

Optimizing your web pages for Search Engines: Part Two

You also need to check whether your site is listed in the Yahoo! Directory. You have to pay to get a commercial site into the Yahoo! Directory, so you may already know if you're listed there. If someone else has registered the site with Yahoo!, you may not know whether it's there. Here's how to find out: 1. Point your browser to dir.yahoo.com. This takes you directly to the Yahoo! Directory search page. 2. Type your site's domain name into the Search text box. All you need is yourdomain.com, not <http://www.yourdomain.com> or anything else. 3. Make sure that the Directory option button is selected and then click Search. If your site is in the Yahoo! Directory, Yahoo! displays that site's information on the results page. The others To find out if your site is listed with the other search systems, simply go to the search site, type the domain name, and click the Search button. In the case of all the major systems, the results page will contain an entry for the specified domain if it's in the index. Other results may be placed on the page, too, but the first entry is the specified website if the information is available. What if you're not listed? First, if your site is not in Yahoo! or the Open Directory Project, you have to go to those systems and register your site. What if you search for your site in the search engines and can't find it? If the site is not in Google, you have a huge problem. If it's not in the smaller search engines, that's less of a problem, of course, and quite common. Here are two possible reasons why your site is not being indexed in the search engines: The search engines haven't found your site yet. The solution is relatively easy, though you won't get it done in an hour. The search engines, whether or not they have found your site, can't index it. This is a much bigger problem! We will go into this in more detail a little later. How to tell if your site is invisible. Some websites are virtually invisible. A search engine might be able to find the site, by following a link, for instance. But when it gets to the site, it can't read it or, perhaps, can read only parts of it. We know of a client who built a website that had only three visible pages; all the other pages, including all the pages with product information, were invisible. How does a Web site become invisible? We talk about this subject in more detail later in this article, but here's a brief explanation. In most cases, the problem is that the site is dynamic, that is, a page is created on the fly when a browser requests it. The data is pulled out of a database, pasted into a web page template, and sent to the user's browser. Search engines often won't read such pages, for a variety of reasons explained in detail later. How can you tell if this is a problem? Take a look at the URL in the browser's location bar. Suppose that you see something like this: <http://www.yourdomain.edu/rodent-racing-scores/march/index.php>. This address is okay. It's a simple URL path made up of a domain name, two directory names, and a filename. Now look at this one: <http://www.yourdomain.edu/rodent-racing/scores.php?prg=1> The filename ends with `?prg=1`. This parameter is being sent to the server to let it know what information is needed for the web page. If you have URLs like this, with just a single parameter, they're probably okay, especially for Google; however, a few search engines may not like them. Here's another example: http://yourdomain.com/products/index.html?&DID=18&CATID=13&ObjectGroup_ID=79 This one's a real problem, even for Google, which does a good job of indexing dynamic pages. This URL has too much weird stuff after the filename: `?&DID=18&CATID=13&ObjectGroup_ID=79`. That's three parameters — DID=18, CATID=13, and ObjectGroup_ID=79 — which is too many. Google cannot or will not index this page. Another problem is caused by session IDs — URLs that are different every time the page is displayed. Look at this example: <http://yourdomain.com/buyAHome.do;jsessionid=07D3CCD4D9A6A9F3CF9CAD4F9A728F44> Each time someone visits this site, the server assigns a special ID number to the visitor. That means the URL is never the same, so Google won't index it. Google probably can get to databased pages, but chooses not to. If Google sees links to a page that appears to be dynamic, it doesn't know whether the URL will change between sessions or whether many different URLs point to the same page. Google doesn't want to overload the site's server and also doesn't want garbage in its index. If you have a clean URL with no parameters, the search engines should be able to get to it. If you have a single parameter in the URL, it may or may not be okay. Two parameters also may or may not be a problem, although they're more likely to be a problem than a single parameter. Three parameters are almost certainly a problem. If you think you have a problem, then we will show you how to resolve this in Part Three.

About the Author

Webs Hosters have been in the website hosting business for 10 years. Their customer base ranges from large corporations to small businesses, but their priority has always been to provide honest affordable hosting for all. www.webshosters.com

Source: <http://www.magiwebs.com>