



An Introduction to Search Engine Optimization

Search engines are one of the primary ways that Internet users find Web sites. That's why a Web site with good search engine listings may see a dramatic increase in traffic. Everyone wants those good listings. Unfortunately, many Web sites appear poorly in search engine rankings or may not be listed at all because they fail to consider how search engines work.

In particular, submitting to search engines is only part of the challenge of getting good search engine positioning. It's also important to prepare a Web site through "search engine optimization." Search engine optimization means ensuring that your web pages are accessible to search engines and focused in ways that help improve the chances they will be found.

How Search Engines Work

The term "search engine" is often used generically to describe both crawler-based search engines and human-powered directories. These two types of search engines gather their listings in radically different ways.

Crawler-Based Search Engines

Crawler-based search engines, such as Google, create their listings automatically. They "crawl" or "spider" the web, then people search through what they have found. If you change your web pages, crawler-based search engines eventually find these changes, and that can affect how you are listed. Page titles, body copy and other elements all play a role.

Human-Powered Directories

A human-powered directory, such as the Open Directory, depends on humans for its listings. You submit a short description to the directory for your entire site, or editors write one for sites they review. A search looks for matches only in the descriptions submitted.



Changing your web pages has no effect on your listing. Things that are useful for improving a listing with a search engine have nothing to do with improving a listing in a directory. The only exception is that a good site, with good content, might be more likely to get reviewed for free than a poor site.

The Parts of a Crawler-Based Search Engine

Crawler-based search engines have three major elements. First is the spider, also called the crawler. The spider visits a web page, reads it, and then follows links to other pages within the site. This is what it means when someone refers to a site being "spidered" or "crawled." The spider returns to the site on a regular basis, such as every month or two, to look for changes.

Everything the spider finds goes into the second part of the search engine, the index. The index, sometimes called the catalog, is like a giant book containing a copy of every web page that the spider finds. If a web page changes, then this book is updated with new information.

Sometimes it can take a while for new pages or changes that the spider finds to be added to the index. Thus, a web page may have been "spidered" but not yet "indexed." Until it is indexed -- added to the index -- it is not available to those searching with the search engine.

Search engine software is the third part of a search engine. This is the program that sifts through the millions of pages recorded in the index to find matches to a search and rank them in order of what it believes is most relevant.

Major Search Engines: The Same, But Different

All crawler-based search engines have the basic parts described above, but there are differences in how these parts are tuned. That is why the same search on different search engines often produces different results. Now let's look more about how crawler-based search engine rank the listings that they gather.



How Search Engines Rank Web Pages

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 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 click-through 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.

Partner with TriDigital for your Search Engine Marketing Needs

TriDigital Inc. develops and manages Search Engine Marketing (SEM) Programs for businesses and organizations who want to increase qualified traffic to their Web site. Our team has been involved in interactive marketing and e-commerce since its inception and has deep knowledge of current online marketing strategies. The commercialization of Internet search is the most significant recent development on the Web, and TriDigital's SEM Programs are designed to help our clients develop a cost-effective strategy and take advantage of this exciting new marketing opportunity.



Paid placement of "sponsored listings" on search engines allows businesses to reach a huge national audience of consumers searching online for specific products and services. The ability to dynamically bid for specific "keywords" on a "pay per click" (PPC) basis results in qualified, measurable traffic at a predictable cost. With proper tracking and analysis, advertisers can continuously improve their campaigns and the return on their online marketing investment.

TriDigital can perform Search Engine Optimization of your site to improve and track your rankings on unpaid "editorial" listings of the major search engines including Google, Yahoo, AOL, and MSN.

TriDigital offers comprehensive, cost-effective Search Engine Marketing Programs to help our clients acquire new customers and grow their businesses.

Partner with TriDigital. Make things happen.

This report was compiled by TriDigital Inc. from a series of articles appearing on SearchEngineWatch.com.