

PowerMapper

one-click site mapping

PowerMapper 5.0 User Manual

PowerMapper 5.0 User Manual	1
Overview	2
Introduction to PowerMapper	2
How PowerMapper Works	2
About Site Maps	3
Visual Site Map Styles	4
Mapping Sites and Folders.....	5
Mapping Options	5
Mapping Password Protected Pages.....	5
Using Site Maps.....	6
Changing Map Style.....	6
Renaming Site Map Pages	6
Adding Notes to Site Maps.....	6
Restructuring Site Maps	6
Exporting Site Maps.....	7
Using Google Sitemaps	7
Importing and Visualizing Site Data	7
Advanced Options	8
Blocks	8
Links	9
Tools	10
Advanced.....	10
Backup and Restore	11

PowerMapper

one-click site mapping

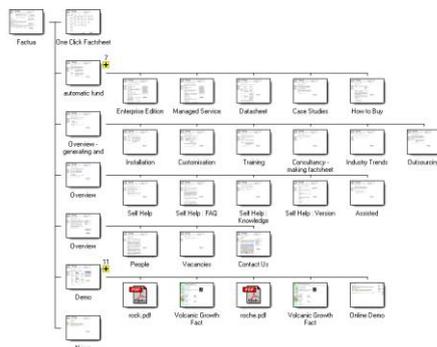
Overview

[Pro] indicates features only available in PowerMapper Professional Edition.

Introduction to PowerMapper

PowerMapper is a one-click site mapping tool that eliminates weeks of intensive manual labor by automatically creating site maps and content inventories.

It produces visual site maps in a variety of styles, and allows data export to other applications and databases for additional processing.

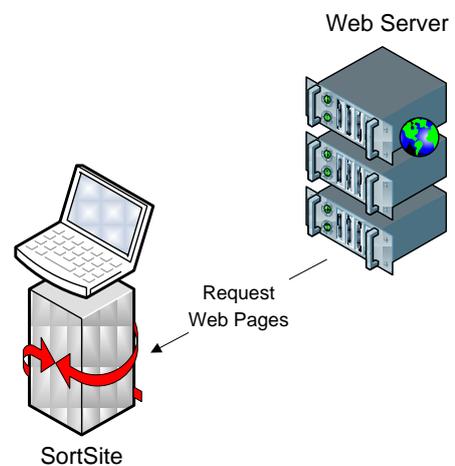


Example PowerMapper Site Map

How PowerMapper Works

PowerMapper scans each page of a site, capturing page thumbnails and metadata, and then examines the page for links to un-scanned pages. It scans the newly found pages and repeats this process until all pages have been scanned.

Pages are requested from the web server in the same way that a web browser requests them. This means the product works with password protected pages as well as server technologies like ASP.NET, JSP, ColdFusion and SharePoint.



This process is called web crawling, and is used by search engine agents like Googlebot to find pages to index. Like other well-behaved web crawlers, PowerMapper obeys the Robots Exclusion Protocol (robots.txt) which describes areas of sites that are off limits to web crawlers.

In addition, you can add URL patterns to ignore:

- *.pdf - to ignore all PDF files
- *logout* - to ignore all URLs containing 'logout'

For more information see Advanced Options: Blocks below.

