



Inspyder Power Search User's Reference Manual

Copyright © 2009 Inspyder Software Inc.

Inspyder Power Search (Power Search) User's Reference Manual

Copyright © 2009 Inspyder Software Inc.

All rights reserved. No parts of this work may be reproduced in any form or by any means - graphic, electronic, or mechanical, including photocopying, recording, taping, or information storage and retrieval systems - without the written permission of Inspyder Software Inc.

Products that are referred to in this document may be either trademarks and/or registered trademarks of the respective owners. The publisher and the author make no claim to these trademarks.

While every precaution has been taken in the preparation of this document, the publisher and the author assume no responsibility for errors or omissions, or for damages resulting from the use of information contained in this document or from the use of programs and source code that may accompany it. In no event shall the publisher and the author be liable for any loss of profit or any other commercial damage caused or alleged to have been caused directly or indirectly by this document.

Printed May 2009 in Canada.

Table of Contents

Part 1 Introduction

Part 2 Quick Start Guide

Part 3 Usage

3.1	Button Bar & Menu	3
3.2	Project	4
3.2.1	Query Options	5
3.2.2	Query Types	6
3.2.3	Exporting Results	8
3.2.4	Advanced Project Settings	9
3.2.5	Excluded Pages	10
3.2.6	Passwords and Forms	12
3.3	Scan Results	14
3.4	Query Library	14
3.5	Project Import and Export	15
3.5.1	Exporting Project Files	15
3.5.2	Importing Project Files	16
3.6	Scheduler	16

Part 4 About Inspyder Software

Part 5 License Agreement

Index

1 Introduction

Power Search from [Inspyder Software Inc.](#) is a Windows based utility to search web sites for words and phrases. Inspyder Power Search is like having Ctrl-F ("Find") for your entire website. If you have ever wanted to know which of your pages contained a particular word, phrase or filename, the answer is here.

Most modern websites contain scripts that generate content when a page is loaded. Content may come from a database, other files, or be pulled from a different website or server. If that's the case, then searching the local files on your hard drive isn't the same as searching your finished website. This is where Power Search is different. Power Search searches the pages delivered by your web server, providing results from the complete site, not script source code and incomplete content.

Applications of a client side search utility like Power Search include:

- Search Engine Optimization - Quickly see which pages contain a particular keyword.
- Find Obsolete Content - Find references to obsolete products, materials, or employees.
- Search Another Site - Find information on a site that isn't yours. Quickly check for copyright notices and other text hidden in HTML code.
- Locate Commonly Misused Words - Find all occurrences of a misused word or phrase (such as "over their").

With Power Search you can crawl password protected and form submission pages with ease. Power Search supports both HTTP authentication and posting form data, making it easy for you to search all types of protected pages that make up your site.

Power Search can also be configured to exclude certain pages from its search, such as message boards and other pages you might wish to ignore.

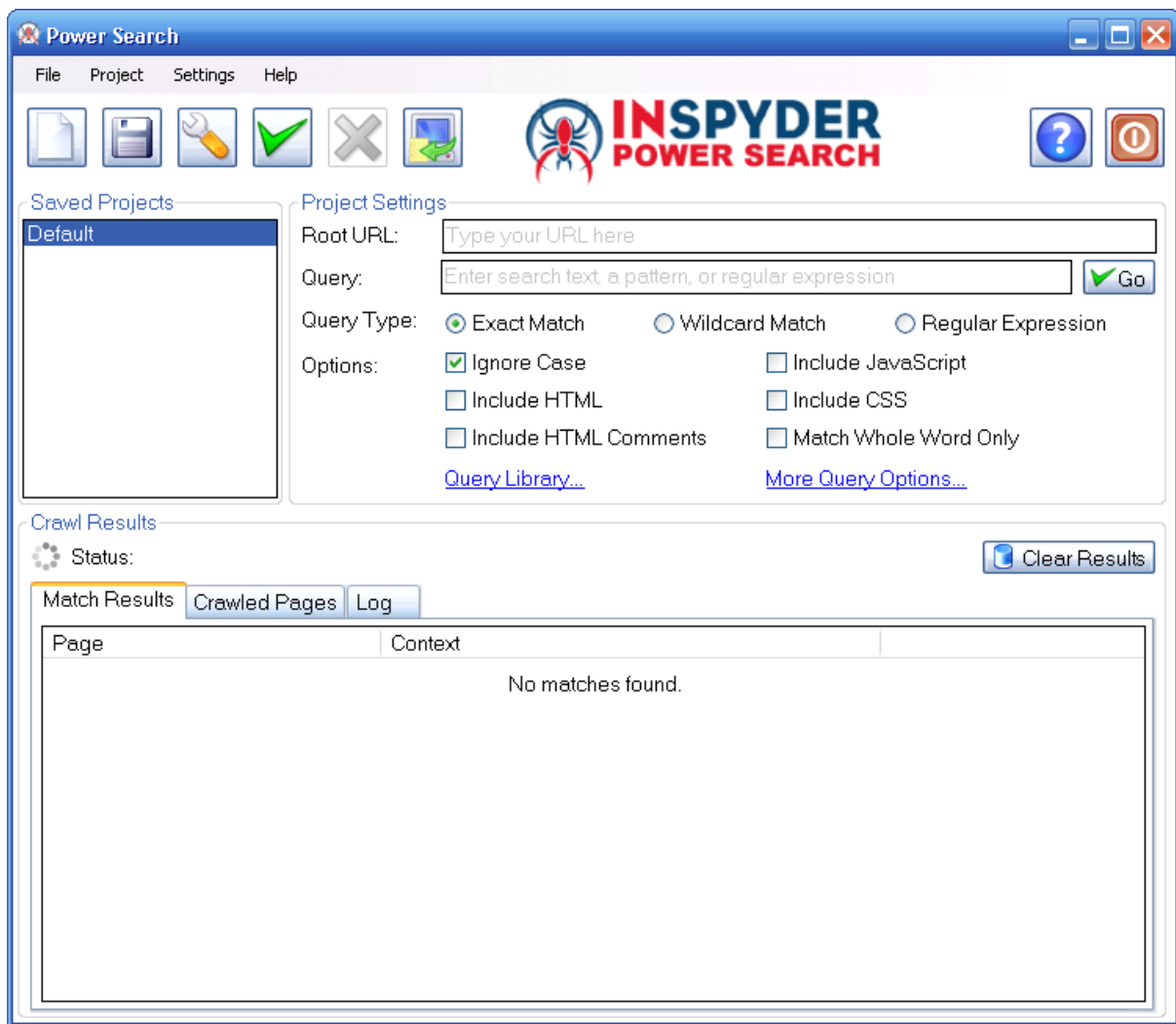
Features and Benefits

- Ease of use - Just enter your homepage URL, a word or phrase and start searching
- Choose to include HTML, scripting and style sheet syntax in the search
- Searches pages served by your web server, regardless of backend technology (PHP, Perl, ASP, ASP.NET, etc.)
- Construct wildcard searches
- Search password protected pages and submit forms - Power Search provides robust facilities for crawling password protected pages.
- Crawl only within your site - Power Search crawls only pages below the root page, so you don't have to be concerned about crawling pages that aren't yours.
- Export the results to a CSV format file for subsequent review.

