

A comprehensive guide to XML Sitemaps:

What are they? Why do I need one? And how do I create one?

A little background and history

A sitemap is a way of collecting and displaying the information about the structure of a website, identifying the links and the data within that website. Sitemaps were originally created for the users of a website so that they could see an index of that site for easier navigation to the pages they wanted to visit. However, in 2005 Google released an XML format for sitemaps that was designed for search engines, allowing them to index a site easily and find the data faster and more efficiently. Google did this in response to the increasing size and complexity of websites. Having a single index file saved time over crawling every page on a website to determine its structure. As blogging became more popular and bloggers updated and added new content to their blogs more frequently, sitemaps became even more important as a way of identifying new content and how it was categorized. Google made their new XML protocol public under the Creative Commons license to encourage other search engines to support it. Yahoo and Microsoft both officially support the Google XML sitemap format.

So what does an XML sitemap contain?

In its simplest form, a Sitemap is an XML file that lists URLs for a site along with additional metadata about each URL (when it was last updated, how often it usually changes, and how important it is, relative to other URLs in the site). Search engines use this information to more intelligently crawl a website or blog. Web crawlers usually discover pages from links within a site as well as from other sites. Sitemaps supplement this data to allow crawlers to pick up all URLs in the sitemap and learn about those URLs using the associated metadata. Using the sitemap does not guarantee that web pages are included in search engines, but provides hints for crawlers to do a better job of crawling your site.

Is there anything I should know about XML sitemaps?

Yes, there is.

In order for search engines to find an XML sitemap it must be placed in the root directory of your web server, i.e. `http://example.com/sitemap.xml` and the file must be called `sitemap.xml`.

If you create a compressed copy of your sitemap then that file must be compressed using Gzip and the compressed file must be called `sitemap.xml.gz`.

A sitemap cannot exceed 10MB and it cannot contain more than 50,000 URLs. If your sitemap is larger than 10MB or contains more than 50,000 URLs you can create a small number of sitemaps and a sitemap index file.

XML sitemaps must use UTF-8 encoding.

All URLs listed in a sitemap must reside on the same host as the sitemap. For example, if the sitemap is located at <http://www.example.com/sitemap.xml>, it can't include URLs from <http://subdomain.example.com>. If the sitemap is located at <http://www.example.com/myfolder/sitemap.xml>, it can't include URLs from <http://www.example.com>.

All URLs listed in a sitemap must be specified in full along with the protocol (http or https). For example, <http://www.example.com/> is a valid URL for a sitemap, whereas www.example.com is not. If you use both http and https for the same URLs then you must only include one of them. You cannot have duplicate URLs in a sitemap.

You cannot include URLs with a session ID in your sitemap.

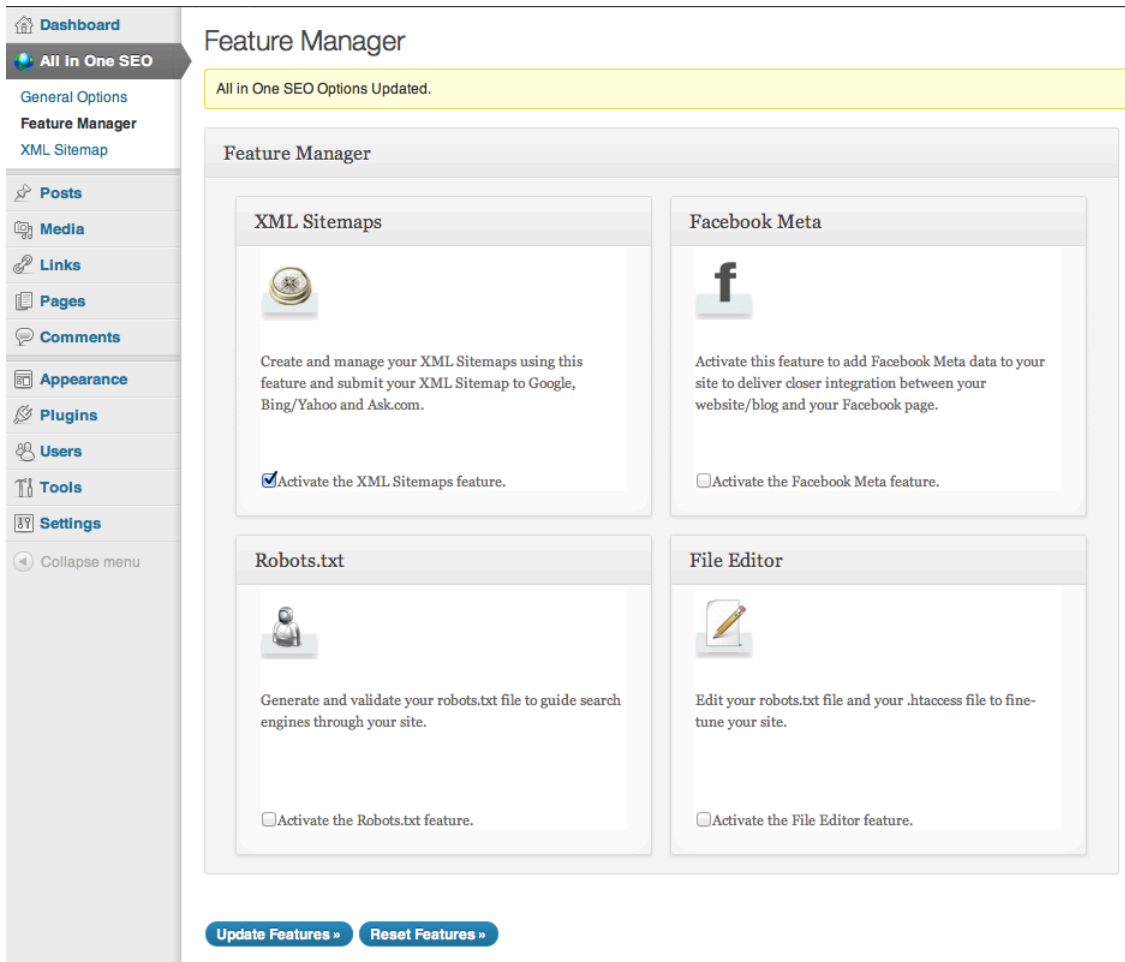
If your site uses frames you should include the frameset URLs as well as the URLs of the frame contents.