PowerMapper

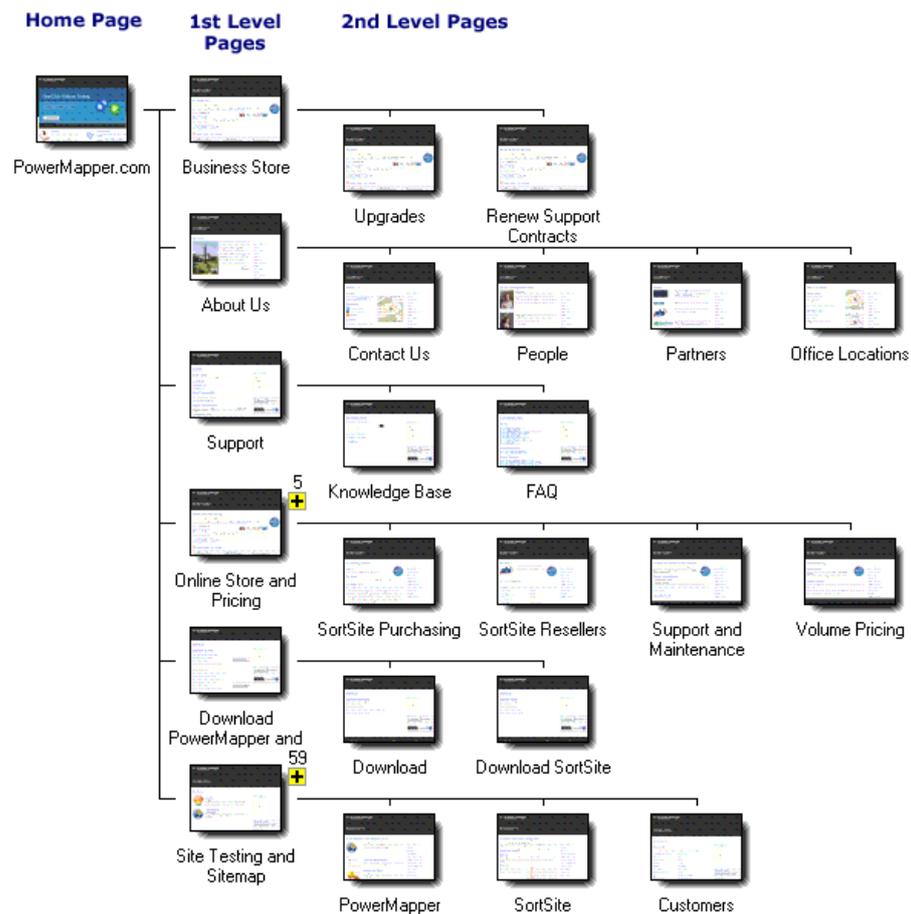
one-click site mapping

About Site Maps

Visual Site Maps are used by people. XML Sitemaps are used by search engines.

There are two main types of site maps:

- Visual Site Maps are graphical or textual representations of a web site's structure, and work very similarly to an organization chart.
- XML Sitemaps are lists of pages on web site, and are designed to let search engines like Google know which pages your site has. There are several special-purpose extensions available for XML sitemaps including: image sitemaps (which describe images, and are used by Google Image Search); mobile sitemaps (for listing mobile web pages); and video sitemaps (for describing video sites like YouTube).



This shows one of the visual site map styles produced by PowerMapper. This style represents the site structure visually with thumbnails of each page, and displays lower level pages towards the right side of the map.

PowerMapper

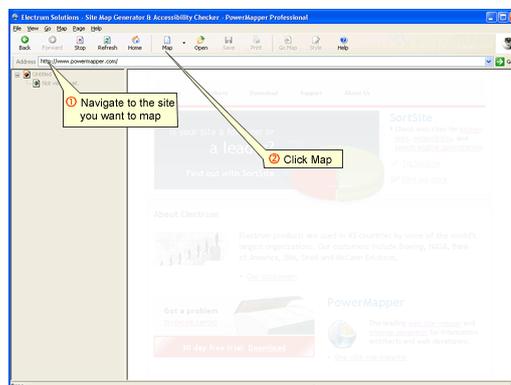
one-click site mapping

Mapping Sites and Folders

To Map a site:

- 1) Navigate to the site using the built-in web browser
- 2) Click the Map button

The software will scan each page in turn from the page displayed in the browser (the start page) and display a map once it's done.



Mapping Options

Clicking the down arrow to the right of the Map button displays a menu that lets you:

- Map Entire Site – maps the entire site from the start page (the default action for the Map button)
- Map Folder – maps pages at the same level and below the start page (if you start mapping at www.example.com/folder/file.htm then all pages starting with www.example.com/folder will be mapped)

You can landing pages to the scan using the New command and the Additional Pages box.

Mapping Password Protected Pages

To map a password protected (authenticated) site:

1. Login to the site in the usual way (either via a login page or the popup password dialog).
2. Once you've logged in navigate back to your home page and map your site.

Note: If you have a logout link inside the secure area, then the scanner may be logged out if it follows the logout link. To prevent this happening go to Block Links (click the Options command on the View menu and go to the Links tab) and add something like the following:

```
*logout*
```

This stops the scanner following links with "logout" in the URL. You will need to adjust this to match the logout links (if any) used on your site.