Power Search requires Windows XP/2003/Vista/2008.

2 Quick Start Guide

- Step 1** The default project is now loaded but does not contain any settings. Enter the URL of the website you wish to search in the 'Root URL' field.
- Step 2** Enter your query text in the Query field and select the match type (for a simple search, use "Exact Match")
- Step 3** Click the "**Go**" button.
- Step 4** Verify your results in the Crawl Results section and make corrections to your website.



Hint: Try modifying the [Advanced Project Settings](#) to refine how Power Search crawls the website. The Advanced Project Settings are accessible from the main menu bar. Click "**Project | Advanced Project Settings...**" at the top of the screen.

3 Usage

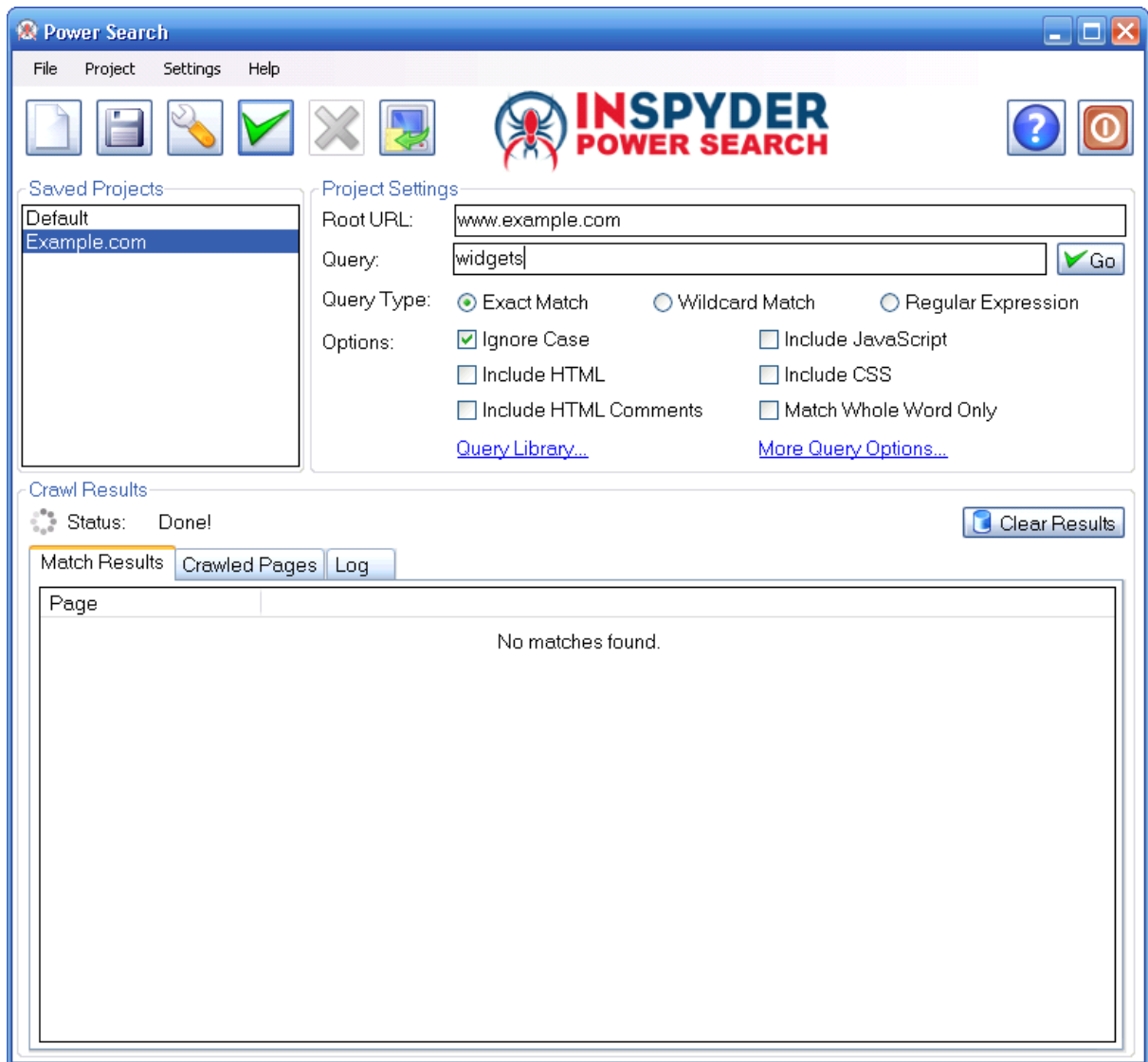
The Power Search user interface area is broken down into four main areas:

Button Bar These buttons provide quick access to frequently used functions within Power Search.

Saved Projects This is a list of the currently saved projects. When you create a new project, it will appear in this list. Projects can be deleted or renamed by right clicking the project name in the project list. To switch projects, click on a different project name in the list.

Project Settings This is where the settings for the currently select project will appear.

Crawl Results This is where the output from Power Search will be presented.



The following sections describe the individual fields and controls of Inspyder Power Search's user interface.

3.1 Button Bar & Menu

This section describes the functions of the button bar controls. These are the most often used features of Power Search and are placed on the button bar for convenience. Some of this functionality is available by right-clicking on controls in other parts of the user interface, or from the menu bar at the top of the screen (File, Project, etc.).



The **New Project** button is used to establish a new website that you wish to analyze. You will be prompted for a shortcut name, which will appear in the Projects list, and the root URL of that web site.



The **Save Project** button is used to save the current project settings. The main menu contains a **"Save As"** command if you want to replicate the current project. As well, you can right-click on any project name in the Saved Projects list to delete or rename the project.



This button allows you to access the advanced configuration options for the current project. For more information please refer to the [Advanced Project Settings](#) section of this document.



This button starts the crawling/searching process on the current Root URL. The crawl can be interrupted at any time by clicking on the **Stop** button.



Use this button to stop the web site checking process.



This button provides a way to export the crawl information. For more information please refer to the [Exporting Results](#) section of this document.



Provide access to the help file and About screen. The About screen contains version and your registration information.



Closes the program.

The main menu contains most of the functionality of the button bar, but extends that capability with a few additional commands. The most notable menu additions are the "**File | Save As**" and "[Settings | Scheduler](#)" commands.

