Inspyder OrFind
User's Reference Manual

Copyright © 2015 by Inspyder Software Inc.

Inspyder OrFind (OrFind)
User's Reference Manual

Copyright © 2015 by Inspyder Software Inc.

Printed July 2015 in Canada.

Table of Contents

1   Introduction

OrFind from [Inspyder Software Inc.](#) is a Windows based utility to analyze web sites and provide detailed results on orphaned or unused files. OrFind is a convenient mechanism to free up disk space by identifying and removing obsolete or misplaced files. OrFind works by comparing your website content that is accessible by browsing your site, with the files that are present on your hard drive or server. Any files that reside on your server but are not accessible through the website are reported as orphans. OrFind also keeps track of file adds, changes and removals. Regular use of OrFind will help detect any file changes on your server which are the result of unauthorized access.

Features and Benefits

- Finds files linked through HREF, IMG, SRC tags and even files included in CSS
- Crawls dynamic content to include files missed by static orphan analysis tools
- Schedule regular web crawls and email the results automatically
- Automatically delete orphans
- Exclude files or subdirectories containing supporting script files
- Export orphan list for manual cleanup

OrFind requires Windows XP/Vista/7/8 or Windows Server 2003/2008.

2   Quick Start Guide

When you first run OrFind the default project is loaded but it does not contain any settings. The following steps will guide you to setting up OrFind for a first run on your website.
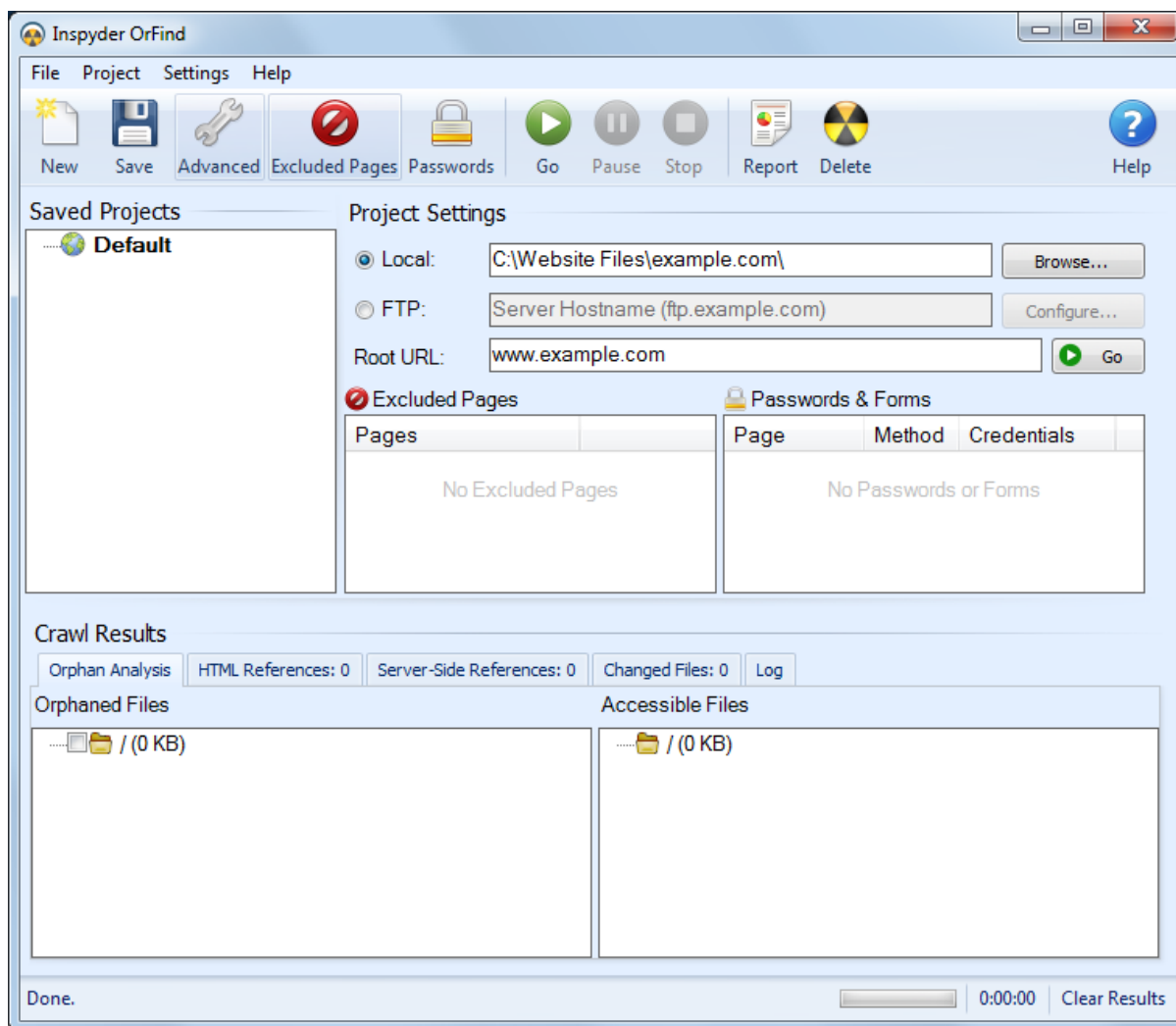
Step	Description
Step 1	Select either "**Local**" or "**FTP**" depending on where your website files are stored. If you select "**Local**" go to **Step 1a**, if you select "**FTP**" go to **Step 1b**.
Step 1a	Click "**Browse...**" to select the folder where your website files are stored on your local hard drive.
Step 1b	Enter your **FTP server hostname** and click "**Configure...**" to specify your FTP login credentials.
Step 2	'In the '**Root URL**' field enter the URL of the website that corresponds to the path configured in Step 1.
Step 3	Click the '**Go**' button.
Step 4	Review your orphaned files in the "**Orphaned Files**" tree. Files which are not orphaned are shown in the "**Accessible Files**" tree.
Step 5	**(Optional)** Check one or more files or folders in the **Orphaned Files** tree then click "**Delete**" to backup and delete these files from your hard drive or server.

