

Search Engines and Spam

By Daniel Bazac

Search Engines and Spam

Daniel Bazac
danielbazac@hotmail.com

Web Design in New York

http://www.web-design-in-new-york.com/search_engines_and_spam_1.html

What is Spam?

Spam (for Search Engines) refers to the use of any illegal technique to improve a page' ranking in the Search Engine Results Pages (SERPs). Who decides what 's illegal? The SEs, who else?

Tons of articles have been written about search engines and spam. Why the need for another article? Because, believe it or not, at the end of 2001 MOST of the major search engines are still vulnerable to MANY unethical techniques used by malicious Webmasters.

Spam has been known for a long time. For years, Webmasters tried to cheat the search engines. Today, many SEs say they know all the tricks and penalize those pages. As you will see, most SEs only SAY that they will punish spammers, but in fact they DON'T.

What is Search Engine Optimization?

Search Engine Optimization, Placement, Positioning, Ranking, or whatever you want to call it, is the process of designing a Web page that can be easily indexed by the SEs and receive high rankings in the SERPs.

My Experience with Search Engines

Recently, using some major U.S. SEs, I made a search for a Web design company in New York. Let me to share with you my highly disappointing experience.

The search engine results pages (SERPs) in most of the search engines were full of Web pages that used one or more spamming techniques. The most frequently encountered illegal strategies used by many SEs were:

- Keyword stuffing
- Page redirect
- Mirror domains

Keyword Stuffing

Keyword stuffing is the adding of many relevant, sometimes irrelevant, words in the:

Keywords META tag

One of the criteria some search engines use to rank pages in the results list is the presence of relevant words in the Keywords META tag. From use to abuse is a little step, so some webmasters place many words, repeating them many times in the hope that the page will rank higher. The record was a page that had 1,150 words (no mistake, WORDS not characters) in the Keyword Meta tag. The word "design" was repeated 209 times. Unbelievable, right?

What are Search Engines' positions regarding this spamming technique? One of the Web's most important Search Engine, in their submission guidelines, states that it will "exclude submissions" with "excessive keywords". It looks like 1,150 words are not "excessive" enough, because THAT SE and many others index the page.

Besides stuffing words in the Keywords META tags, some Webmasters also add lots of words in the:

Visible page

Another criteria some search engines rank sites in the results list is the so-called "word frequency": the more times a word is repeated in the content of a page, the higher are the chances it will be near the top in the results list.

Some Webmasters often abuse these criteria by placing words or phrases- most of the times- at the bottom of the page, many, MANY times. Do you recall the site that has 1,150 words in the Keywords Meta tags? Well, the same site just "copied and pasted" those words in the copy of the page.

You might think that it's weird to see a bunch of words inserted in the page. Well, Webmasters have found two methods to work around this problem: "tiny text" and "hidden- or invisible- text".

"Tiny text" means that the Webmaster places the text in a very small size, most of the time hardly legible. "Invisible text" means the use of the same color for text AND for background, so users will NOT see the words but SE's spider WILL see, so the page will rank higher than it deserves.

The Search Engines' position regarding this spamming technique? One of the Web's most important Search Engine, in its guidelines writes: "We must sometimes exclude submissions" of "pages with text that is not easily read, either because it is too small or is obscured by the background of the page". Another SE's will "significantly downgrade a page's ranking" ... "if words cannot be read due to their small size or color".

Despite these statements, both SEs, as well as other SEs, index pages with text in the same color as the background. To illustrate the total lack of respect for users with little fear for SEs, one Web page has a whopping 936 words in BLACK text on a WHITE background. Incredible!

Page Redirect

One page can be redirected either by using the "refresh" META tag or using cloaking techniques. Why the need to redirect a page? One of the legitimate reasons to redirect a page is having a new URL for the site. But some Webmasters abuse page redirection to obtain higher rankings for their pages.

How does it work?

Using the "refresh" META tag:

This technique consists in building 2 pages: one, highly "optimized" - read "spammed"- with many, MANY words in the KW and Description META tag and also in the Title tag. Most of the time the copy of the page is also highly "optimized" with KW stuffing- many times with invisible text.

