

Putting Robots and Spiders To Good Use

If you have ever used a search engine such as Yahoo at <http://www.yahoo.com/> , you may wonder how they work. Is it magic? Well, yes and no. While it is true that most databases on the Internet (such as Yahoo) rely on users to submit URLs for addition to the database, descriptions and keywords that describe what is contained within that website can help your website find its way into many databases. Most websites are found (or indexed) by the use of META tags in HTML documents, mainly within either the index.html or default.html page of that website.

Most caretakers of websites are unaware of the power of the META tag. It allows them to virtually and effortlessly get their website indexed by hundreds, if not thousands, of robots and spiders that scour the Internet every day to index the Internet, as if to take a picture of what is available on the Web. These robots and spiders record information that they find contained on websites and record it in their databases, which then become available to millions of visitors of search engines on the Web. When a visitor is looking for information about a subject, all the websites that contain that information will be made available to the visitor to research and visit, all because of a simple META tag.

Time To Put On Your Thinking Cap

So what does all of this mean to you? Simple. Get familiar with the META tag, and use it as a tool to drive visitors to your website. This information can direct users to everything from business websites to those sites devoted to the family pet. The possibilities are endless, and by using the META tag, it will help you gain a little more visibility amongst the sea of websites that are currently available on the Internet.

The META tag should be placed within the HEAD section of the document. It can contain a number of variables. Note: The words in red are the variables that you can change to suit your website. See example below...)

<META PLACE THE VARIABLE FROM THE LIST BELOW YOU WISH TO USE HERE>
<META REMEMBER THAT YOU CAN HAVE MANY META TAGS IN ONE HTML FILE>

HTTP-EQUIV="text" CONTENT="text"

This is an addition or replacement of a header sent by the server software. A header is a part of the hypertext transfer protocol, which is used to transfer www pages. A header can include, for example, the length of the page, the server software, or the date the page was last changed. A frequently used additional header is "refresh." (see below)

HTTP-EQUIV="refresh" CONTENT="n [; URL=url]"

This will reload the page every nth second. If URL is specified, that page will be opened instead of the current. This tag is often used in "welcome" pages, or in web-based chats. Since not all browsers support this tag, a standard link to the page should also be used.

HTTP-EQUIV="expires" CONTENT="day, date year time timezone"

Some search engines won't keep the page content in the search database after this time. A newspaper may, for example, not want the search engine to keep a copy of an old placard. The content can be, as an example, "Sat, 03 Jan 1998 12:00:00 GMT".

HTTP-EQUIV="Content-Type" CONTENT="mime-type [; charset=charset]"

This is useful if you wish to define another MIME-type than the server default settings, or when you use special characters that are not entities. You may then set CONTENT to "text/html; charset=ISO-8859-1". Note that including this META tag may cause your web page to load twice. You may want to avoid it for this reason.

NAME="keywords" CONTENT="word1,word2..."

This will tell a search engine such as Yahoo the searchable keywords of the page, for example, "horses, riding."

NAME="description" CONTENT="text"

A short description of your page used by search engines, as an example.

NAME="copyright" CONTENT="text"

A short copyright description of the text, like "Copyright by Me 1998".

NAME="distribution" CONTENT="global/local"

Defines if the page is an index page (global) or not (local).

NAME="robots" CONTENT="none/nofollow/all/noindex"

This is a great way to tell search engines which pages should be indexed.

NOFOLLOW will index the present page only, NOINDEX will index linked pages only, ALL will index both the present and the linked pages and NONE won't index any pages at all.

NAME="generator" CONTENT="text"

This tag, often added by WYSIWYG editors, indicates which program was used to create the page.

NAME="language" CONTENT="language"

Some search engines may take notice of this tag to find out what language the page is written in, such as French, English, or Swedish.

NAME="text" CONTENT="text"

There are many other META tags not mentioned here. You can use just about any tag you'd like, and if the browser or search engine doesn't understand it, no harm no foul.

example:

```
<HTML>
<HEAD>
<META NAME="description" CONTENT="The World's Best Computer">
<META NAME="keywords" CONTENT="Apple, Macintosh, Hardware, Software">
<META NAME="copyright" CONTENT="Copyright 1998 by Apple Computer">
<TITLE>Welcome to Apple Computer</TITLE>
</HEAD>
<BODY>
Contents of the website...
</BODY>
</HTML>
```

As you can see in the above example, you can only use each META tag once. That's OK though, since you can use as many as you want, and fill in the information with whatever you choose!

Hot Tip

One of the best places to start the promotion of your new or existing website is submit It! at <http://free.submit-it.com/> . Here you can get a listing of hundreds of websites that are devoted to helping to simplify the process of getting your site into the databases of the most popular and

most used search engines.

Since most of the websites that track and store website information use spiders and robots to gather the information, this is where the use of META tags will be fully realized. So the next time someone is looking for information on a particular subject matter, your website will now be included in the search results!

e Colophon

These are just some of the many ways you can add a bit more life to your new or existing website, and you can experiment with these tips to produce just the right look for your site.

If you are interested, you can visit my redesigned personal website at <http://junior.apk.net/~rjl/> and take a look at some of my current work. Also, if you or anyone you know is looking to have a website created or re-done, I can help you create an award-winning site. Contact me at rjl@apk.net to make your dreams a reality.

robert j. la follette

rjl@apk.net

<http://junior.apk.net/~rjl/>

<http://www.applewizards.net/>