3   Toolbar

This section describes the functions of the tool bar shortcuts. These are the most often used features of OrFind and are placed on the tool bar for convenience. Some of this functionality is available by right-clicking on controls in other parts of the user interface, or from the main menu bar at the top of the screen (File, Project, etc.).
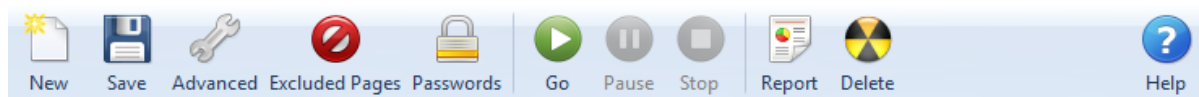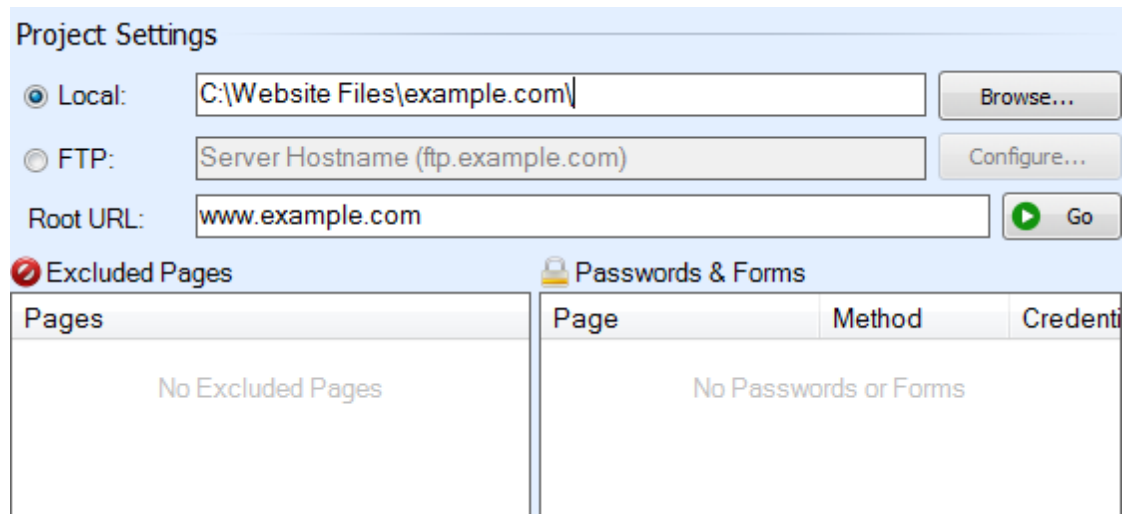
- **New** - The **New Project** button is used to establish a new website that you wish to analyze. You will be prompted for a shortcut name, which will appear in the Projects list, and the root URL of that web site.
- **Save** - The **Save Project** button is used to save the current project settings. 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.
- **Advanced** - This button allows you to access the Advanced Project Settings for the current project. For more information please refer to the Advanced Project Settings section of this document.
- **Excluded Pages** - Opens the Excluded Pages window which can be used to control which pages the crawler visits on a website.
- **Passwords** - Opens the Passwords and Forms window which is used for configuring access to password protected websites.
- **Go** - This button starts the crawling process. The crawl can be interrupted at any time by clicking on the **Stop** button.
- **Pause** - Pauses the crawl/analysis
- **Stop** - Use this button to stop the crawling process.
- **Report** - Opens the report wizard to create PDF, Excel or CSV reports. For more information please refer to the Exporting Results section of this document.
- **Delete** - Opens the delete window for configuring backup and removal of selected orphaned files from your server or hard drive.
- **Help** - Opens the help file.