That's a lot to know, is there an easy way to create and manage my XML sitemap?

Yes there is. The upcoming new release of All in One SEO Pack (free version and Pro version) available from <http://semperplugins.com/plugins/all-in-one-seo-pack-pro-version/> lets you create, update and manage your XML sitemap, and submit it to Google, Microsoft and Yahoo. You no longer need a separate plugin to do this or need to know anything about XML sitemaps.

So how does the All in One SEO Pack XML Sitemaps feature work?

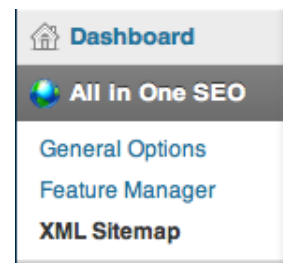
Shortly we will be releasing a new version of All in One SEO Pack and All in One SEO Pack Pro version that includes an XML Sitemap Generator. To make use of this great new feature you first have to activate it by going to our new Feature Manager and activating XML Sitemaps as shown in the screenshot below:



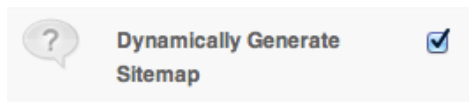
Using All in One SEO Pack to create and submit XML Sitemaps

Once you have activated XML Sitemaps you will see a sub-menu item appear under All in One SEO. Click on the sub-menu item shown in the screenshot to the right to go to the new options page for the new XML Sitemaps feature.

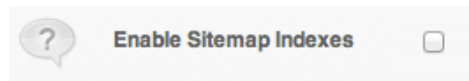
Here you can set the name you want for your XML Sitemap. By default the name is set to “sitemap” and the “.xml” extension will be added for you.



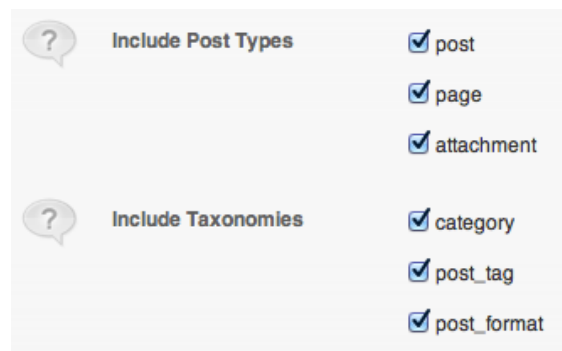
You can select whether you want the XML Sitemap to be dynamically generated. If you select this then the XML Sitemap will not be created until it is requested by a search engine crawling your site. This means that the XML Sitemap will always be up to date whenever a search engine visits your site. This is the recommended setting for the majority of users.



If you have a large site with over 50,000 URLs or if your XML Sitemap is over 10MB in size you can select the “Enable Sitemap Indexes” and the plugin will break your XML Sitemap into smaller files and create an index file for you.

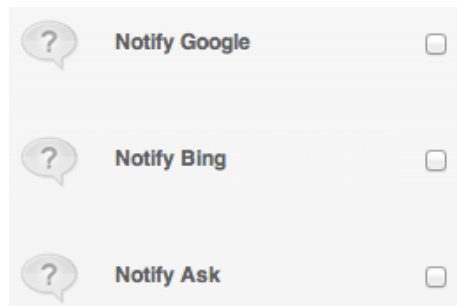


Include Post Types and Include Taxonomies allow you to select which post types and which taxonomies you want included in your XML Sitemap. So if you use Custom Post Types and/or Custom Taxonomies on your site you can select them here to include or exclude them. The default post types of posts, pages and attachments and the default taxonomies of category, tag and post format are included by default for you.



There are buttons for Create XML Sitemap and Create Compressed Sitemap so that you can manually create your sitemap.xml and sitemap.xml.gz files.

There are checkboxes to select which search engines you want to notify about your XML Sitemap. The four major search engines are available to select – Google, Bing/Yahoo, and Ask.com.



And finally there are the Update Options and Reset to Defaults buttons where you can save any changes you make to the XML Sitemap options and reset the options back to the defaults.

We think this is the easiest and most intuitive solution for creating XML Sitemaps for your WordPress site or blog available but we'd love to hear your feedback. Please feel free to comment on this new feature and all the other new features in All in One SEO Pack at <http://semperplugins.com/blog/xml-sitemap-generator-coming-to-all-in-one-seo-pack/>.