

Inspecting Your Web Site for Spam

by Jim Hedger, SEO Manager, [StepForth Search Engine Placement](#) Inc.

Spam is an overused word. Describing everything from luncheon "meat" foodstuffs to the flood of advertising Email we all receive each day, the simple, four-letter word is as widely used as many other deceptively descriptive four-letter words are. Suffice it to say that SPAM is not counted amongst the most honorable words in the English language. The IT world has forced the English language to evolve very quickly as old words are applied against new ideas. The word "Spam" has become synonymous with words such as "junk" and "waste". In the search engine optimization field, the word Spam is used to describe illegal techniques used to try to fool search engine spiders. This section will outline a few of the more frequent Spam techniques used by webmasters or unscrupulous SEO practitioners. Check your websites for any incidents of Spam and do your best to get rid of them as they could seriously damage your search engine rankings.

Hidden Text

Some webmasters hide text by making it the same color as the background of the page they are working on. The thinking is that one can put a collection of keywords on their website that are "invisible" to most site viewers but will be perceived and recorded by search engine spiders. This technique actually worked on first generation search engines but was heavily exploited by the adult entertainment industry, forcing search tools to make this technique totally illegal and punishable by banishment. The easiest way to find hidden text on a site is to hold down the left mouse button and drag your mouse across the entire body of the page you are looking at. This will highlight all text and images on the page, including any text that may be the same color as the background. If you find hidden text on your site, delete it or make it visible by changing the text color.

Off Topic or Redundant Text

Websites exist to spread messages about specific things, be they products, people or politics. If your site uses text that does not relate to the topic of the site in any way, it will likely be considered Spam. If this type of Spam appears on your site too often, your site will likely be considered banned.

Keyword Repetition

Meta-Tags

Any excessive repetition of specific words in the keywords meta-tag will be considered spam and will likely cause a website to rapidly lose position on search engine indices. A common rule of meta-tags is to only use the same word a maximum of three times in the Keywords tag.

Body Text

While there is no common rule for the frequency of use for keywords in the body text, it is important to note that each word used on your site must be relevant to the topic of your site and must be used in a proper context. In other words, you can't just bulk a number of words in the lower section of your website and hope search engines reward your efforts with high rankings. Unless each word used has a good excuse for being there, it will likely be considered spam.

Transparent or Hidden Links

Many search engine spiders work by following all links found on any page they come across. A technique that once produced strong results was to hide links by using a tiny, invisible image link, generally phrased as a 1 pixel X 1 pixel dot. This technique simply doesn't work anymore and will likely get a site using it banned for spamming. Unfortunately, this technique is a bit more difficult to find unless you know what you're looking for. The easiest way to look for a hidden image link is to examine the source-code of your website by clicking on the word View (File Edit View Insert...) in the upper text menu on your browser. From the drop-down menu, choose Source. This will bring up a notepad document with the source-code of the page. Go to the word Edit and choose Find. Enter the following information into the dialog box that opens, "width="1" height="1" border="0"". If the Find feature offers anything up in return, chances are you should take a close look at it. If it is a link to another URL, you should immediately remove it.

Pages Generated by Mechanical Means

There are dozens of instant page creators on the market today. Most of them are junk and produce pages that are easily spotted by search engine spiders. More often than not, an instant solution is going to get you in trouble as search engines are deluged with sites that offer no real new information but are only there in order to try to manipulate search engine listings. Website design is becoming so simple that the use of mechanically generated sites is unnecessary.

Cloaking

Cloaking is a technique that offers one set of data to search engine spiders and another set of data to live-human users. Cloaking is a highly advanced technique and is not likely to be used without the knowledge of the website owner or Webmaster. If you think cloaking techniques are being used on your website, [ask a SEO professional](#) to evaluate and advise.

Duplicate Content

With the advent of affiliate marketing schemes, there are tens of thousands of sites with duplicate content out there. As they are in the business of providing accurate and relevant results, search engines do not like sites that repeat what another site has already said. If your site is one of ten thousand duplicate sites, chances are you're not going to rank well anyway, especially since your site does not get to remain in the search engine's database for very long.