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.

4   Project

A project contains settings that are typically associated with a single website. This section describes the options, fields and features of the Project Settings area.

For OrFind to work properly it must have access to your website (through HTTP, like a web browser does) and access to your actual website files (through FTP or local file access). It compares the files found on your website through HTTP with the files found through local access and determines which files are not linked from your website's HTML content.

If your website's files are stored locally on your hard drive, select the "**Local**" radio button and click "**Browse...**" to locate the folder where your website's files are saved.
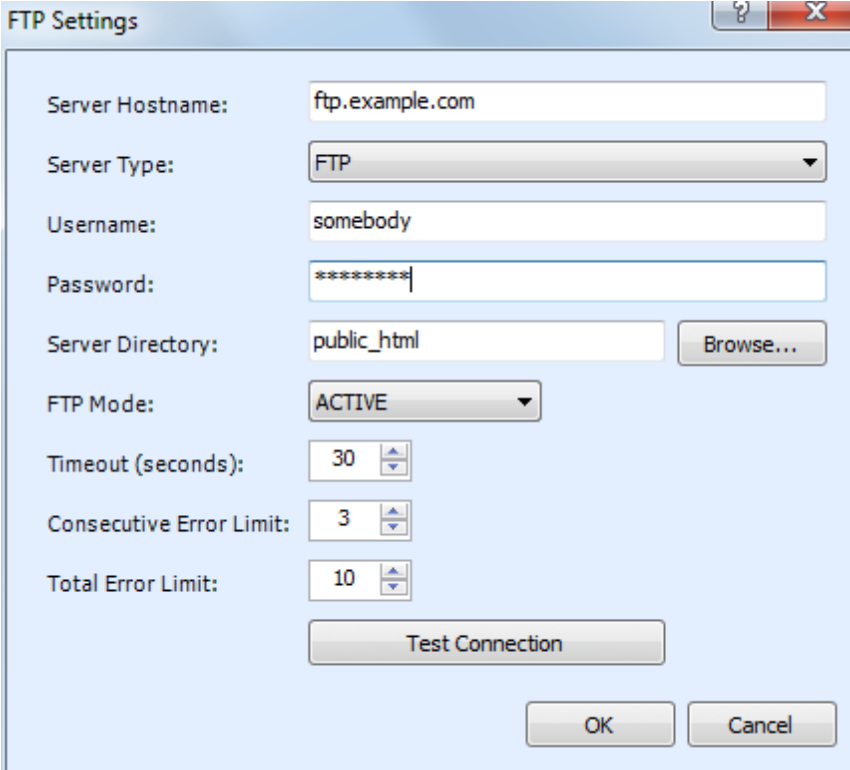
If your website's files are accessible by FTP, SFTP or FTPS choose the "**FTP**" radio button and click "**Configure...**". More information on the FTP specific configuration is available in the FTP Settings section.

Finally enter the URL of the website you are analyzing in the Root URL field. Make sure the Root URL you choose corresponds directly to the files stored at the path configured (Local or FTP). If your website files are stored in a different folder than the initial folder when you connect (such as "public_html" or "wwwroot"), then be sure to configure that in your FTP settings. This is one of the most common configuration mistakes. (**Tip:** If OrFind identifies all your files as orphans, it was either unable to crawl your website or it is configured to search the wrong local/FTP folder for website files.)

When your Root URL and file access parameters are configured, click the "Go" button to begin the crawl.

4.1   FTP Settings

The FTP settings section allows you to configure how OrFind communicates with your FTP server. These settings have no effect if your are not using the FTP connection method. To access the FTP settings, click "**Configure...**" beside the FTP hostname field on the main OrFind window.

