

Ten Tips for Avoiding Google Hell

— By Liana Evans

Have you ever wondered just what makes a website disappear from the search engines? Maybe one day your website is ranking well, and then a few weeks later it's like someone hit a switch and sent your site into the nether regions of the search engine rankings. Perhaps your site has never even managed to make it into the rankings at all? Welcome to the world of Search Engine Marketing — in particular, Search Engine Optimization (commonly referred to as SEO).

In a nutshell, SEO is the practice of optimizing a website's pages to ensure that they are search-engine friendly and that they target their intended key audiences. Each of the major search engines (Google, Yahoo!, MSN, and Ask) sends out "spiders" or "crawlers." These are programs that come to your website and "collect" all the information about each web page and take it back to the search engine's database. The search engine houses the information about the pages it collected and then presents it when a query from a searcher matches what your page is about. Each search engine applies its own algorithm to determine a page's relevancy to the searcher's query, and returns a set of results.

Here are some very basic tips that webmasters, or online marketers responsible for promotion of a website, should keep in mind. These fundamental strategies can help overcome many issues that webmasters and marketers encounter when dealing with search engines — issues that can inhibit a site from ranking, and prevent pages from being crawled or even push them into the supplemental index.

1. Use the robots.txt File

The robots.txt file is probably one of the easiest tactics webmasters can employ. Using it can help ensure that the right folders are being found and crawled by the search engines. Use of robots.txt can also assist webmasters in keeping the search engine spiders out of places they shouldn't be in. This simple text file is placed in the root directory of your website and basically acts as a "roadmap" for the search engine spiders to follow when accessing your site. You can use specific commands to allow or block folders from being crawled by all search engines spiders or a select few.

To create a robots.txt file, just use your notepad program, and follow the standard specifications for setting up instructions for the search engine spiders to follow. To get help understanding how to set up a robots.txt file, take a look at Robotstxt.org, a great resource for listing out specific robots and the syntax needed for excluding robots and folders. Referring to the specifications at Robotstxt.org will also help you avoid the pitfalls of possibly blocking folders or allowing the wrong folders to be indexed by the search engines.

Two of the four major search engines, Google and Yahoo!, allow and encourage webmasters to validate their websites. This is a simple task that involves uploading files to the root of your website directory, or placing information within the website's meta tags.

When a website is validated, it helps to build a level of trust with the search engine. Validation also allows the search engine to know that this indeed is a current, truly valid and functioning website.

Yahoo! allows you to sign up for their [SiteExplorer service](#) (login required) and monitor your site's authentication. Blog feeds can also be submitted if the website has a blog. And Yahoo! will notify the email connected with the account when any actions are taken on the authenticated websites connected to the account.

[Google's Webmaster Central](#) is a great resource that allows webmasters to validate their websites, and offers a lot of other tools for monitoring a website's indexing with Google. Signing up is free. Once a website is validated through the Webmaster Central tool, a webmaster can monitor and control how often a Google search engine spider comes to "crawl" the website. Webmasters can also monitor links coming into the site, high-ranking keywords for the site, and most-clicked-on query results.

Both Yahoo! Site Explorer and Google's Webmaster Central allow webmasters to submit sitemaps — the subject of our next tip.

Sitemaps work similarly to the robots.txt file, in that they act as a roadmap for search engines' spiders. A sitemap tells the search engines which pages of the website should be crawled. If the site is small, you can put all your page URLs into the sitemap file; however, you'll need a different approach if your site is a couple thousand pages or more.

For larger sites, you'll want to give the spiders a "guide" to follow, so make sure your higher-level category and subcategory pages are listed in the sitemap file. By including these higher-level pages in the sitemap, you let the spiders follow the links that will lead to deeper pages within your site.

A protocol for sitemaps was set up through a cooperative effort from Google, Yahoo!, and MSN. The specifications for setting up a sitemap file either through a .txt file or an .xml file can be found at the Sitemaps.org website. Following these specifications will help webmasters to ensure they are avoiding some common pitfalls. Currently, both Google and Yahoo! allow you to submit your sitemap file.

One of the biggest factors in having a website rank for keywords or phrases is more than simply the number of links acquired coming into the website — how those links are actually worded matters immensely.

Search engines consider a link to a website as a "vote of confidence." However, the search engine really takes into account just how the link is formed. For example, links to your site that are worded as just "click here" won't be as powerful as a link worded "Buy blue widgets at Blue Widgets' Website." The value of the link is greatly enhanced by using keywords within the link, rather than just "click here."

If you have the ability to suggest or influence the formation of a link that points to your site, make sure to use the keywords you are hoping to rank for. Not only does it benefit your website; it can also help the site that is linking to you.

Webmasters and search marketers: Take heed of where you are acquiring links from. Links to your website from [bad neighborhoods](#) can negatively affect your website's relevancy to the search engine. ("Bad neighborhoods" are defined on the Webmaster Guidelines page in the foregoing link.) If the linking site seems shady or spammy, or if something just "doesn't seem right," it might be better to trust your instincts until you can thoroughly check out the site before acquiring the link from it.

Selling links on a website is a fairly controversial subject. On one hand, webmasters should be allowed to do what they want with their websites. On the other hand, the search engines view this tactic as high-risk behavior if the links are not denoted as "paid" in some way. Then there is the whole issue of just what is considered a payment.

