
SEARCH MARKETING

Date: May 1, 2004

PREFACE

Let me acknowledge right off the bat that I am not a Search Marketing guru. As a "Jack-of-All-Trades" and master of none, I leave this specialization to the experts. In this layman's opinion, that distinction falls to Danny Sullivan, editor of Search Engine Watch.

The information below represents my best attempt to copy and paste, summarize and synthesize the vast wealth of knowledge that he imparts on www.searchenginewatch.com. Because Search Engines are continually changing their algorithms and repackaging their service offerings, I highly recommend periodically visiting this website for the latest school of thought on this technology segment.

INTRODUCTION

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" (SEO), which means ensuring that your web pages are accessible to search engines and focused in ways that help improve the chances they will be found.

SEARCH ENGINE TYPES

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.

There are three major directories:

1. Open Directory (www.dmoz.org)
2. Yahoo (www.yahoo.com)
3. LookSmart (www.looksmart.com)

Crawler-Based Search Engines

Crawler-based search engines have three major elements:

1. Spider

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. The spider returns to the site on a regular basis, such as every month or two, to look for changes.

2. Index

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, it is not available to those searching with the search engine.

3. Search engine software

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.

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.

There are four major crawler-based search engines:

1. Google (www.google.com)
2. All the Web (www.alltheweb.com)
3. Inktomi (www.inktomi.com)
4. Teoma (www.teoma.com)

HOW SEARCH ENGINES RANK WEB PAGES

Crawler-based search engines determine relevancy by following 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:

1. Location/Frequency Method

One of the main rules in a ranking algorithm involves the location and frequency of keywords on a web page. 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.

2. Spice in the Recipe

All the major search engines follow the Location/Frequency Method 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.

3. 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 that webmasters cannot easily influence.

A. 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.

B. Click Through Measurement

In short, this means that a search engine may watch what results someone selects for a particular search, and 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.

SEARCH ENGINE RESULTS CHART

Major search engines generally provide listings from a variety of sources, which they may get from third-party search providers or through their own efforts.

1. Type of Main Results

While search engines display results from many different sources, usually the results from one particular source will be most dominant. These are considered to be the "main" results for that search engine. For example, in a search at Google, the main results are typically editorial listings that come from Google having crawled the web. Types include:

- A. **Crawler** - the main results are compiled by having crawled the web
- B. **Human** - the main results come from listings compiled by human editors
- C. **Paid** - the main results come from paid listings

2. Provider of Main Results

Some search engines gather their own listings for the main results they display. For example, Google crawls the web itself for the main results it shows. Other search engines use third-party search providers for their results. For instance, the main search results at AOL come from Google's crawler-based listings, rather than from work inside AOL.

3. Paid Results

Every major search engine has paid listings that are also presented alongside its editorial results. For example, Overture provides paid listings to many different partners.

4. Directory, Backup & Other Results

Most search engines, where the main results come from crawling the web, will also provide human-powered "directory" results in some way. For example, in a search at Google, "category" links that lead to human-compiled information often appear at the very top of the search results page.

For search engines where the main results come from human work, it's common for them to have a "backup" or "fallthrough" partnership with a crawler-based search engine. For example, if a search at MSN Search fails to find a match in information from LookSmart, then matches from Yahoo-owned Inktomi provide answers. This provides backup against having no matches at all.

Search Engine	Type of Main Results	Provider of Main Results	Paid Results	Directory &/or Backup Results
AllTheWeb	Crawler	AllTheWeb (Overture-owned)	Overture	n/a
AltaVista	Crawler	AltaVista (Overture-owned)	Overture	LookSmart
AOL Search	Crawler	Google	Google	Open Directory
Ask Jeeves	Crawler	Teoma (Ask-owned)	Google	Open Directory
Google	Crawler	Google	Google	Open Directory
HotBot	Crawler	Inktomi (Yahoo-owned)	Overture	Also available: AllTheWeb, Google, Teoma
LookSmart	Human	LookSmart/Zeal	LookSmart	Backup from Inktomi
Lycos	Crawler	AllTheWeb (Overture-owned)	Overture	Open Directory
MSN Search	Human	LookSmart/Zeal	Overture	Backup from Inktomi
Netscape	Crawler	Google	Google	Open Directory
Overture	Paid	Overture	Overture	Backup from Inktomi

Open Directory	Human	Open Directory	n/a	n/a
Teoma	Crawler	Teoma	Google	n/a
Yahoo	Crawler	Google	Overture	Yahoo

SEARCH ENGINE PLACEMENT TIPS

Queries on a crawler-based search engines often turn up thousands or even millions of matching web pages. In many cases, only the 10 most "relevant" matches are displayed on the first page. Naturally, anyone who runs a web site wants to be in the "top ten" results. This is because most users will find a result they like in the top ten. Being listed 11 or beyond means that many people may miss your web site.