Server Hostname - This is the hostname of your FTP server (it may be the same as your website's domain name, such as www.example.com). Do not include any protocol prefix or any path information.
    Valid:
        ftp.example.com
    Invalid:
        ftp://ftp.example.com
        ftp.example.com/public_html

Server Type - This setting specifies the protocol to use while communicating with your server. FTP, FTPS and SFTP are supported.

Username and **Password** - These are the username and password OrFind should use to login to your FTP server.

Server Directory - If you must change directories after connecting with your FTP server to access your website's files, enter that directory here. Use the 'Browse' button to have OrFind connect to the server, allowing you to find the folder graphically. The "**Browse...**" button becomes active when the Server Hostname, Username and Password fields are configured.

FTP Mode - Indicates if OrFind should use Passive or Active FTP when connecting to your server.

Timeout (Seconds) - The number of seconds to wait for a server response before an FTP timeout occurs.

Consecutive Error Limit - The number of consecutive FTP errors that can occur before aborting. If an error occurs before the Consecutive Error Limit is reached, OrFind automatically retries the operation.

Total Error Limit - The total number of FTP errors that can occur before aborting. This number should always be larger than the Consecutive Error Limit.

Test Connection - When you click this button OrFind will attempt to connect to your server using the FTP settings you have configured. It will verify that the "Server Directory" folder exists on the server.

4.2   Advanced Project Settings

The Advanced Project Settings dialog allows you to modify the behaviour low level behaviour of OrFind. Any changes made to the Advanced Project Settings will not take effect until you crawl your website again.
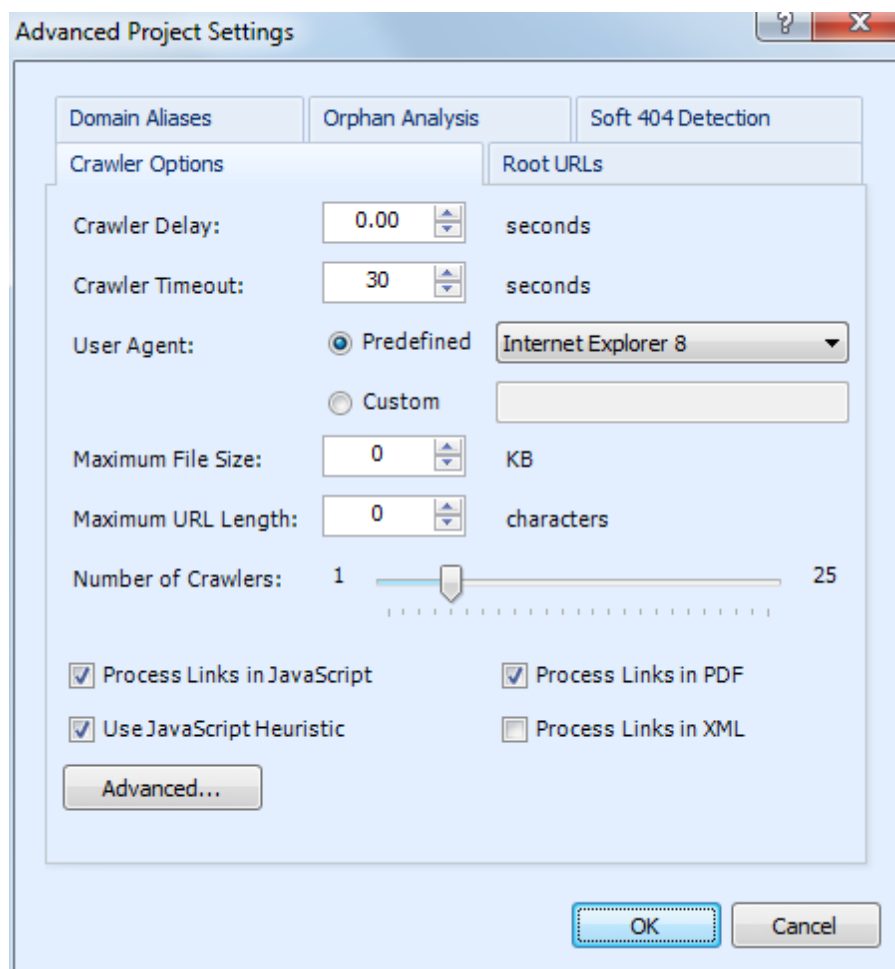
Orphan Analysis Tab

Default Filenames - This list identifies to OrFind which files your web server will look for when no filename is specified. For example, when you browse "http://www.example.com/" the server is loading an specific file (such as "default.html") even though it is not explicitly provided ("http://www.example.com/default.html"). The **Default Filenames** list allows you to tell OrFind which files your web server will look for when a path is referenced without an explicit filename. OrFind will then use

these filenames in the order provided to determine which file was accessed if no filename is explicitly provided by a particular link.

Parameter Separator - This tells OrFind which character your web server uses to determine where a filename ends and the URL parameter list begins. In most cases this is the "?" character. OrFind must strip parameters from filenames before performing orphan analysis.

Log Level - This option allows you to change the level of detail that is included in the "Log" result tab. If you encounter problems while crawling, it may be useful to change this value to "All" to provide more information on any problems that may occur.

Server-Side Code Analysis - If this option is enabled OrFind will inspect your code for server-side includes. This process can slow down orphan analysis (because all source files must be analyzed on your FTP server or local disk), but provides more accurate orphan analysis.

Currently OrFind supports the following server-side code types:
• ASP (.asp)
• ASP.NET (.aspx, .ascm, .asmx, .ashx, .cs, .vb)
• ColdFusion (.cfm)
• JSP (.jsp)
• PHP (.php)
• Server-Side Includes (.shtml)

Note that Server-Side Code Analysis only performs static code analysis. Dynamic references (files linked by string concatenation or runtime loading, for example) and binary references will not be detected.

Delete Read-Only Files - If this option is enabled OrFind will remove orphaned read-only files (by default these files will not be removed, even if selected).

Note: The "**Case Sensitivity**" option from version 3 has moved to "Crawler Options | Advanced...".

Crawler Options Tab

The Crawler Options (along with Root URLs) allows you to configure how OrFind crawls your website. It is important that OrFind crawls your site as thoroughly as possible to ensure that no false orphans are reported.

Crawler Delay - The Crawler Delay value specifies how long the crawler should pause between HTTP requests to your server. If OrFind places too high a load on your server, increase this value to slow the crawl down.

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.

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

Maximum File Size - This option limits the amount of data the crawler will download from a single URL. This is useful for websites that have either very large HTML pages (1 MB or more) or script

errors that cause pages to generate infinite amounts of HTML. If you receive "Out of Memory" errors while crawling your website with OrFind, try setting a limit of 1024 KB. For most websites setting this value is not required (use 0 for no limit).

Maximum URL Length - This option limits the maximum URL length the crawler will follow. This provides a mechanism to limit the crawler from crawling endlessly on websites that have infinitely recursive links. If you notice that OrFind seems to be crawling similar URLs over and over again, try setting this value to limit the problem. For most websites setting this value is not required (use 0 for no limit).

Crawler Threads - The crawler threads slider is used to adjust the number of simultaneous HTTP requests sent to your webserver. By default, OrFind is set to use 5 threads. Increasing this value can make OrFind crawl your site more quickly, but puts more load on your local computer and on your web server. If you find that your website is running slowly while crawling with many threads, you may need to decrease this value.

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

Use JavaScript Heuristic - If this option is checked OrFind 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 for OrFind to automatically discover all your pages.

Process Links in PDF - If this option is checked OrFind will attempt to locate URLs within PDF documents.

Process Links in XML - If this option is checked OrFind will attempt to locate URLs within XML documents.

Advanced Crawler Options

To access the Advanced Crawler Options window, click the "Advanced..." button on the Crawler Options tab page.

Use HTTP Compression - This option tells OrFind to use compression when communicating with your web server. If you see error messages that indicate problem with the "Zip Header", try disabling this feature.

Use HTTP 1.0 - This option tells OrFind to use HTTP 1.0 (rather than 1.1) when communicating with your web server. This should be unchecked unless you experience difficulties crawling a website.

Use HTTP Keep-Alive - This option tells OrFind to keep the HTTP connection (if possible) between multiple requests. Disabling this feature can have a negative impact on performance, but may be necessary with some older servers.

Obey "nofollow" Meta Tag - If checked the crawler will not follow URLs on pages with the no-follow meta tag.

Ignore Protocol Prefix - If checked OrFind will ignore the "http://" or "https://" protocol prefix when determine if a URL has already been crawled. With some websites that mix secure and non-secure sections together is may be necessary to uncheck this option to get a complete crawl.

Ignore HTTP Content-Type Encoding - This tells OrFind to ignore the HTTP "Content-Type" header and use the document encoding (if specified). You should only enable this option if instructed by technical support.

Parse Form "Action" Attribute - By default OrFind does not include URLs found in the "Action" attribute of HTML forms. If this option is checked OrFind will include these URLs when crawling.

Accept Cookies - This option tells the crawler to accept cookies from the website. The default is enabled.

