first are the most important and should be given more weight than those you type later in your search string. Google does not seem to care. When Google looks at the keywords you have entered, it tries to match results-page words in the same order first and then in other orders. For example, if you want to find a glossary of fishing terms, type in *glossary fishing*. You can also enter *fishing glossary* and get similar results, though Google will return pages with the phrase *fishing glossary* higher than pages with the phrase *glossary of fishing*. Similarly, if you search for *biography Herbert poet George* you will get pages discussing the biography of the poet George Herbert. If you search for *poet George Herbert biography*, you'll get mostly the same pages, with some pages containing the phrase "poet George Herbert" higher up in the results. Remember that Google's Page-Rank™ technology that determines the importance of each page interacts with the location, order, and frequency of words on the pages to produce the ordered results. Thus, think *order of appearance* rather than *order of importance* when constructing your queries. Feel free to shuffle your search phrases around to see what differences in results you get.

Tip 3.2 Punctuation Rules

Google ignores most punctuation in search phrases. You need not enter a period, exclamation point, comma, or question mark. Hyphens between words have an effect on the ordering of results, but Google will return both the hyphenated form and the non-hyphenated form (as in Obsessive Compulsive Disorder versus Obsessive-Compulsive Disorder or Sapir-Whorf Hypothesis versus Sapir Whorf Hypothesis). The one punctuation mark that Google takes most seriously is the apostrophe. A search for *containers* (plural) yields quite different results from a search for *container's* (possessive singular). Google ignores the apostrophe in *containers'* (possessive plural).

Similarly, the search *John O'Hara* includes results with the apostrophe while the search *John OHara* includes results without it. Try searching on *Greshams Law* and *Gresham's Law* and you will see the importance of the apostrophe.

3.3 Stop Words

When you type in an ordinary phrase search, Google will ignore what are known as stop words, those small words like *the*, *with*, and *when*. Stop words: (1) are very common; (2) usually add little or no meaning to the query; (3) do not help differentiate one page from another. When Google ignores a stop word, it tells you with a little message, such as, “**‘With’** is a very common word and was not included in your search.” If one or more ordinary words are important to your search, you have three choices:

- *Don't worry, because they may appear anyway.* The search *dancing wolves* returns pages containing *dancing with wolves*, and a search on *dancing dark* returns pages with *dancing in the dark*. The expressions are so common that they simply appear all connected.
- *Use a plus sign (+) in front of the stop word or words you want Google to include.* For example, if you want to search for Ivan the Terrible, you can force Google to include the stop word *the* by entering *Ivan +the Terrible*. See Chapter 5 for more information about using logical operators in your queries.
- *Enclose your phrase in quotation marks.* If your search phrase contains several stop words that might be ignored, enclose the phrase in quotation marks to create an exact phrase search (all the words in that exact order). For example, try searching first for *I won't put up with that* and then “*I won't put up with that*” to see the difference. See section 3.4 below for more information.

The elimination of stop words makes searching much more accurate and effective because those little words make up a very large proportion of almost any page and therefore do not help distinguish one page from another. The novel *Jane Eyre* by Charlotte Bronte, for example, contains just under 187,000 total words. The ten most common words in the novel (the, I, and, to, a, of, you, in, was, it) account for nearly a fourth of that total. That is, just ten words out of the 12,662 different words Bronte uses account for nearly a quarter of all the words used in the novel. A search on *+the*, which tells Google to look for every page containing the word *the*, returns 3,750,000,000 (three and three quarter billion) pages.

Tip 3.3 Finding Information Tools

Google is great for locating other non-text information besides pictures and drawings. Try some of the following queries or model your own after them and see what you get:

- “*periodic table*”
- “*map of Kenya*”
- “*web cams*”
- “*satellite images*”
- “*stock market graphs*”
- “*normal distribution calculator*”
- “*amortization calculator*”

3.4 Exact Phrase Searches

Usually a search phrase consists mostly of distinctive words. However, several common words together can make up a distinctive phrase, or a search phrase may include both distinctive and common words. If your search phrase includes several common words or is made up entirely of them, use quotation marks around it to perform an exact phrase search. If you simply enter the words *I want to be with a friend*, Google will search on *want* and *friend* and ignore the rest of the words (they are all stop words). You will get pages containing *want* and *friend*. On the other hand, if you use quotation marks to make an exact phrase search, “*I want to be with a friend*” you will get only those pages containing all of those words, in the order you typed them.

Exact phrase searches are useful for more than including stop words. Any time you want an exact pattern of words, instead of simply all the words to appear somewhere on each result page, an exact phrase search is useful. For example, a search on *image pixelation* will return pages where the word *image* and the word *pixelation* both appear, anywhere on the page, in any arrangement. But a search on “*image pixelation*” will return pages where the word *image* immediately precedes the word *pixelation*.

Tip 3.4 Search Within Your Results

A quick way to refine a query and reduce the number of hits is to look for something specific within your search results. At the bottom of the results page, Google offers a hyperlink labeled “Search within results.” Clicking on that brings up another search box, where you can enter a new search term or phrase to narrow down your results. For example, search on *mitogenic rays* and you’ll get about 740 results. Now do a search within those results for *Allison effect* and you’ll get about 14 results, a reduction of 90 percent. Searching within your results is especially useful if you notice that Google has returned pages of interest widely interspersed with pages that you don’t want. Check a few of the useful pages for a common word or phrase and use that to filter out the non-useful ones, or add some of the other search words from your list. Searching within results is also

useful, of course, when you get an enormous number of hits and want to reduce the list.

Use exact phrase searches to find the following items:

- People's names. "*Sherlock Holmes*" or "*George Armstrong Custer*"
- Book or article titles. "*The History of Rasselas*" or "*The Road Less Traveled*"
- Phrase names. "*Boeing 747*" or "*Magnus the Good*"
- Multi-word concepts. "*elasticity of demand*" or "*stratospheric ozone depletion*"
- Famous sayings. "*An oral agreement isn't worth the paper it's written on*" or "*Who's on First*" (Remember Google will search for only the first ten words.)
- Laws, rules, products or other items with letter and number identifiers. "*Epson PowerLite 52C*" or "*Business and Professions Code Section 1600*"
- Asking questions. "*How long do turtles live?*" See the tip below, "One-Stop Explanations" for more information.

You'll notice that Google often returns an exact phrase when you type in a two- or three-word query even without quotation marks. However, that is not guaranteed, and you will also get pages containing the words in relative positions other than the exact phrase, often before further examples of the phrase. And you may miss occurrences of the phrase if it contains stop words. So, to guarantee that you get just what you want, use quotation marks when you want an exact phrase.

Tip 3.5 One-Stop Explanations

To get the definition of a word you can go to an online dictionary or you can try an encyclopedia. However, you can often get several good definitions or explanations by asking Google a question. Some researchers will tell you not to use natural language queries in Google, because that is not how the engine works. However, writing a Google query that resembles a standard question that might be found in a FAQ (frequently asked questions database) often returns excellent results. A FAQ is a list of questions and answers related to a particular topic, covering the most common subjects or questions related to that topic. For example, try some of the following queries and see what you get (be sure to include the quotation marks for an exact phrase search):

- "*What is a transducer?*"
- "*How does a computer virus work?*"
- "*Why is the sky blue?*"
- "*How many square feet are in an acre?*"

In the next chapter, we'll see how you can build a more complex query using Google's Advanced Search Page.

© 2003 by McGraw-Hill/Dushkin, Guilford, CT 06437, A Division of The McGraw-Hill Companies.

Copyright law prohibits the reproduction, storage, or transmission in any form by any means of any portion of this publication without the express written permission of McGraw-Hill/Dushkin, and of the copyright holder (if different) of the part of the publication to be reproduced. The Guidelines for Classroom Copying endorsed by Congress explicitly state that unauthorized copying may not be used to create, to replace, or to substitute for anthologies, compilations, or collective works.