The tips below will help you come closer to this goal:

1. Pick Keywords

How do you think people will search for your web page? The words you imagine them typing into the search box are your target keywords.

Each page in your web site will have different target keywords that reflect the page's content. For example, say you have another page about the history of stamps. Then "stamp history" might be your keywords for that page.

Your target keywords should always be at least two or more words long. Usually, too many sites will be relevant for a single word, such as "stamps." This "competition" means your odds of success are lower. Pick phrases of two or more words, and you'll have a better shot at success.

2. Position Keywords

Make sure your target keywords appear in the crucial locations on your web pages. The page's HTML title tag is most important. Failure to put target keywords in the title tag is the main reason why perfectly relevant web pages may be poorly ranked.

Build your titles around the top two or three phrases that you would like the page to be found for. The titles should be relatively short and attractive like newspaper headlines for your pages. They appear in search engine listings, and a short, attractive title may help make users click through to your site.

Search engines also like pages where keywords appear "high" on the page. To accommodate them, use your target keywords for your page headline, if possible. Have them also appear in the first paragraphs of your web page.

Keep in mind that tables can "push" your text further down the page, making keywords less relevant because they appear lower on the page.

Large sections of JavaScript can also have the same effect as tables. The search engine reads this information first, which causes the normal HTML text to appear lower on the page. Place your script further down on the page, if possible.

3. Relevant Content

Changing your page titles is not necessarily going to help your page do well for your target keywords if the page has nothing to do with the topic. Your keywords need to be reflected in the page's content.

In particular, that means you need HTML text on your page. Sometimes sites present large sections of copy via graphics. It looks pretty, but search engines can't read those graphics. That means they miss out on text that might make your site more relevant. Some of the search engines will index ALT text and comment information. But to be safe, use HTML text whenever possible. Some of your human visitors will appreciate it, also.

Finally, consider "expanding" your text references, where appropriate. For example, a stamp collecting page might have references to "collectors" and "collecting." Expanding these references to "stamp collectors" and "stamp collecting" reinforces your strategic keywords in a legitimate and natural manner. Your page really is about stamp collecting, but edits may have reduced its relevancy unintentionally.

4. HTML Links

Often, designers create only image map links from the home page to inside pages. A search engine that can't follow these links won't be able to get "inside" the site. Unfortunately, the most descriptive, relevant pages are often inside pages rather than the home page.

Solve this problem by adding some HTML hyperlinks to the home page that lead to major inside pages or sections of your web site. This is something that will help some of your human visitors, also. Put them down at the bottom of the page. The search engine will find them and follow them.

Also consider making a site map page with text links to everything in your web site. You can submit this page, which will help the search engines locate pages within your web site.

Finally, be sure you do a good job of linking internally between your pages. If you naturally point to different pages from within your site, you increase the odds that search engines will follow links and find more of your web site.

5. Frames Can Kill

Some of the major search engines cannot follow frame links. Make sure there is an alternative method for them to enter and index your site, either through Meta tags or smart design.

6. Dynamic Door Blocks

Generating pages via CGI or database-delivery? Expect that some of the search engines won't be able to index them. Consider creating static pages whenever possible, perhaps using the database to update the pages, not to generate them on the fly. Also, avoid symbols in your URLs, especially the ? symbol. Search engines tend to choke on it.

7. Build Links

Every major search engine uses link analysis as part of their ranking algorithms. This is done because it's very difficult for webmasters to "fake" good links, in the way they might try to spam search engines by manipulating the words on their web pages. As a result, link analysis gives search engines a useful means of determining which pages are good for particular topics.

By building links, you can help improve how well your pages do in link analysis systems. The key is understanding that link analysis is not about "popularity." In other words, it's not an issue of getting



lots of links from anywhere. Instead, you want links from good web pages that are related to the topics you want to be found for.

Here's the simple means to find those good links. Go to the major search engines. Search for your target keywords. Look at the pages that appear in the top results. Now visit those pages and ask the site owners if they will link to you. Not everyone will, especially sites that are extremely competitive with you. However, there will be non-competitive sites that will link to you -- especially if you offer to link back.