Case Sensitive URLs - This option tells OrFind whether or not to use case-sensitivity when comparing file/folder names. This option is useful when applying OrFind to Linux/Unix based web sites.

Root URLs Tab

Use this page to specify Additional Root URLs. Additional Root URLs are used the 'seed' the crawler before checking your site. You only need to specify Additional Root URLs if your site is not fully cross-linked. For example, if you had some pages that were only linked to by an external website, then you may want to add one of those pages as an Additional Root URL so that the crawler is able to find them and index them before performing orphan analysis.

Domain Aliases Tab

Domain aliases tell OrFind that your site may use one of these as alternate domain names. If you consistently use the same domain name, or only use relative paths on your website, then you do not need to enter anything here. However, if you use multiple domain names ("www. example.com" and "example.com") to refer to the same site on the same server, than it is necessary to enter those alternate names here. You do not need to include the domain name given in the Root URL.

4.3   Soft 404 Detection

Soft 404 Detection allows OrFind to detect which links are broken even if your web server does not correctly return a "404 - Not Found" HTTP header. This is useful if you use a content management system that could report "Missing Pages" but don't necessarily return an HTTP status error. A common example of this is when an invalid product ID number is specified in some online shopping cart systems.

If you find that when crawling your website OrFind crawls lots of incorrect or invalid URLs it's a good indication that you need to configure Soft 404 Detection. Another way to test this is to change your Root URL to point to a page on your website you know does not exist, for example:
http://www.yoursite.com/thispagedoesnotexist
(Replace "yoursite.com" with your actual hostname). If OrFind doesn't detect an error on this bogus Root URL, then you'll need to configure Soft 404 Detection.

To configure Soft 404 Detection just enter some unique text that appears on your website's error page. In the example screenshot below we have configured "Page Not Found" and "Oops...sorry" as our Soft 404 Text. Any time OrFind sees a page with this text on it, it will treat it as an error page. To find out what text your website uses, use a regular web browser to access a URL on your website that you know does not exist.

Soft 404 Text:

Page Not Found
Oops...sorry

New Text: Example: File Not Found   Add
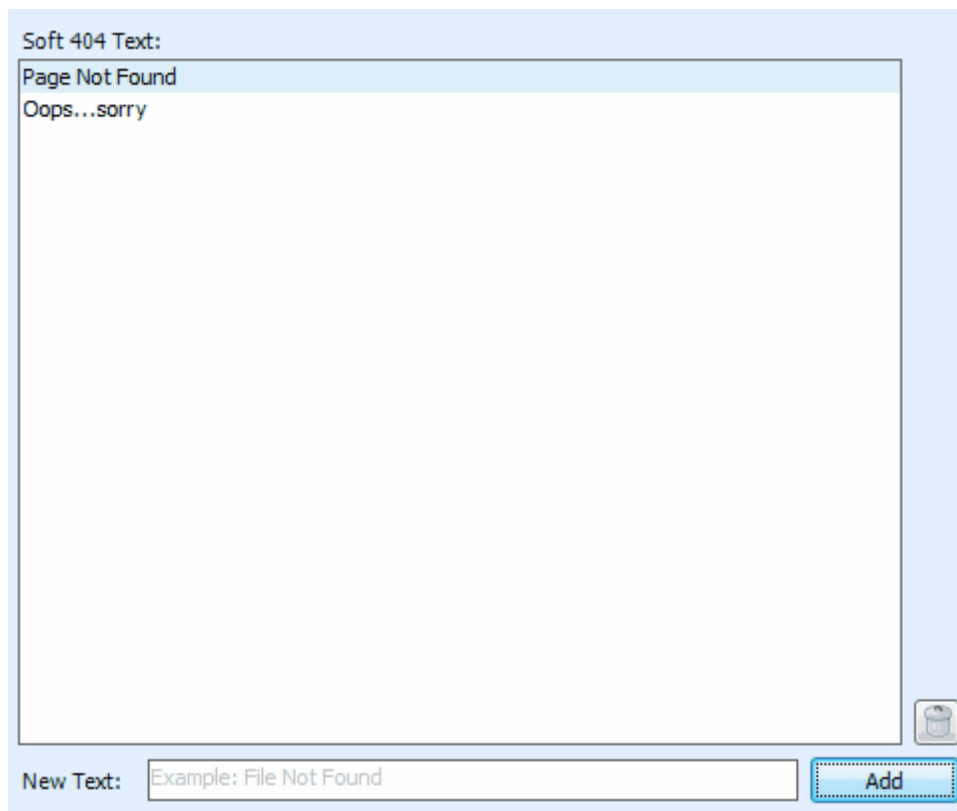
4.4   Clear Page Change History

The Clear Page Change History menu item (from the "**Project**" menu) allows you to delete the history file that OrFind uses to determine which files have been added, removed or changed since the last crawl. After cleaning the change log and crawling your website again, all files will appear as "New" in the "Changed Files" result tab.

Clearing the Page Change History only affects the current project.

4.5   Excluded Pages

Often it is necessary to exclude some sections of a website from crawling. To do this, OrFind provides the ability to exclude specific URLs or groups of URL from crawling. To access the Excluded Pages window, click "**Project | Edit Excluded Pages**" from the main menu bar.

The Excluded Pages window contains full or relative 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. To add a new URL to the list of Excluded Pages enter the full or relative path in the **Relative Path** field. (As you type the "Actual URL" field will show how your URL is combined with the project's Root URL.) Click the **"Add Rule"** button to save the new URL to the list.

To remove a URL from the list, select it then click **"Delete URL"** (under Actions), or press the "Del" key on your keyboard.
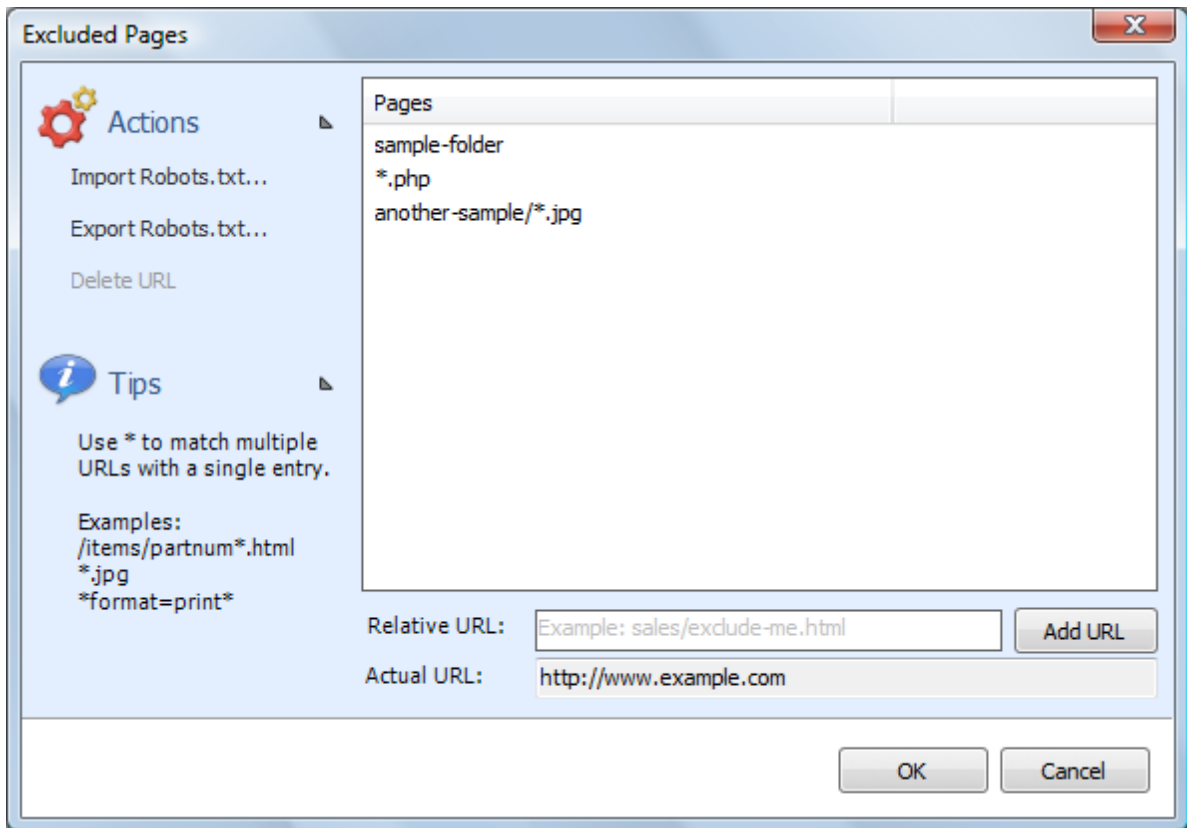
When a directory (like "/blog") is specified as an Excluded Page, then all of its sub-directories will be excluded (such as: www.example.com/blog/index.html and www.example.com/blog/archives/january2005.html).

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 the Relative Path field.