3.2 Project

This section describes the options, fields and features of the Project Settings group.

Project Settings

Root URL:

Query:

Query Type: Exact Match Wildcard Match Regular Expression

Options: Ignore Case Include JavaScript
 Include HTML Include CSS
 Include HTML Comments Match Whole Word Only


[Query Library...](#) [More Query Options...](#)

Root URL - Specify the address of the homepage for the website to be analyzed. You can enter an IP address or the URL of the web page. The URL does not need to begin with the "http://" prefix. If your website requires the use of SSL, than you must specify the "https://" prefix.

Query - Enter the text you wish to locate on the website here. The Query can be a specific word or phrase, a wildcard match or a Regular Expression. If a Wildcard match or Regular Expression is used, a special format can be used to capture data from the website into the results. For more information, see [Query Types](#).

Go Button - Use this control to begin the analysis of the Root URL. This button has the same



functionality as the  control on the button bar. It is included in this location for convenience.

Query Type - The type of text that is entered in the Query field. For more information, see [Query Types](#).

Ignore Case - If this option is checked, the query will not be case sensitive.

Include HTML - If this option is checked, all HTML code will be included in the search. By default this option is unchecked, so only visible text on the website will be searched (HTML markup will be ignored).

Include HTML Comments - If this option is checked then all HTML comments will be searched. This could be used to search for hidden text, such as copyright notices which are often placed inside HTML comments.

Include JavaScript - If this option is checked then JavaScript code will be included in the search (this included inline JavaScript as well as JavaScript source files (*.js)).

Include CSS - If this option is checked then CSS code will be included in the search (this includes inline CSS, CSS in "<style>" tags, and CSS source files (*.css)).

Match Whole Word Only - If this option is checked Power Search will only return query matches that are surrounded by word boundary characters (such as spaces). For example, of the query were "ton" and Match Whole Word Only were enabled, Power Search would **not** match the word "button". Match Whole Word Only is only available is the Query Type is set to Exact Match.

Query Library - This link provides access to Power Search's list of predefined queries. When selecting a Query from the Query Library, all existing query settings (such as Ignore Case and Include HTML) will be set to match the needs of the newly selected query.

More Query Options - This link provides access to additional Query options not listed in the main Project Settings area.

For advanced users it may be of value to note that Power Search Projects and settings are stored in folder:

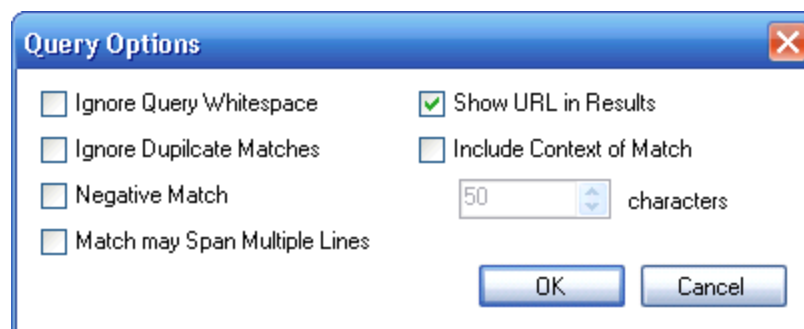
Windows XP:

C:\Documents and Settings\[username]\Application Data\Inspyder Power Search\

Windows Vista:

C:\Users\[username]\appdata\roaming\Inspyder Power Search\

3.2.1 Query Options



Ignore Query Whitespace - If this option is checked then any whitespace (such as " " or tabs) in the query will be ignored. This is typically used to help make complicated Regular Expressions easier to read.

Ignore Duplicate Matches - If this option is checked, Power Search will only show totally unique results (any result where all columns fields already exist will be hidden). This option is useful if you need to extract a list of unique names, email addresses or some other information and don't wish to manually clean duplicate results from the data.

Negative Match - If this option is checked only pages that do **NOT** contain the query or pattern will be shown. When this option is selected "Include Context of Match" becomes disabled (since a negative match cannot have a context). Show URL in Results will also become disabled (and checked), the URL is the only result returned when using a negative match (essentially providing a list of pages that do not contain the search query).

Match may Span Multiple Lines - If this option is checked then "*" in a Wildcard match will include newline (\n) characters. This is the same as the Regular Expression "Single Line" mode.

Show URL in Results - This option prints the URL that the match occurred on in the output. By default this option is enabled.

Include Context of Match - If this option is enabled the text around the match is shown in a column called "Context". Essentially this shows where on the page the match was found. The length of the context is determined by the number specified in the numeric field (by default, 50 characters).

3.2.2 Query Types

3.2.2.1 Exact Match

The Exact Match search type is used to find specific text on webpages. This functionality is typically used to find references to specific items, such a source files, old product or employee names, HTML code snippets, copyright notices, etc.

By utilizing the Negative Match feature, Power Search can also provide a list of pages that DO NOT contain a particular piece of text. For example, you could find all pages on a website that do not include a link to the company privacy policy.

Exact Match will find the entire text as it is entered into the Query field on the website. For more advanced matches see [Wildcard Match](#) and [Regular Expressions](#).

3.2.2.2 Wildcard Match

Wildcard searches are a way of easily matching patterns of text. The pattern can be very simple, or very complex. Wildcard support has been included in Power Search to make it easier to find complicated items, or groups of items, on your website.

Basic Wildcards

The basic wildcard character (*) lets us match zero or more characters in the search pattern. For example; "a*.html" matches:

```
a.html  
ab.html  
abc.html
```

To match a single character, we use the special character "?". For example; "a?.html" will match the

following:

```
aa.html  
ab.html  
ac.html
```

It will NOT match words like:

```
a.html  
abc.html
```

Ranges

To match a character within a range (for example, any number character) we can use ranges. A range is encapsulated within square brackets, [], and contains the set. For example, if we wanted to match 'inventoryX.html' where the 'X' is any digit, we could write `inventory[0123456789].html`. This will match the following:

```
inventory1.html  
inventory2.html  
inventory3.html
```

It will NOT match words like:

```
inventory12.html
```

As a shorthand, the following is also valid: `inventory[0-9].html`.

Multiple ranges within a single set of brackets is acceptable (for example, the letters A to F and the numbers 0 to 9). For example; `inventory[A-F0-9].html` will match

```
inventory1.html  
inventoryA.html  
inventoryB.html  
inventoryF.html
```

It will NOT match words like:

```
inventoryG.html
```

It's important to note that a range only matches a single character. To find `inventoryXX.html`, where "XX" was a number from "00" to "99", the following would have to be used: `inventory[0-9][0-9].html`. If the search is case sensitive, then [A-Z] and [a-z] are different ranges.