These pages called: "Doorways Pages", "Gateway", "Entry" or "Bridge" pages most of the time display a "Click here to enter the site" or simply "Enter" or sometimes "select Flash or HTML" message on the entire page. Using the "refresh" META tag the Webmaster will "redirect" the user to a second page, which will appear after a determined number of seconds after seeing the first page. If the time is set to "0" (zero) seconds, the viewer will NOT see the first page, but directly the second one.

Why two pages? The first one shown to SEs as highly "optimized" will help the page rank very high in the SERPSs, by cheating the SEs. The second one is "nicer", not too much spam, a good page for viewers.

Cloaking technique is probably the most controversial spamming technique.

It also uses 2 pages, one for the SE and another for the viewer. The big difference between this technique and the one above is that in the first case it is possible to see the code of the first page (shown to SEs), but in the second case it is not.

Hiding the Meta data (especially keywords) in the code can provide a huge advantage in this current highly competitive market, as there are known cases of Webmasters that stole a competitor's keywords in an attempt to rank higher.

What are search engines' positions regarding redirecting? Do not submit any site with an address that redirects to another address" or "Your site cannot mirror or redirect to another Web site" or we "may permanently ban from our index any sites or authors who engage in cloaking to distort their search rankings" or simply: "Don't cloak".

Pretty clear, right? Then why some Webmasters insist in using cloaking techniques? To hide their pages' code from prying eyes. Bad guys are not afraid to spam- most of the time, SEs only SAY: "don't do that". If you do, there will be NO- or sometimes very little- punishment.

Mirror domains

Consists in building hundreds or even thousands of pages with SAME content but with altogether DIFFERENT URLs (Web addresses). The advantage is clear: by finding the "right" tricks to cheat the SE algorithm one can "dominate" the SERPSs with pages one after another. One company had 62 pages in the Top 100 results list.

What are the search engines positions regarding this spamming technique? "Do not submit mirror sites". "Your site cannot mirror or redirect to another Web site". "Do not submit" . . . "the same pages from multiple domains". But in practice most of the SEs are vulnerable to this technique.

Questions for:

Search Engines

When are you going to be SERIOUS about your job?

When will you PUNISH the spammers? Have CLEAN indexes? It is not difficult. A dialog with the Webmasters community might result in opinions such as THIS (<http://www.webmasterworld.com/forum5/830.htm>) one:

"I can't understand why the search engines aren't professional enough to put their anti-spam efforts into a detailed agreement and have anyone that wants to be listed sign their agreement. Such an agreement would spell out very clearly what is and what isn't allowed. Anyone breaking the rules would be subject to specific penalties or banning, but would be notified and have a chance to fix the problem especially if the infraction was not too serious".

Clear enough?

Why do you accept advertising from sites which use spamming techniques? So, if a page cannot achieve high rankings, for few buck it becomes "featured" site or listing and voilà!, on the top of the list. Is this the latest trick for a poor site to be listed high? Now, on one hand, in your guidelines you tell Webmasters not to spam, on the other hand, if a spammer PAYS, that's no problem. Good bye relevancy, hello profits! Here is a comment by a high-ranking official from a major SE: "the more we take payment for listings, the more you'll get great results" Excuse me? "Big" pockets' sites are NOT always more relevant than "poor" pockets' sites.

I know I'm not the first (<http://www.commercialalert.org/releases/searchengine.html>) one to ask but, when are you going to make a CLEAR distinction between paid listings and real results? The user needs to make an informed decision. The user needs to be helped, not confused.

Webmasters

Why use techniques seen by the SEs as spam? Why take the risk of having pages penalized, or even permanently banned, from indexes? Two months of glory, then a new domain, blacklisted again- poor way to market any business.

Think it's impossible to get higher rankings WITHOUT using spam? My answer is NO. Do your homework, read thoroughly and abide to SEs guidelines, cross your fingers and you'll be # 1.

Have you ever seen one of your page's ranking lower than a page which uses spamming techniques? I bet you did. Now, how did you react? Did you report it to the SE or you said "nah, that's no use anyway"? Or perhaps you thought it's not nice to turn in the Webmaster. OK, let me ask you something: would you like, that while waiting in line to buy a ticket to a movie, somebody, who comes AFTER you, go in the front and buy a ticket BEFORE you do?