PowerMapper

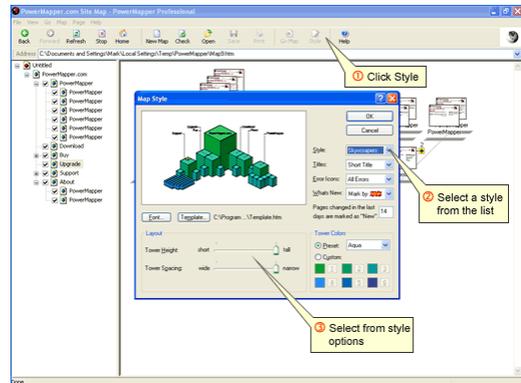
one-click site mapping

Using Site Maps

Changing Map Style

To change map style after mapping a site:

- 1) Click the Style button on the toolbar.
- 2) Make changes to the map style using the Style window.
- 3) Click OK.



Renaming Site Map Pages

To rename a page:

1. Right-click on one of the page icons in the Navigation bar on the left side of the PowerMapper window.
2. Choose Rename from the pop-up menu.

Adding Notes to Site Maps

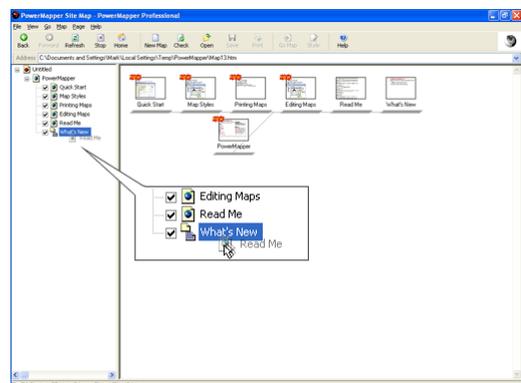
To add a note to a map:

1. Right-click on one of the page icons in the Navigation bar on the left side of the PowerMapper window.
2. Choose Properties from the pop-up menu.
3. Go to the Notes tab, add a note, and then click OK.

Restructuring Site Maps

To restructure a map:

1. Drag one of the page icons in the Navigation bar on the left side of the PowerMapper window.
2. Drop the icon onto a new location in the Navigation bar.



PowerMapper

one-click site mapping

Exporting Site Maps

You can export site maps in a variety of formats using **Export** on the File menu. Supported export formats including: HTML; PNG Images; Excel CSV; and Google sitemap format.

Using Google Sitemaps

To use a sitemap in Google:

1. Export from PowerMapper using the Export command on the File menu, choose the Google sitemap option, and name the file sitemap.xml
2. Upload sitemap.xml to the root directory on your web server (usually the same folder as the home page)
3. Login to Google Webmaster Tools (see Webmaster Tools help if you don't have an account yet)
4. Click "Add" in the Sitemap column of the dashboard
5. Choose "Add General Web Sitemap" and type in the filename sitemap.xml

Importing and Visualizing Site Data

You can import and visualize site data produced by other tools like Google Analytics and Google Webmaster Tools.

To visualize analytics or SEO data:

1. Map the site.
2. Choose Import from the File menu. You can import CSV files exported by Google Analytics and Webmaster tools. See below for supported file formats.
3. Choose Style from the toolbar, then choose from the Data Style dropdown list, then click OK.

Supported file formats:

- Google Analytics Content Report (CSV format)
- Google Webmaster Tools external links report (CSV format)
- Any CSV file with a URL in the first column, and data values in the remaining columns:

```
Page, Jun, Jul, Aug
/index.htm, 126, 101, 126
/service/, 66, 87, 102
/support/index.htm, 80, 80, 80
/support/faq.htm, 0, 23, 54
```

PowerMapper

one-click site mapping

Advanced Options

Application options are changed using the Options command on the View menu.

Blocks

Allows you to prevent pages being visited during the scan.

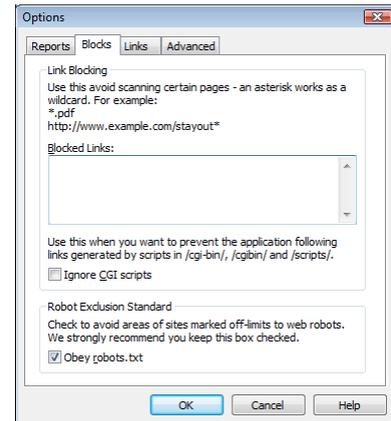
Obey Robots.txt

When checked the web crawler will obey this protocol.

Note: *We strongly recommend you keep this checked.*

Ignore CGI scripts

Check to ignore links generated by scripts in the /cgi-bin/, /cgibin/ and /scripts/ directories.