If you are selling links on your page, the search engines suggest that you tag the link with the `rel=nofollow` attribute. This basically tells the search engines that you are not "passing confidence" to the site you are linking to. But this action essentially defeats the whole purpose of selling links on a website, and therein lies the controversy.

[Google's quality guidelines](#) state, "Don't participate in link schemes designed to increase your site's ranking or PageRank." If your website sells links for this purpose, and it is discovered by the search engines, your website may receive some sort of penalty that will affect the site's ranking. And in Google, this could possibly affect the site's [PageRank](#).

Search engines want to display the most relevant web pages in response to a searcher's query. Therefore, when pages have identical or very similar content, some of those pages will likely be disregarded by the search engine. Some websites don't even realize they're creating duplicate pages because of the use of session IDs, or multiple variables in a URL taking site visitors to the same page. These are just a few ways duplicate content can be created and cause a substantial headache for marketers and webmasters alike.

Google states in its webmaster guidelines:

"Allow search bots to crawl your sites without session IDs or arguments that track their path through the site. These techniques are useful for tracking individual user behavior, but the access pattern of bots is entirely different. Using these techniques may result in incomplete indexing of your site, as bots may not be able to eliminate URLs that look different but actually point to the same page."

Be aware of the following other ways of accidentally creating duplicate content:

- "Printer-friendly" pages
- Manufacturers' product descriptions that are reused by other retailers
- Mirrored websites
- Syndicated articles

All of these situations can get a web page ranked lower in the search engine, or possibly get it pushed into Google's supplemental index — neither of which you want to happen for any website. There are some simple ways to avoid this, such as blocking folders from the search engine spiders with the Robots.txt file (see tip #1).

Redirects are generally used by webmasters to take visitors who land on old, out-of-date web pages (usually from prior bookmarks) to the pages that are most current within the website's new structure. There are two types of redirects: a Temporary Redirect, technically known as a 302, and a Permanent Redirect, technically known as a 301.

Using the wrong kind of redirect when a web page is being moved can cause the page to lose its rank within the search engine. If a web page is being moved permanently to a new URL, the Permanent Redirect/301 should be used.

A Permanent Redirect tells the search engine to pass all the old page's "juice" onto the new page. This means all the value acquired by links that still point to the old page will be passed onto the new (redirected) page. If a Temporary Redirect/302 is used, no value is passed onto the new page.

Webmasters of sites that dynamically generate content based on variables in a URL should be cognizant that search engines might not be capable of crawling some web pages. The spiders/crawlers sent out by the search engines are very "childlike" in nature. Should they run into any "roadblocks," they will stop their crawl.

Having too many parameters in your URL is one of the roadblocks that can stop crawlers from going deeper. In general terms, most search engines have a difficult time crawling URLs with more than four parameters.

If having too many parameters is an issue with your website, it might be wise to discuss with your tech team some alternatives for passing the variables to the server to generate the dynamic pages. Tools such as [ISAPI Rewrite](#) can be a powerful assist in efforts to resolve issues with variables.

One of the most important pieces in optimizing a web page is the page title tag: It tells search engines what the page is all about. A lot of relevancy is placed in a page's title tag, so careful crafting of that tag will not only reflect a company's position but also capitalize on the use of keywords and phrases.

Unless your company is a household name like Coca-Cola or Toyota or Ritz® crackers, it's unlikely that people will be searching on your company's name to find the services or products on your website. Keeping this in mind, it's clear that you should avoid title tags like "XYZ Company Services" or "About Us — XYZ Company."

Along with the title tag, the page's meta description is an important piece of the marketing puzzle. The meta description has little value in determining relevancy to a search engine, but it is displayed when a search engine displays its results page, right below the page's title tag. Customizing both the page's title tag and meta description to focus on the actual content of the page goes a long way toward getting searches to click on your listing in the search results.

Let's say you've discovered that your website has been removed from Google's index, or that a penalty has been imposed on your website's pages. What's a webmaster or search marketer to do? Are you banned or penalized forever?

Thankfully, Google can be a forgiving search engine. If you address and fix the issues that Google has identified as the reason for the banning or penalty, you then have recourse: the [Reinclusion Request](#). Google offers you the ability to submit this request to have your website reevaluated to verify whether the offending practices Google found have been "cleaned up."

The one thing the reinclusion request won't do is guarantee that Google will include your website in the top 10 when a search is done on keywords or phrases.

There are a few other issues that can give webmasters and search marketers some hassles when it comes to getting their websites to rank in the search engines. In general, though, we've covered some common problems that can be resolved rather easily. If your site is experiencing any search ranking issues that are particularly difficult to work out, don't be afraid to join the webmaster users groups on both Google and Yahoo! — both groups are friendly and helpful. You'll also find that employees from these search engines will lend a helping hand when they can.

Google and Yahoo! are currently the only two engines that have dedicated actual resources and created programs to communicate with webmasters and site owners. Both MSN and Ask are currently in the process of developing similar tools. MSN does have its own set of tools called [AdLabs](#), but these tools are highly geared toward pay-per-click advertising.

So if you ever find yourself in "search engine hell," stop and take a look at the ten situations listed in this article and compare them to what's going on with your website. Hopefully there's a nugget or two of information here that will help you climb out of trouble and into higher rankings!