Alternates

Matches that take the form "one or the other" can be specified by using alternates. Alternates are enclosed in curly brackets, "{}", and each alternate is separated by a comma. For example, to find all of the words that start with "ins" and end with "pyder" or "ite" use the following syntax: `Ins{pyder,ite}`. This will match the following:

```
Inspyder  
InSite
```

But it will NOT match:

```
Inspect
```

Literals

To match a special character as is (such as '*' and '?'), then it must be preceded with a backslash ('\'). For example; 'question\?' will match the following:

```
question?
```

But it will NOT match:

```
question!
```

The same holds true for the '{', '}', '[' and ']' characters.

Extracting Data

In addition to Wildcard matches, Power Search provides a powerful mechanism for extracting data from websites. By including "#DataName#" in your search pattern (where "DataName" is the name of the column you wish to extract data into), any text found in place of "#DataName#" will be captured and saved.

For example, the title text of a webpage is stored between HTML title tags as follows:

```
<title>Inspyder Software Inc.</title>
```

We could create a Wildcard Match like:

```
<title>#WebpageTitle#</title>
```

(Remember to check "Include HTML") When we run this search on the website, each page's title will be extracted and stored in a column called "WebpageTitle".

It's possible to include multiple fields in the Wildcard Search. For example:

```
<IMG*src="#Source#"*alt="#AltText#"*>
```

This search will extract the "src" and "alt" attributes from all images (where the alt attribute is present). We could enhance this pattern to make the "alt" attribute optional:

```
<IMG*src="#Source#"*{alt="#AltText#",}*>
```

By adding the Alternate syntax ("{alt="#AltText#",}") the match includes "alt="#AltText#" OR nothing" (since nothing follows the comma).

3.2.2.3 Regular Expression

In addition to Exact Match and Wildcard Match, Power Search has the ability to apply Regular Expressions to the webpages it crawls. The syntax of Regular Expressions is beyond the scope of this manual, but many excellent tutorials are available online, including:

<http://www.codeproject.com/KB/dotnet/regextutorial.aspx>

To capture data on a website using a Regular Expression, the expression must include one or more Named Groups. For example, a Regular Expression to match email addresses looks like:

```
[a-z0-9\.]*@[a-z0-9]*\.[a-z0-9\.]*
```

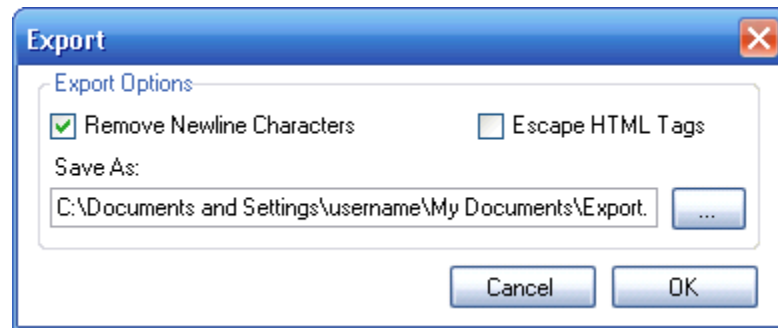
This regular expression will return all pages that contain at least one email address. To capture (or parse) the addresses into a column called "Email", we need to augment the expression as follows:

```
(?<Email>[a-z0-9\.]*@[a-z0-9]*\.[a-z0-9\.]*)
```

The Regular Expression syntax used by Power Search is the Microsoft .NET syntax.

3.2.3 Exporting Results

In order to facilitate additional analysis of collected data, Power Search provides the ability to export (or save) collected data to CSV files. The CSV (Comma Separated Value) format can be opened by most spreadsheet and database applications (for example, Microsoft Excel and Microsoft Access). It is also possible to open a CSV file with a simple text editor such as Notepad.

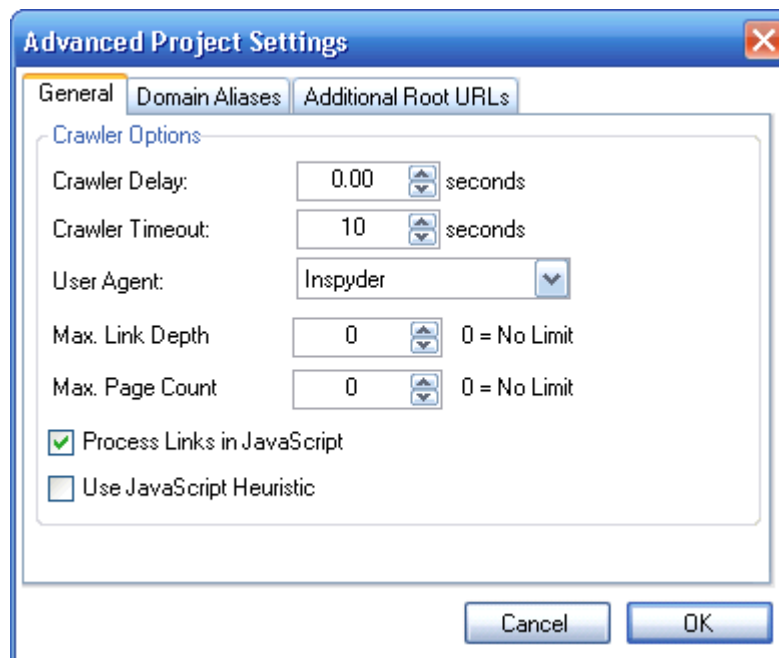


Remove Newline Characters - If this option is checked, newline characters will be removed from the results. This should always be checked unless you specifically require the newline characters in your results. Failure to remove the newline characters may prevent the data from importing into your application correctly.

Escape HTML Tags - If this option is checked Power Search will escape any HTML entities in your export with their HTML escape sequences. This should only be checked if your application is importing Power Search's files as HTML content by mistake. (In most cases, this can be safely left unchecked.)

3.2.4 Advanced Project Settings

To access the advanced project settings, click the 'Wrench' icon from the button bar, or choose "Project | Advanced Project Settings" from the main menu bar.



Crawler Options

Crawler Timeout - The Crawler Timeout value specifies the number of seconds that the crawler should wait for a response from your webserver. If your server is slow to respond or has long running scripts, increase this value.

Crawler Delay - The Crawler Delay value specifies how long the crawler should pause between HTTP requests to your server. If Power Search places too high a load on your server,

increase this value to slow the crawl down.

User Agent - The User Agent value indicates how Power Search should identify itself to your webserver. By default, Power Search identifies as the Inspyder Crawler. If your site uses browser detection scripts, you may wish to have Power Search masquerade as Internet Explorer or FireFox.

Max Link Depth - This option limits the depth of the crawl for the current project. The depth of the crawl is the number of links (or 'mouse-clicks' for a user browsing your website) from the root URL. For example, if the link depth limit was 1, then all the pages that could be reached in one mouse-click from the root URL would be crawled. If the link depth is 0, then there is no limit.