Or perhaps THIS (http://www.globalserve.net/~iwb/search_engine/spamdex.html) is the solution?

A site will be nominated and posted on a listserv which any member can second then the offender will be notified of his conviction by the spamdex police.

They will have a week to clean it up and re-submit and remove the offending listing from the search engines.

After a week the offender will be reported to the search engines. The search engines will have a month to act or they will be added to the list as an accomplice to the activity.

Huh? What do you think?

Site Owners

When are you going to wake up? What am I talking about? Well, I'm talking about a report, which says, "nearly 46 percent of the marketers surveyed said they allocate less than 0.5 percent of their annual marketing budgets on search engine optimization (SEO) services". Now, read it again, and again.

So, you spend 95,5 % on gizmos such as Flash and Java Scripts and less than 1% for SEO. Let me tell you something: bells and whistles will never help a page rank high in SERPs. NEVER. In fact it will work against.

Which are your rankings, by the way? When was the last time you checked, if ever? Are you in the Top 30 results for your strategic keywords when the user runs a search? You're not? Practically you don't exist for your prospects. Still wonder why so many .com CLOSE? Are you going to be next?

Do you need a Ph.D. to realize that the SEO- your Web site "salesman"- is more necessary than ever? One of these days you'll discover the power of search engine marketing: the cheapest yet most effective way to promote your site to more than 300 million prospects. Bright future for SEO. I can see a bright future for the SEO, don't you?

When are you going to understand that SEOs is NOT only optimizing the Keyword Meta Tag? SEO is a highly specialized, time-consuming and sometimes expensive but absolutely NECESSARY task. It's both an art and a science to position a Web page near the top in SERPs. So, forget about your in-house SEO "expert", shop around and find a reputable SEO firm or consultant.

Dos and don'ts when you talk to the SEO guy:

If a Web Marketer suggest you to get rid of that flashy or framed Home Page, just DO IT! Or, don't blame HIM for the consequences.

If a SEO says something about cloaking and doorways, run, and I mean it.

Don't even think telling a SEOs, that you'll pay him, AFTER the pages show up in the search engines. Enough said.

Don't be a kid by asking him to GUARANTEE you Top 10 positions! No honest SEO will do it. The only thing a SEO can guarantee is an increase of the number of pages indexed and an increase of the current rankings. With some luck your sales will sky rocket.

Last thing: please, avoid those submission tools and their hype: "submit your site to 500,000 SEs for \$24.95". Submit your Web site by hand. Period.

IN CONCLUSION:

Since the summer of 1995, when I worked for the first time with a Search Engine, I saw some SEs disappearing, and lots of new SEs appearing.

Competition is good, but users might be confused with thousands of SEs. Besides, not everybody knows sophisticated searching techniques such as the Boolean operators, so there is a lot of frustration out there. If we add the irrelevancy of the results in most of the SEs, we have a pretty sad picture.

A clean index should be the main priority for any search Engine. If a SE gives irrelevant results, the user will switch to another SE. Can a SE afford losing users in the today's fierce competition between SEs? I don't think so.

To wrap up, I am not saying here that ALL the Search Engines are vulnerable to ALL the known spamming techniques. All I say is that MOST of the Search Engines are vulnerable to MOST of the bad tricks.

I also believe that is NOT important to list WHICH SEs are vulnerable to WHICH spamming techniques or WHICH sites spam WHICH SEs. The important thing is it STILL happens.

At not only in the American SEs. I've checked also Spanish, German, French and Italian SEs. Less spam, but still there.

From now on, monthly I'll make an update of this article. I can only hope that one day I can say "everything is fine". Be back!

Copyright 2001

Daniel Bazac is the Web Marketer for Web Design in New York, a design, redesign and promotion company. He's specializing in Search Engine Optimization (SEO). He's been online from 1995 and he's also a professional Internet Information Researcher. You may contact Daniel Bazac at <mailto:danielbazac@hotmail.com> or visit his Web site: <http://www.web-design-in-new-york.com/>

[Get-Articles.com](http://www.get-articles.com) : 1000's of reprintable business and internet marketing-related articles.

[Submit your article for reprint.](#)