Articles: Google Sitemaps Explained

Once again I seem to be writing about Google. The reason Google keeps cropping up in these articles is that:

- Google is a innovative company always coming up with new ideas to keep one step ahead of the competition.
- Search Engines are the most important way that most websites generate new traffic.
- Google is the most important of the Search Engines.

One of Googles latest implementations is a method of allowing web-designers and website owners to create specially formatted sitemap documents that inform and direct the Google robots. These sitemaps are a quick and easy way to keep your site constantly indexed and updated within the Google database and to ensure that all of your sites pages are crawled by the Google search engine robot. The program is called 'Google Sitemaps' and this is what Google themselves have to say about it:

"Google Sitemaps is an experiment in web crawling. Using Sitemaps to inform and direct our crawlers, we hope to expand our coverage of the web and improve the time to inclusion in our index. By placing a Sitemap-formatted file on your webserver, you enable our crawlers to find out what pages are present and which have recently changed, and to crawl your site accordingly.
Google Sitemaps is intended for all web site owners, from those with a single web page to companies with millions of ever-changing pages. If any of the following are true, then you may be especially interested in Google Sitemaps:

- You want Google to crawl more of your web pages.
- You want to be able to tell Google when content on your site changes. "

What is a Google Sitemap?

The Sitemap protocol requires the sitemap to be present on your web-server in the form of an XML document. XML is simple code like

HTML and it is used to syndicate your content to all interested parties. You may have seen it in use for syndicating weblog entries via RSS to a news-reader. In the case of Google Sitemaps, the XML document is syndicated to Google and their software uses it to ensure that the pages of your website are crawled and indexed.

Before the introduction of Google Sitemaps, website-owners had to rely on the Google robot to find all of a websites links in order to make sure that all the pages were indexed. The introduction of Google Sitemaps now gives website owners some control over this process. IN additon, the XML format of the sitemap document also gives you control over several key variables.

If we examine a very simple chunk of code from a basic Google sitemap XML document we can see the variables we now have control over.

```
<url>
   <loc>http://www.yourdomainname.com/</loc>
   <priority>1.0</priority>
   <lastmod>2005-07-06T18:00:00+00:00</lastmod>
   <changefreq>weekly</changefreq>
</url>
```

This part of code describes one page of a website, so a typical Google Sitemap document would contain similar chunks of code for every page contained within the website. As you can see each web page has 4 variables:

Location – Simply the URL of the web page.
Priority – A number from 0.0 to 1.0 allowing you to set the priority of a particular page within your website. This number is a relative setting and relates only to those pages within your site. It allows you to instruct Google to pay more attention to particular key pages within your website.
Last Modified – This tells Google when your web pages were last modified so preventing the robot from having to index pages that haven't changed since its last visit.
Change Frequency – This allows you to tell Google how often the content of a page is likely to change. You can set it to never, yearly, monthly, weekly, daily, hourly and always.

More information on the XML Protocol used with Google Sitemaps can be found at:
https://www.google.com/webmasters/sitemaps/docs/en/protocol.html

How do I create a Google Sitemap?

There are a number of ways to create a Google Sitemap document for your website.

The simplest but least controllable way is to use an online XML generator that will spider the pages in your website, and automatically create the XML file for you. With most you then have to upload the file to your web-server and inform Google of its presence. There are plenty of these scripts popping up and many of them are completely free.

A list of these online XML Google Sitemap generators can be found at http://code.google.com/sm_thirdparty.html. The main disadvantage to using these online generators is that they need to be recreated each time you add new pages to your website. This won't be a problem for many website owners who rarely add new pages, but for those who are constantly adding new pages another approach may be better.

If you would like a little more control over the various parameters stored within your Google Sitemap XML document then a script that you configure and then upload to your web-server may be the answer for you. These are written in various scripting languages such as PHP or Perl and give you more control over your Google Sitemap. They do require some knowledge of scripting and installation to get them working which is beyond the scope of this article. Many however can be set up to run at regular intervals and not only spider your complete site and automatically generate your XML Google Sitemaps document but also upload it to the relevant place on your web-server and ping Google to tell them that the sitemap exists.

Finally you could use Googles own Sitemap generator which is a Python script and takes a little more knowledge to install and configure on your web-server. It also requires that Python 2.2 is installed on the server. More information on this method can be found at https://www.google.com/webmasters/sitemaps/docs/en/sitemap-generator.html

Note: These automatic Sitemap generators work by following the internal links within your website, any orphaned pages that are not linked to will not be included in your sitemap.

How do I submit my Sitemap to Google?

Whichever method you use to generate your Google Sitemaps document, you then need to submit it to Google. Most of the online generators and scripts will either do this for you or give you an option to do it once your XML document has been uploaded.

First, you should create a Google Sitemaps Account (which requires you to have a Google Account). This account enables Google to provide you with useful status and statistical information. The My Sitemaps page lets you know if there are problems with your Sitemap or with any of the URLs listed in it. Your Google Sitemaps account will also allow you to re-submit your Sitemaps document when you make changes to it.

Once your Sitemaps account is set up simply use the online forms to inform Google of the location of your new Sitemaps document and your site will soon be indexed.

Conclusion

Google Sitemaps give website owners the opportunity to inform Google about all the pages of their website. It should ensure that no pages are missed and also allows a certain degree of control over the relative importance of individual pages. Simply generating a Google Sitemaps document won't necessarily give you higher rankings within the search engines as you will still be competing with other websites for those top spots. Both on page Search Engine Optimisation and off-page promotion will still be essential. However, sitemaps make sure all your pages are crawled and indexed quickly by Google, and may therefore give you a competitive advantage over those websites that don't have a Google Sitemap.

As part of the suite of Search Engine Optimisation services offered by Pixelwave Design I can create and manually configure a XML Google Sitemaps document for your website. Please contact me if you would like a Google Sitemap for your website.

Alan Cole
Pixelwave Design
info@pixelwave.co.uk
http://www.pixelwave.co.uk
Tel: 01970 871072

.