To limit the crawl by path depth (such as "no pages deeper than 3 subdirectories") you can create an entry in the Exclusion List that looks like: `"/**/*/*"`. This would exclude any URLs that had 3 or more '/' characters.

Max Page Count - This option limits the crawl by the number of pages crawled. This feature is useful if you are unsure of the website's size and only wish to survey a fixed number of pages.

Process Links in JavaScript - If this option is checked Power Search will attempt to locate URLs within JavaScript code.

Use JavaScript Heuristic - If this option is checked Power Search will use a more aggressive technique for locating links within JavaScript. If a site makes heavy use of JavaScript (such as for the menu system or other navigation) it may be required to enable this feature.

Domain Aliases

Domain aliases are used to tell the crawler that your site has more than one domain name. This feature is useful if your site is referenced internally by more than one domain. For example, many sites use "www.example.com" and "example.com" interchangeably. If the root URL has the "www" prefix, links using the other domain name will be treated as external to the site (and ignored) unless it is specified as a domain alias.

Additional Root URLs

Additional Root URLs are used to 'seed' the crawler if your site is not fully cross-linked. If you have doorway pages, or a disconnected section of your site (to which no links point), you can enter these pages as Additional Root URLs. The crawler will process these pages while crawling your site.

To add a new Root URL, simply enter it in the URL field and click 'Add'. To remove a URL from the list, select it and press the delete key ('Del') on your keyboard.

3.2.5 Excluded Pages

Often it is necessary to exclude some sections of a website from crawling. To do this, Power Search provides the ability to exclude specific URLs or groups of URL from crawling.

The Excluded Pages window contains full or partial URLs that should be skipped when crawling. This is important if you have pages or scripts that generate a large amount of redundant content (such as a message board or article pages that are formatted for printing). You might not want these included in the crawl.

When a directory (like `"/blog"`) is specified as an Excluded Page, then all of its sub-directories will be excluded (such as: `www.example.com/blog/index.html` and

www.example.com/blog/archives/january2005.html).

Wildcard matching is also supported through the use of the '*' (asterisk) character. The '*' will replace zero or more characters as shown in the examples below. The '*' can be used multiple times in a string.

*inventory/warehouse*index* will exclude the following:
inventory/warehouse1index
inventory/warehouse125index
inventory/warehouseindex

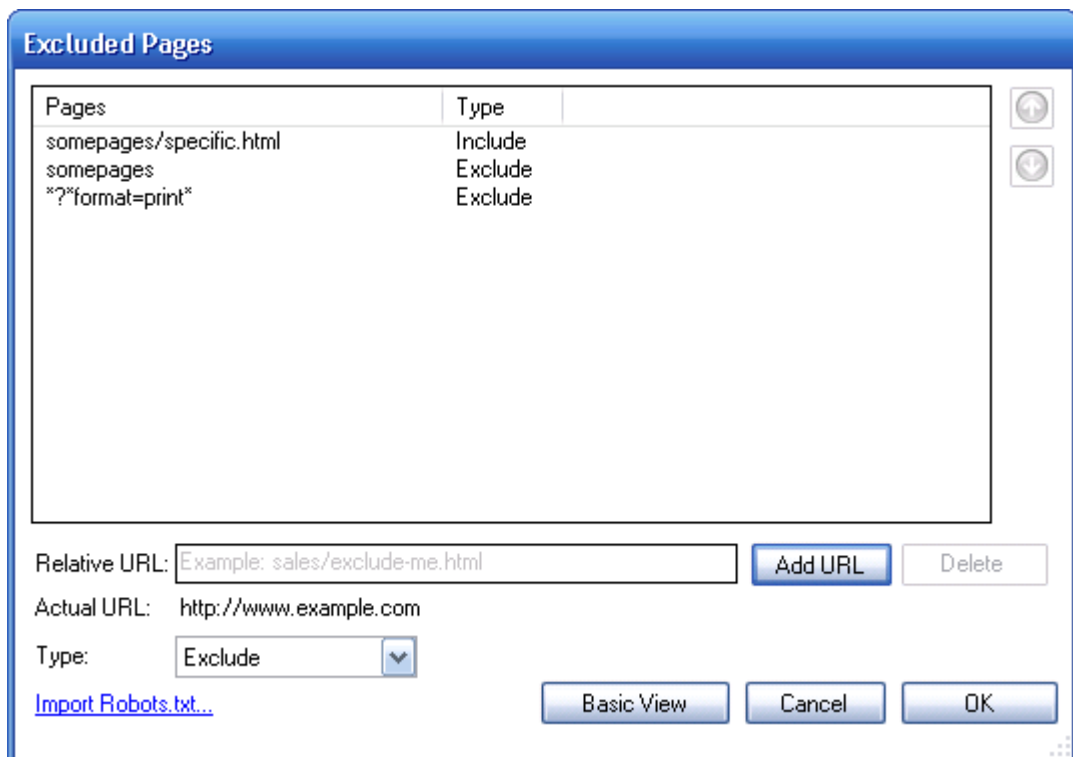
userreports/.txt* will exclude all files with an extension of ".txt"

It is also possible to use the Exclusion List to exclude URLs that contain a particular parameter. For example, to exclude all urls that contain "format=print" in the URL parameter list, you could enter: **?*format=print**. This will exclude all URLs that contain 'format=print' somewhere in the URL parameter list (after the '?').

To edit the Excluded Pages list, click **Project | Excluded Pages** from the main menu bar. Internal URLs should be relative to the root URL. URLs beginning with '/' are taken from the domain root (this is not recommended). URLs **not** beginning with '/' are taken from the start of the crawl. For example, if the root URL is "www.example.com/somepage/":

'apage.html' becomes 'www.example.com/somepage/apage.html'
 '/apage.html' becomes 'www.example.com/apage.html'

As a shortcut, a page can also be added to the exclusion list by right clicking the URL in the Results tab.

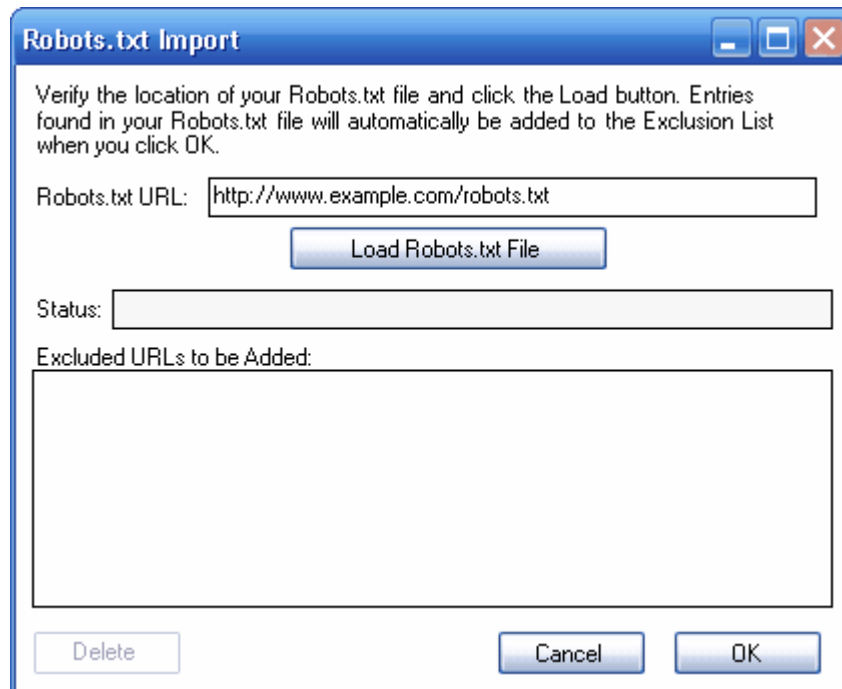


By clicking the "Advanced View" button on the exclusion dialog you are able to mark a page or section as specifically **included**. The list is matched from top to bottom. If you want to exclude all pages under

"www.example.com/content/" except "www.example.com/content/important" than you would create an entry that includes "/content/important" above the rule that excludes "/content". You can move the rules up and down with the arrows on the right.

3.2.5.1 Importing Robots.txt

Some websites already contain a special file called "robots.txt" which tells crawling software which URLs (or paths) to ignore. Power Search does not automatically read and follow robots.txt because it is often necessary to crawl the entire website. However, Power Search has the ability to import the robots.txt file from the website into the Excluded Pages list. To access this feature, click "Import Robots.txt" on the [Excluded Pages](#) window.



When the Robots.txt Import window opens, it will automatically populate the "Robots.txt URL" field based on your Root URL. You can modify this field if necessary. Click "Load Robots.txt File" to fetch and load the file from the website.

If Power Search is unable to load the file, an error message will appear. Double check the URL and try again.

If Power Search is able to load the file, the results will be placed in the "Excluded URLs to be Added" field. You can remove any unwanted entries now by highlighting the entry and click the "Delete" button. If you click "Cancel" at any time no changes will be made to your "Excluded Pages".

When you click "OK", any entries found in the "Excluded URLs to be Added" list will be automatically added to your "Excluded Pages" list.

Only entries under the "*" user-agent will be loaded from the robots.txt file. Entries for other user-agents will be ignored.

3.2.6 Passwords and Forms

If your website is password protected or contains content hidden behind a form submission, you can configure Power Search to access that content using the "Passwords and Forms" feature.

Passwords and Forms allows you to create rules which Power Search will use each time it crawls your website. If a page is encountered with a form rule, it will post the specified information (usually a username and password) to the server when that page is encountered.

Password protected pages can be configured by clicking **Project | Passwords and Forms** from the main menu bar. A window will appear that will let you enter the relative URL of the password protected page, as well as the authentication method and the credentials. If the HTTP method is selected then a username, password and domain can be entered into the appropriate fields. Note that the domain field is optional.

There are two methods for supplying login information to a website. The first (and most common method) is the POST method. If you enter your login information into the website itself (or are filling out some type of web based form) use this method. Click the 'Wizard' button to launch the Inspyder mini-browser. Use this browser to log into your website. When you complete the login form, Power Search will ask if you want to save the form data, click 'Yes' and Power Search will automatically create the necessary rule for you.

The send method is the HTTP method. This is used when the server (rather than the website) is providing the authentication method. If you see a popup login window (with username and password fields) when you access your website with Internet Explorer or FireFox, then your website is using this authentication method. To configure this rule, enter the Relative Path of your login page (if it is your Root URL, just enter "/"), along with the username and password in the corresponding fields. Click 'Add' to save the rule.

Refer to the Excluded Pages section of the manual for more information on how to format the Relative URL field.

Page	Method	Credentials
/admin/ login.php	HTTP POST	user1:***** username=user1&passw

Relative URL:

Actual URL:

Method:

Username:

Password:

Domain (Optional):

3.3 Scan Results

This section describes the options, fields and features of the Scan Results group.

Clear Results

Click to reset the contents of the **Scan Results** section, including all of the tabbed dividers.

Match Results

The Match Results tab contains the search results found by Power Search. When searching for an Exact Match, the Match Results will show the URL that contains each match along with the Context of the match. (The Context is where on the page the Query text was found.)

If you are using Wildcard or Regular Expressions, then the Match Results will adjust to include any data fields (or Named Groups) that are specified in the Query text. It is also possible to disable the URL and/or Context when using either of these two Query types.

For information on saving (Exporting) the data from the Match Results tab, see [Exporting Results](#).

Crawled Pages

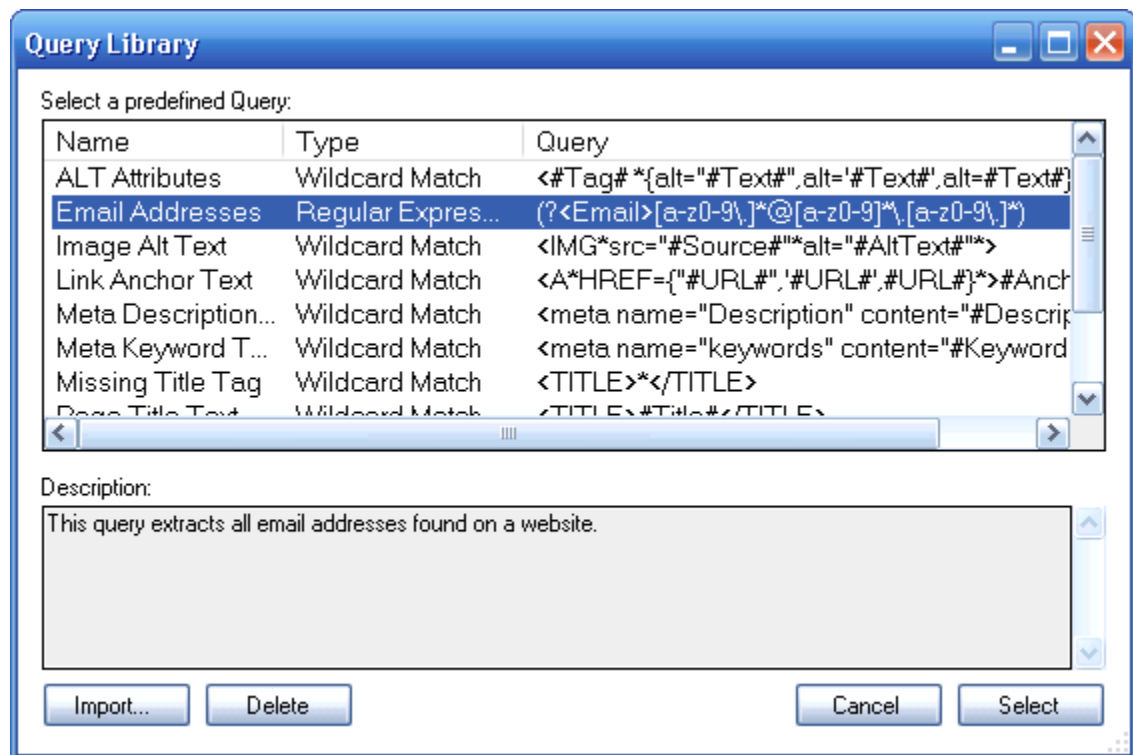
The Crawled Pages tab contains a list of each page that was crawled and analyzed by Power Search. Only HTML, JavaScript and CSS files are analyzed. (Power Search will analyze any file containing HTML content regardless of the file extension, PHP, ASPX, and other formats are supported.)

Log

The Log tab contains information that can be useful when troubleshooting crawling problems. The log indicates each time a crawl starts and stops, as well as information regarding errors that have occurred.

3.4 Query Library

The Query Library provides access to a set of predefined, commonly used Queries. When you select a Query from the list, an English language description of what that query does is presented in the Description field.



It's possible to add new Queries to the library using the "Import..." button. Occasionally Inspyder may publish new queries to our website. You can add these to your library by downloading them, then Importing them with this functionality. It is also possible to import an existing Power Search Project file into the Query library (the non-Query components of the Project will be removed automatically).

To remove a Query from the Query Library, select it and click the Delete button.

When you click the "Select" button, the settings from the currently selected Query will be applied to your Project. Your Root URL, Excluded Pages and other settings not related to the Query will remain unchanged, however the Query text and any other special settings will be adjusted (such as case sensitivity) for the query to work correctly.

3.5 Project Import and Export

Power Search provides the ability to import and export projects. These features are useful if you need to move projects to a different PC, wish to backup your existing projects, or copy projects between different Inspyder applications.

3.5.1 Exporting Project Files

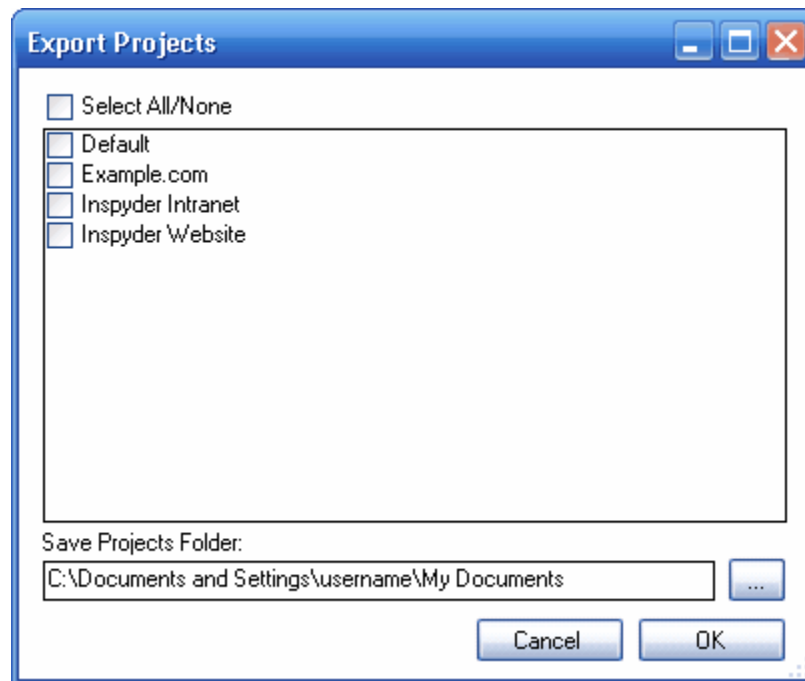
To export one or more projects, click **"File | Export Project Files..."** from the main menu bar. The Export Projects window will open. Here you can select one or more project files to export and the location where they will be saved.

Each exported project is saved to a unique file. If you chose to export multiple projects, multiple files will be created in the selected output folder. The filenames will have the following format:
"Project Name"."Application".proj

"Project Name" is the name of the project in Power Search. "Application" will be the Inspyder application abbreviation and is used to determine which application generated the project export. All Inspyder project files end with the ".proj" file extension.

Exporting projects from Power Search will not remove them from your list of saved projects. Exporting a project makes a copy of the project file. Any changes made to the project file will not be reflected unless you Import the project back into Power Search.

If you wish to remove a project after exporting it, right click the Project name in the list of Saved Projects and select 'Delete Current Project'.



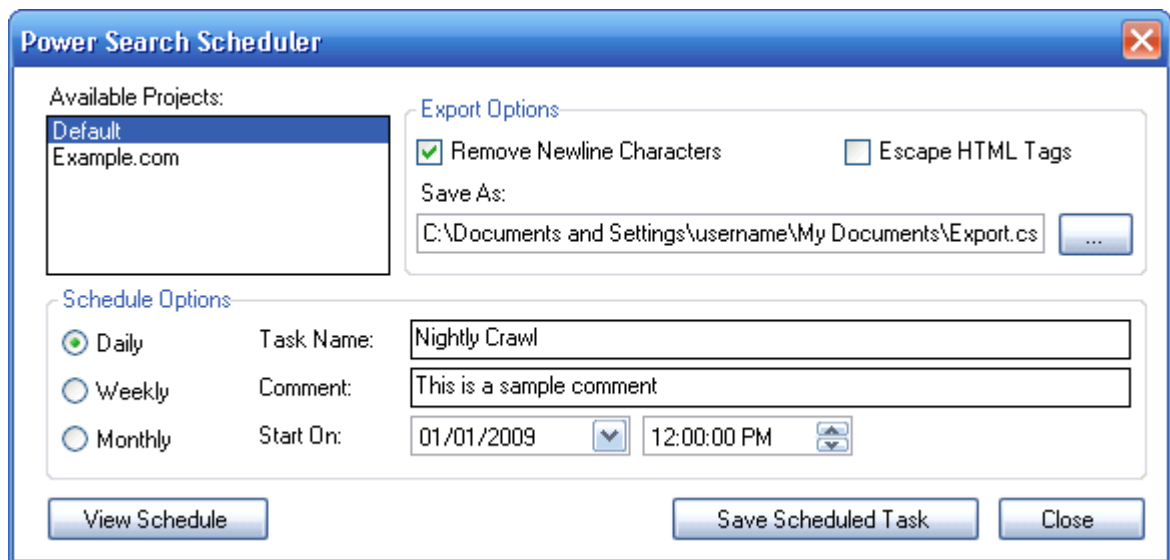
3.5.2 Importing Project Files

To Import one or more project files into Power Search, select "**File | Import Project Files...**" from the main menu bar. An Open File window will appear that will enable you to select one or more Inspyder project files.

It is possible to import a project file from an Inspyder application other than Power Search. If you do, only the common elements between applications will be imported (such as Root URL, Excluded Pages, etc.). Application specific settings will be ignored. If you are importing a Power Search project back into Power Search, then all settings will be used.

3.6 Scheduler

Power Search has the ability to run in an unattended mode and automatically export the results. To configure Power Search for scheduled operation simply select "**Tools | Scheduler**" from the main bar. The dialog below contains the choices from the [Export Options](#) dialog as well as those necessary for timing the execution of Power Search. Once the Schedule Options are filled in and the **Add Task** button is clicked the task will be scheduled using the Windows Task Scheduler.



Use the **View Schedule** button to view all the tasks submitted to the Windows Task Scheduler as well as **to modify the scheduling parameters**. If you double-click the scheduled entry in the Windows Scheduled Tasks dialog a new dialog will be displayed which contains all of the parameters for that task.

The actual command used to invoke Power Search is found in the 'Run:' field. You can modify the parameters used by Windows when it starts Power Search by **carefully** editing the contents of the Run field. The following are the command line arguments:

-rooturl=	(Optional) Overrides the Root URL configured in the Project
-escape_html	Indicates that the output should have HTML content escaped (for more information, see Exporting Results)
-remove_newlines	Indicates that the output should have newline characters removed (for more information, see Exporting Results)
-output=	Specified the output file for the results (i.e. C:\MyFile.html or someone@company.com)
-project=	The project name to be analyzed (i.e. -project=Inspyder Local)

Note that if the command line argument's parameter contains one or more spaces, then the enter argument must be enclosed in quotes. (i.e. "-output=My Report.csv")

Note to Windows XP and 2003 users: The scheduled tasks will not run on an account with a blank password. For a workaround to this please see our online Knowledge Base.

Scheduled Execution Return Codes:

These are the codes returned to the Windows Task Scheduler after the Power Search task has been executed. You can view them in the Scheduler window to help troubleshoot any problems.

- 0** Success (no errors)
- 1** Not registered
Solution: Run the program interactively and enter your registration information when prompted.
- 2** Export Error
Solution: Try running the project and exporting in interactive mode. Check for file in-use, disk full and other possible errors.
- 3** Crawler Error

Solution: The crawler detected some type of unhandled error during crawling. Run this project in interactive mode and check for errors (corrupt HTML, etc.).

4 About Inspyder Software

Founded in 2004, Inspyder Software is a leading provider of web crawling technologies for content analysis, SEO and website management. These technologies have formed the foundation of our software products and made 'Inspyder' a recognizable brand in the software industry.

Our products are used all over the world by large and small businesses, from independent web designers to Fortune 500 companies. Our user base includes customers in over 50 countries around the world.

For more information regarding future products, custom development or consulting services, please contact us at sales@inspyder.com.

5 License Agreement

This legal document is an agreement between you, as licensee and Inspyder Software Inc. ("Inspyder"), as licensor. You should carefully read the following terms and conditions before using this product. Using this product indicates your acceptance of these terms and conditions.

DEFINITIONS: "Product" means (a) all of the contents of the files, disk(s), CD-ROM(s) or other media with which this Agreement is provided, including but not limited to Inspyder or third party computer information or software; written documentation or documentation files; and (b) upgrades, modified versions, updates, and additions. "Use" or "Using" means to access, install, download, copy or otherwise benefit from using the functionality of the Product in accordance with the documentation. "You" and "Your" means the purchaser of the Product.

GRANT OF LICENSE: The Licensor grants to You a non-transferable non-exclusive license to use the product on a single computer. You may physically transfer the product to more than one computer provided the product is used only on a single computer. You may not modify, adapt, translate, reverse engineer, decompile, disassemble or create derivative work based on: (a) the product; (b) written material associated with the product; (c) the concepts or technology utilized in this product.

COPY RESTRICTIONS: You may not copy the product including any product that has been modified, merged or included with other products except as specified in this Agreement, nor may You copy any written materials associated with the product. You shall be held legally responsible for any copyright infringement that is caused or encouraged by Your failure to abide by the terms of this license.

TRANSFER RESTRICTIONS: The product is licensed only for You, and You may not transfer the product or the license to use the product without Licensor's prior written consent. Any authorized transferee of the product shall be bound by the terms of this Agreement. In no event may You transfer, assign, rent, lease, sell or otherwise dispose of the product on a temporary or permanent basis except as expressly provided for herein. The product cannot be resold.

USAGE RESTRICTIONS: The product is licensed only for Your personal use. You cannot use the product to provide a service to others. To be clear, you can use the product as part of your web development activities, including web development for others. However, you cannot use Power Search to provide an exclusive service such as using Power Search as part of a

web checking or analysis service.

LIMITED WARRANTY: Except as expressly set forth herein, the product is provided 'AS IS' without warranty of any kind, either express or implied, including, but not limited to the implied warranties of merchantability and fitness for particular purpose. Inspyder Software Inc. does not warrant that the function of this product will be error free. However, Inspyder Software Inc. does warrant the media on which the software is furnished to be free from defects in material and workmanship under normal use for a period of 30 days from the date of delivery to You.

LIMITATIONS OF REMEDIES: The Licensor's entire liability and Your exclusive remedy shall be: (a) the replacement of any media not meeting Licensor's 'Limited Warranty' and which is returned to Licensor or an authorized representative of Inspyder Software Inc. with a copy of Your receipt; or (b) if Inspyder Software Inc. is unable to deliver replacement media which is free of defects in materials or workmanship, You may terminate this Agreement by returning the product and Your money will be refunded. In no event will Licensor nor anyone involved in the creation, production or distribution of the product be liable to You for any direct, indirect, consequential or incidental damages including any lost profits, lost savings, lost business revenue or other commercial or economic loss arising out of the use or inability to use the product even if Licensor or any authorized representative of Licensor has been advised of the possibility of such damages or for any claim by any other party.

JURISDICTION: The laws of the Province of Ontario, Canada shall govern this Agreement.

All rights are reserved; Copyright © 2009 Inspyder Software Inc.



Inspyder Software Inc.
202-3060 Mainway,
Burlington, Ontario,
Canada, L7M 1A3

Tel: 888.732.6134

Web: <http://www.inspyder.com>
Email: support@inspyder.com