

# Google Sitemap



This community version of this module was [deprecated](#). The Forge version is for those wanting to use Google Sitemap in Magnolia 6.

Magnolia Compatibility	
6.2	✓
6.1	✓

The Google Sitemap app creates an XML sitemap file that lists URLs for each page. Sitemaps are used to tell search engines which pages they should index. This improves search engine optimization (SEO) by ensuring that all site pages are found and indexed. This is particularly important for sites that use dynamic access to content such as Adobe Flash and for sites that have JavaScript menus that do not include HTML links. Where navigation is built with Flash, a search engine will probably find the site homepage automatically, but may not find subsequent pages unless they are provided in a Google Sitemap format.

Note that using Google Sitemaps does not guarantee that all links will be crawled, and even crawling does not guarantee indexing. Nevertheless, a Google Sitemap is still the best insurance for visibility in search engines. Webmasters can include additional information about each URL, such as when it was last updated, how often it changes, and how important it is in relation to other URLs in the site. Google Sitemaps adhere to the [Sitemaps protocol](#) and are ready to be submitted to search engines.

- [Installing](#)
- [Source Code](#)
- [Issues, Bugs and Improvements](#)
- [Configuration](#)
  - [Workspace](#)
  - [Node types](#)
- [Creating a sitemap](#)
- [Editing sitemap entries](#)
- [Editing virtual URI entries](#)
- [Publishing](#)
- [Viewing the sitemap](#)
- [Exporting the sitemap](#)
- [Sitemap template](#)
- [Virtual URI mapping](#)
- [Adding to robots.txt file](#)
- [Submitting to search engines](#)

## Installing

Maven is the easiest way to install the module.

```
<dependency>
  <groupId>info.magnolia.google</groupId>
  <artifactId>magnolia-module-google-sitemap</artifactId>
  <version>3.1</version> <!-- Magnolia 6.2 -->
</dependency>

<!-- Requires magnolia-core-compatibility -->
<dependency>
  <groupId>info.magnolia</groupId>
  <artifactId>magnolia-core-compatibility</artifactId>
</dependency>
```

## Source Code

The source for this module can be found here: <https://github.com/magnolia-community/google-sitemap>

## Issues, Bugs and Improvements

Report issues here: <https://github.com/magnolia-community/google-sitemap/issues>

## Configuration

The app is based on the [content app framework](#) so its configuration is typical of any content app. The app is configured at `/modules/google-sitemap/apps/siteMaps`.

<b>Node name</b>
modules
google-sitemaps
apps
siteMaps

## Workspace

The Google Sitemap app operates on the `googleSitemaps` workspace that stores the sitemaps.

## Node types

The Google Sitemap module registers a custom `mgnl:siteMap` [node type](#). The Google Sitemap app operates on nodes of this type and on folders.

## Creating a sitemap

Sitemaps are created in the Google Sitemaps app.

To create a new sitemap,

- Click **Add sitemap**. You can arrange your sitemaps in folders if you like.
- **Site Map Properties:**
  - **Name:** The internal name of the sitemap. It is used in the URL that renders the sitemap. See [Viewing the site map](#).
  - **URI:** Optional. Virtual URI that redirects visitors to the actual sitemap page. Useful for shortening a long URI.
  - **Include virtual URIs:** Select to include any defined virtual URIs. [Virtual URI mappings](#) are a Magnolia method of redirecting requests and shortening URLs. The apps reads all virtual URI mappings from the system and lists them in the [Virtual URIs subapp](#).
  - **Site map type:** Two sitemap types are available, `Standard` and `Feature Phone`. [Google recommends](#) that you use separate sitemaps for different content types. [Feature phone](#) sitemaps use compliant mobile-specific tag and namespace requirements.

The screenshot shows a dialog box titled "Edit site map properties". It has three tabs: "SITE MAP PROPERTIES", "SITE SELECTION", and "DEFAULT VALUES SELECTION". The "SITE MAP PROPERTIES" tab is selected. Below the tabs, there is a "SHOW ALL" link. The form contains the following fields:

- "Name \*": A text input field containing "standard-google-sitemap".
- "URI": An empty text input field with a question mark icon to its right.
- "Include virtual URIs": A checkbox that is currently unchecked, with a question mark icon to its right.
- "Site map type": A dropdown menu with "Standard" selected.

At the bottom right of the dialog, there are two buttons: "CANCEL" and "SAVE CHANGES".

- **Site Selection:**
  - **Sites:** Select the relevant site in the **Pages chooser**. You can also select subpages as the root node to, for example, create different sitemaps for site sections.

**i** The root node of the selection will not be included in the site map. Assume you have the following trees: `/a/b/c` and `/a/b/d`. If you select `/a/b` as the root of the sitemap, only pages under `c` and `d` will be included in the map. The root node `b` will not be included.

Edit site map properties ? ↶ ✕  
 SITE MAP PROPERTIES | SITE SELECTION | DEFAULT VALUES SELECTION | SHOW ALL  
 Sites:  SELECT ANOTHER... ?  
ADD  
CANCEL SAVE CHANGES

**i** Sitemap links are generated using the protocol that is defined in your site definition. The default protocol is HTTP. If you want HTTPS define the protocol in [domain mapping](#).

- **Default Value Selection:**
  - **Change frequency:** Select the default value of **Change frequency** to use in the current site map.
  - **Priority:** Select the default priority value to use in the current site map.

Edit site map properties ? ↶ ✕  
 SITE MAP PROPERTIES | SITE SELECTION | DEFAULT VALUES SELECTION | SHOW ALL  
 Change frequency:  ?  
 Priority:  ?  
CANCEL SAVE CHANGES

The default values displayed in the **Default Values Selection** tab are configured in `/modules/google-sitemap/config`.

Node name	Value
modules	
google-sitemap	
config	
changeFrequency	weekly
priority	0.5

## Editing sitemap entries

To edit the individual sitemap entries click **Edit site map entries**. The site pages display in an expandable tree and you can set properties for each page.

The screenshot shows the Magnolia SITEMAPS interface. At the top, there is a green header with the Magnolia logo and navigation icons. Below the header, there is a breadcrumb trail: SITEMAPS > PAGES: STANDARD-GOOGLE-... > /standard-google-sitemap. The main content area displays a table of site map entries:

	Page title	Path	Show in site map	Priority	Change frequency
<input type="checkbox"/>	Tour Types	/travel/tourType	✓	0.5	weekly
<input type="checkbox"/>	Destinations	/travel/destination	✓	0.5	weekly
<input type="checkbox"/>	Tour detail	/travel/tour	✓	0.5	weekly
<input type="checkbox"/>	▶ About	/travel/about	✓	0.5	weekly
<input type="checkbox"/>	Contact	/travel/contact	✓	0.5	weekly
<input type="checkbox"/>	▶ Container for m	/travel/meta	✓	0.5	weekly

On the right side, there is a sidebar titled "Site map entries" with a sub-option "Edit entry properties".

To define properties for the entries click **Edit entry properties**:

- **Priority**: Priority of the page relative to other pages on the site. Values range from 0.0 (low) to 1.0 (high). Default is 0.5. Set the priority of your most important page to 1.0. Setting all pages to 1.0 does not increase the rank of your site in search results since the importance is a relative measure among pages of the same site. A search engine may choose to rank the page higher than other pages of the site based on the value, however. See [priority](#) in [XML Sitemap protocol](#).
- **Change frequency**: Suggested frequency for search engines to crawl the page. Valid values are: `always`, `hourly`, `daily`, `weekly`, `monthly`, `yearly` and `never`. Use the value `always` for pages that change each time they are accessed. Use `never` for archived pages that will never change. See `changefreq` in [XML Sitemap protocol](#).
- **Hide**: Excludes a page from the sitemap. Child pages are not excluded automatically. ⚠ The `hideInGoogleSiteMap` property is stored in the page itself. This means you need to activate the page. Publishing the sitemap only is not enough.
- **Hide children**: Excludes child pages from the sitemap. To exclude both a parent and its children check both boxes.

The "Edit entry properties" dialog box is shown. It has a title bar with a question mark icon, a refresh icon, and a close icon. The dialog contains the following fields:

- Priority**: A dropdown menu with the value "0.5" selected.
- Change frequency**: A dropdown menu with the value "Weekly" selected.
- Hide**: A checkbox that is currently unchecked.
- Hide children**: A checkbox that is currently unchecked.

At the bottom of the dialog, there are two buttons: "CANCEL" and "SAVE CHANGES".

## Editing virtual URI entries

If you included virtual URIs in the sitemap you can edit their properties. Click **Edit virtual URI entries** to open the **Virtual URIs** subapp. The pages display as individual entries (as opposed to a tree) and you can set the same properties that are available for pages, except **Hide children** that is inapplicable.

To	From
<input type="checkbox"/> forward:/.magnolia/activation	/ActivationHandler
<input type="checkbox"/> redirect:/.magnolia/admincentral	/
<input type="checkbox"/> \$1~\$2~cache.\$3	(\.\*\)\.\{4\}-{?\d{2}-}{5}\d{3}\.cach
<input type="checkbox"/> forward:/.magnolia/pages/messages.js?locale=\$1	/\.\magnolia/pages/messages\.\*\)\.js
<input type="checkbox"/> forward:/admintemplates/dialogs/fileThumbnail.jsp	/.\magnolia/dialogs/fileThumbnail.jpg
<input type="checkbox"/> -	-
<input type="checkbox"/> redirect:/rss/?feedPath=\$1	/rssFeeds/(.*)
<input type="checkbox"/> redirect:/rss/?feedPath=\$1&generatorName=plan	/planetFeeds/(.*)
<input type="checkbox"/> forward:/travel/tour?tour=\$1	/tours(.*)\.html

✔ Hide default mappings defined in the `ui-admincentral` module, such as those for accessing AdminCentral. Public users will not access AdminCentral, so these URLs do not need to appear in the sitemap. Also, hide mappings that use regular expressions in the `toURI` property. These are not understood by search engines as regular expressions.

## Publishing

Publish the sitemap to the public instance to ensure that it is accessible to the search engines.

## Viewing the sitemap

You can view the XML sitemap on the author or public instance at `/<CATALINA_HOME>/<contextPath>/sitemaps/<sitemap name>.xml`, for example, <http://localhost:8080/magnoliaPublic/sitemaps/standard-google-sitemap.xml>. Note that a filter mechanism removes duplicate urls.

## Exporting the sitemap

The Export option in the Action bar of the Google Sitemap app will export the sitemap itself (a `<urlset>`) named as `<sitemap name>.xml`. The JCR configuration for such a sitemap can be exported from the JCR Tools app, in which case the XML file will be named as `googleSitemaps.< sitemap name >.xml`.

## Sitemap template

The `siteMapsConfiguration` page template renders the sitemap.

The definition is configured in `/modules/google-sitemaps/templates/pages/siteMapsConfiguration`:

Node name	Value
google-sitemaps	
templates	

pages	
siteMapsConfiguration	
i18nBasename	info.magnolia.module.googlesitemap.messages
modelClass	info.magnolia.module.googlesitemap.model.SiteMapModel
name	GoogleSiteMap
renderType	freemarker
templateScript	/sitemap/pages/main.ftl
title	GoogleSiteMap
visible	false

#### Properties

siteMapsConfiguration	
modelClass	SiteMapModel is the main model class for site map templates.
templateScript	main.ftl ( <a href="#">GIT</a> ) includes two alternative scripts, mainXml.ftl ( <a href="#">GIT</a> ) and mainConfiguration.ftl ( <a href="#">GIT</a> ) that renders Text or XML content dependent on the URL extension.

## Virtual URI mapping

The [virtual URI mapping](#) configuration is in `/modules/google-sitemap/virtualURImapping/sitemaps`. For details see the configuration at the [virtual URI mapping](#) page.

## Adding to robots.txt file

Add the following line in your `robots.txt` file. Include the full URL to the sitemap:

Sitemap: <http://www.example.com/sitemap.xml>

## Submitting to search engines

Submit the sitemap to major search engines via the webmaster tools of each engine or wait for the engines to find the sitemap on their own.

- [Google](#)
- [Bing](#)
- [Baidu](#)