Blocked Links

Allows you to specify URL patterns for pages to block. The syntax is the same as used in robots.txt. Examples:

<code>http://www.domain.com/copyright.htm</code>	Blocks a single page
<code>http://www.domain.com/stayout/*</code>	Blocks all pages in the stayout folder
<code>https:*</code>	Blocks all HTTPS links
<code>*.doc</code>	Blocks links to all pages with a .doc extension
<code>*print_friendly.htm</code>	Blocks links to URLs ending in print_friendly.htm
<code>*action=edit*</code>	Blocks links containing action=edit

PowerMapper

one-click site mapping

Links

Allows you to control which links are followed during the scan.

Java Parameters

Lets you choose whether links in <PARAM> tags are followed. If checked any parameters containing "url", "href" or "link" in their names are treated as URLs.

JavaScript

Lets you choose whether links in JavaScript are followed. For example,

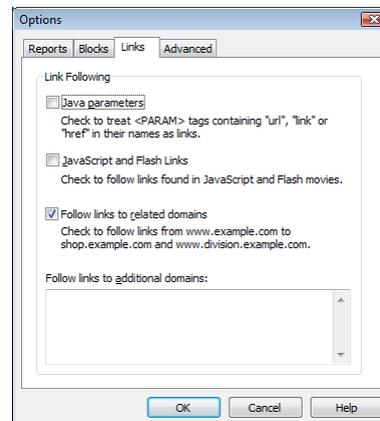
```
function GoTo()
{
    // the URL "bigframe1.htm" is
    // extracted from this JavaScript code
    parent.bigframe.location.href = "bigframe1.htm";
}
```

Follow links to related domains

When checked the scanner will follow links to child domains (e.g. if you scan domain.com then links to store.domain.com and www.domain.com will be followed) and sibling domains at the same level (e.g. if you scan www.domain.com then links to secure.domain.com and www2.domain.com will be followed).

Follow links to additional domains

Links to any additional domains listed in this box will be followed.



PowerMapper

one-click site mapping

Tools

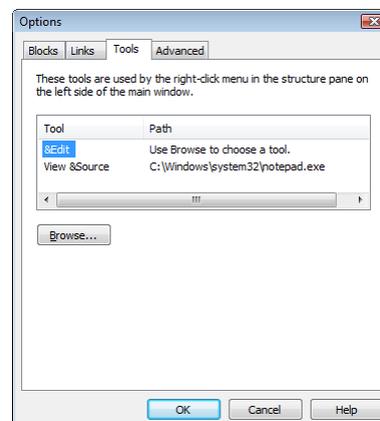
Allows launching third party tools using the Navigation bar right-click menu. The Navigation bar is on the left side of the main window.

Tool List

Lists the tool used for each command on the Navigation bar right-click menu.

Browse

Chooses the tool to use for the selected command in the Tool List.



Advanced

Server Load

Lets you set a delay between loading pages to avoid placing undue load on a server.

Page Timeout

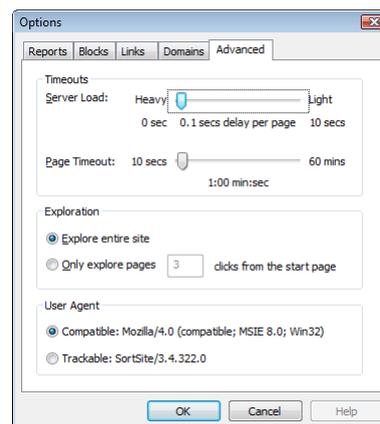
Controls the maximum time spent loading a page - the default setting is 1 minute.

Exploration Depth

Controls how deep the application will go into a web site. The default setting is your entire site, but you can restrict this to the top level pages of your site.

User Agent

Controls the user agent string sent to the web server (which is often recorded in web server logs). The Compatible setting identifies the request as coming from Internet Explorer (IE). Some web servers return different content for different browsers, so this string means content is tailored for IE. The Trackable setting uses a unique user string which can be filtered from web logs.



PowerMapper

one-click site mapping

Backup and Restore

PowerMapper stores site structure in a binary .site file, with rendered maps and page thumbnails stored in an associated folder with the same name as the .site file (minus extension).

Files are only held open while reads and writes are in progress so saved data can be archived and restored using any backup solution.

Saving a file called "Example" produces the following files and folders:

