www.searchenginewatch.com

[Back to Article](#)

How Search Engines Rank Web Pages

By [Danny Sullivan](#), Editor

Search for anything using your favorite crawler-based search engine. Nearly instantly, the search engine will sort through the millions of pages it knows about and present you with ones that match your topic. The matches will even be ranked, so that the most relevant ones come first.

Of course, the search engines don't always get it right. Non-relevant pages make it through, and sometimes it may take a little more digging to find what you are looking for. But, by and large, search engines do an amazing job.

As WebCrawler founder Brian Pinkerton puts it, "Imagine walking up to a librarian and saying, 'travel.' They're going to look at you with a blank face."

OK -- a librarian's not really going to stare at you with a vacant expression. Instead, they're going to ask you questions to better understand what you are looking for.

Unfortunately, search engines don't have the ability to ask a few questions to focus your search, as a librarian can. They also can't rely on judgment and past experience to rank web pages, in the way humans can.

So, how do crawler-based search engines go about determining relevancy, when confronted with hundreds of millions of web pages to sort through? They follow a set of rules, known as an algorithm. Exactly how a particular search engine's algorithm works is a closely-kept trade secret. However, all major search engines follow the general rules below.

Location, Location, Location...and Frequency

One of the the main rules in a ranking algorithm involves the location and frequency of keywords on a web page. Call it the location/frequency method, for short.

Remember the librarian mentioned above? They need to find books to match your request of "travel," so it makes sense that they first look at books with travel in the title. Search engines operate the same way. Pages with the search terms appearing in the HTML title tag are often assumed to be more relevant than others to the topic.

Search engines will also check to see if the search keywords appear near the top of a web page, such as in the headline or in the first few paragraphs of text. They assume that any page relevant to the topic will mention those words right from the beginning.

Frequency is the other major factor in how search engines determine relevancy. A search engine will analyze how often keywords appear in relation to other words in a web page. Those with a higher frequency are often deemed more relevant than other web pages.

Spice In The Recipe

Now it's time to qualify the location/frequency method described above. All the major search engines follow it to some degree, in the same way cooks may follow a standard chili recipe. But cooks like to add their own secret ingredients. In the same way, search engines add spice to the location/frequency method. Nobody does it exactly the same, which is one reason why the same search on different search engines produces different results.

To begin with, some search engines index more web pages than others. Some search engines also index web pages more often than others. The result is that no search engine has the exact same collection of web pages to search through. That naturally produces differences, when comparing their results.

Search engines may also penalize pages or exclude them from the index, if they detect search engine "spamming." An example is when a word is repeated hundreds of times on a page, to increase the frequency and propel the page higher in the listings. Search engines watch for common spamming methods in

a variety of ways, including following up on complaints from their users.

Off The Page Factors

Crawler-based search engines have plenty of experience now with webmasters who constantly rewrite their web pages in an attempt to gain better rankings. Some sophisticated webmasters may even go to great lengths to "reverse engineer" the location/frequency systems used by a particular search engine. Because of this, all major search engines now also make use of "off the page" ranking criteria.

Off the page factors are those that a webmasters cannot easily influence. Chief among these is link analysis. By analyzing how pages link to each other, a search engine can both determine what a page is about and whether that page is deemed to be "important" and thus deserving of a ranking boost. In addition, sophisticated techniques are used to screen out attempts by webmasters to build "artificial" links designed to boost their rankings.

Another off the page factor is clickthrough measurement. In short, this means that a search engine may watch what results someone selects for a particular search, then eventually drop high-ranking pages that aren't attracting clicks, while promoting lower-ranking pages that do pull in visitors. As with link analysis, systems are used to compensate for artificial links generated by eager webmasters.

Learning More

The Search Engine Features Chart has a section that summarizes key areas of how crawler-based search engines rank web pages. The Search Engine Placement Tips page also summarizes key tips that will help you improve the relevancy of your pages with crawler-based search engines.

Search Engine Watch members have access to the How Search Engines Work section. This section provides detailed information about how each major search engine gathers its listings and an additional tips on enhancing your position in their results. Learn more about becoming a Search Engine Watch member and the many benefits members receive by visiting the Membership

page.

Next: Search Engine Placement Tips
Previous - Beginning

July 31, 2003

Search Engine Watch
www.searchenginewatch.com
Danny Sullivan, Editor; Chris Sherman, Associate Editor

Having trouble? Tell us about technical questions or bug reports.
Let us know about newsletter problems or issues.

Jupitermedia is publisher of the ClickZ Network.

JupiterWeb networks:

Search JupiterWeb: Find

Jupitermedia Corporation has four divisions:
JupiterWeb, JupiterResearch, JupiterEvents and JupiterImages

Copyright 2004 Jupitermedia Corporation All Rights Reserved.
Legal Notices, Licensing, Reprints, & Permissions, Privacy Policy.

Jupitermedia Corporate Info | Newsletters | Tech Jobs | E-mail Offers