By searching for your target keywords, you'll find the pages that the search engines themselves are telling you are good, as evidenced by the fact that they rank well. Hence, links from these pages are more important -- and important for the terms you are interested in -- than links from other pages. In addition, if these pages are top ranked, and then they are likely to be receiving many visitors. Thus, if you can gain links from them, you might receive some visitors who initially go to those pages.

8. Submit Your Key Pages

Most search engines will index the other pages from your web site by following links from a page you submit to them. But sometimes they miss, so it's good to submit the top two or three pages that best summarize your web site.

Don't trust the submission process to automated programs and services. Some of them are excellent, but the major search engines are too important. There aren't that many, so submit manually, so that you can see if there are any problems reported.

Also, don't bother submitting more than the top two or three pages. It doesn't speed up the process. Submitting alternative pages is only insurance. In case the search engine has trouble reaching one of the pages, you've covered yourself by giving it another page from which to begin its crawl of your site.

Be patient. It can take up to a month to two months for your "non-submitted" pages to appear in a search engine, and some search engines may not list every page from your site.

9. Verify and Maintain Your Listing

Once your pages are listed in a search engine, monitor your listing every week or two. Strange things happen. Pages disappear from catalogs. Links go screwy. Watch for trouble, and resubmit if you spot it.

Resubmit your site any time you make significant changes. Search engines should revisit on a regular schedule. However, some search engines have grown smart enough to realize some sites only change content once or twice a year, so they may visit less often. Resubmitting after major changes will help ensure that your site's content is kept current.

HOW TO USE HTML META TAGS

Meta tags have never been a guaranteed way to gain a top ranking on crawler-based search engines. Today, the most valuable feature they offer the web site owner is the ability to control to some degree how their web pages are described by some search engines. They also offer the ability to prevent pages from being indexed at all.

Meta Tag Overview

Meta tags are information inserted into the "head" area of your web pages. Other than the title tag, information in the head area of your web pages is not seen by those viewing your pages in browsers. Instead, Meta information in this area is used to communicate information that a human visitor may not be concerned with. Meta tags, for example, can tell a browser what "character set" to use or whether a web page has self-rated itself in terms of adult content.

Here is a typical Meta tag listing of a web page:

```
<HEAD>  
<TITLE>Stamp Collecting World</TITLE>  
<META name="description" content="Everything you wanted to know  
about stamps, from prices to history.">  
<META name="keywords" content="stamps, stamp collecting,  
stamp history, prices, stamps for sale">  
</HEAD>
```

Title Tag

The HTML title tag isn't really a Meta tag, but it's worth discussing in relation to them. Whatever text you place in the title tag (between the TITLE and /TITLE portions as shown in the example) will appear in the reverse bar of someone's browser when they view the web page. For instance, within the title tag of this page that you are reading is this text:

How to Use HTML Meta Tags

If you look at the reverse bar in your browser, then you should see that text being used, similar to this:



Some browsers also supplement whatever you put in the title tag by adding their own name, as you can see Microsoft's Internet Explorer doing in the picture above.

The title tag is also used as the words to describe your page when someone adds it to their "Favorites" or "Bookmarks" lists. For instance, if you added this page to your Favorites in Internet Explorer, it would show up like this:



The title tag is crucial for search engines. The text you use in the title tag is one of the most important factors in how a search engine may decide to rank your web. In addition, all major crawlers will use the text of your title tag as the text they use for the title of your page in your listings.

Meta Description Tag

The Meta description tag allows you to influence the description of your page in the crawlers that support the tag. The text you want to be shown as your description goes between the quotation marks after the "content=" portion of the tag (generally, 200 to 250 characters may be indexed, though only a smaller portion of this amount may be displayed).

Search engines treat this differently. For example, Google ignores the Meta description tag and instead will automatically generate its own description for this page. Others may support it partially.

Meta Keywords Tag

The Meta keywords tag allows you to provide additional text for crawler-based search engines to index along with your body copy. However, for most major crawlers, it doesn't matter because most now ignore the tag.

The Meta keywords tag is *sometimes* useful as a way to reinforce the terms you think a page is important for ON THE FEW CRAWLERS THAT SUPPORT IT. For instance, if you had a page about stamp collecting -- AND you say the words stamp collecting at various places in your body copy -- then mentioning the words "stamp collecting" in the meta keywords tag MIGHT help boost your page a bit higher for those words.

Remember, if you don't use the words "stamp collecting" on the page at all, then just adding them to the Meta keywords tag is extremely unlikely to help the page do well for the term. The text in the Meta keywords tag, FOR THE FEW CRAWLERS THAT SUPPORT IT, works in conjunction with the text in your body copy.