*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 '?').

> **Advanced Tip:** The * matches every character including "/". If you want to match any text but **NOT** the path separator ("/"), use an exclamation point in front of the asterisk: "!*". This special case will match any characters **EXCEPT** the path separator.

The "$" can be used to match the end of the URL, like this:

inventory/index.php$	will exclude the following:
	inventory/index.php
	but will NOT exclude:
	inventory/index.php?product=xyz

Internal URLs should be relative to the root URL. URLs beginning with '/' are taken from the domain root (this is recommended). URLs **not** beginning with '/' are relative to your project's Root URL. 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'

URLs beginning with 'http://' are treated as full URLs, and can be used to exclude links to external sites. For example:

'http://www.anothersite.com' **remains** *'http://www.anothersite.com'*

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

4.5.1 Importing Robots.txt

Some websites already contain a special file called "robots.txt" which tells crawling software which URLs (or paths) to ignore. OrFind does not automatically obey robots.txt because it is often desirable to crawl the entire website. However, OrFind has the ability to import a robots.txt file into the Excluded Pages list. To access this feature, click **"Import Robots.txt..."** under **"Actions"** on the Excluded Pages window.
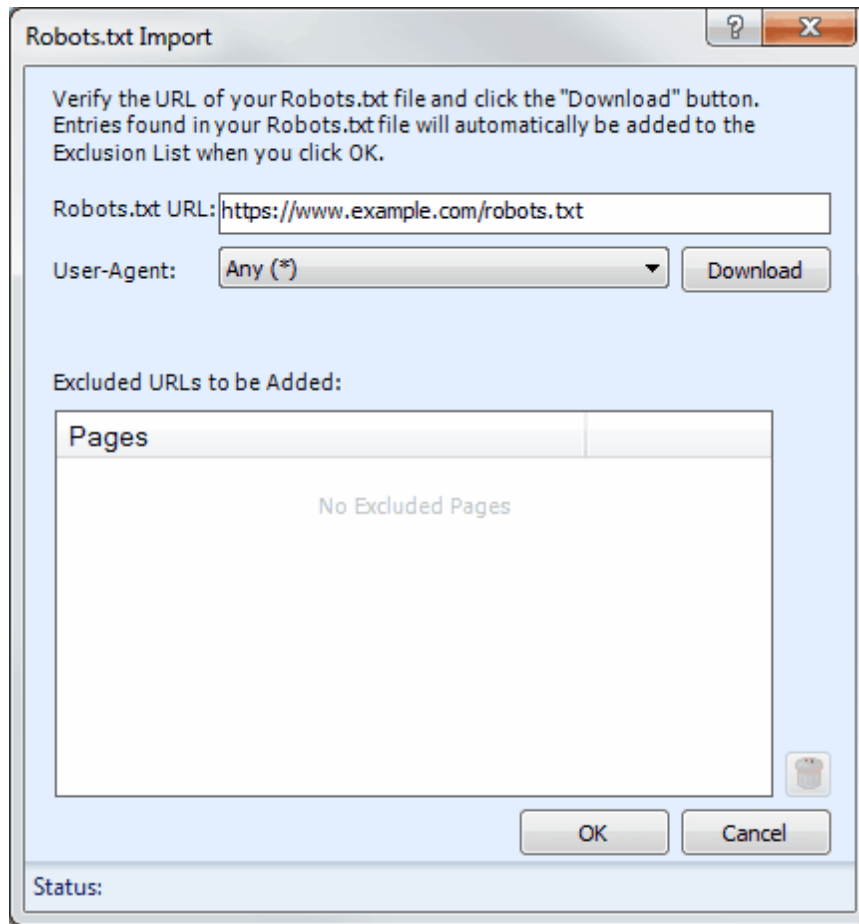
When the Robots.txt Import window opens, it will automatically populate the "Robots.txt URL" field based on your Root URL. You can modify this field if necessary. Click "Download" to fetch and load the file from the website.

The User-Agent field is used to select which rules you wish to import from the robots.txt file. By default the "*" user-agent will be loaded (any bot).

Once OrFind loads the file successfully, the results will be placed in the "Excluded URLs to be Added" list. You can remove any unwanted entries now by highlighting the entry and clicking the "Trash" button (or pressing "Del" on your keyboard). 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.

4.5.2 Exporting Robots.txt

In addition to being able to import an existing robots.txt file, OrFind is also capable of creating (or exporting) it based on the current entries in the Excluded Pages list.

The contents of the Excluded Pages list will be converted into "Allow" and "Disallow" rules and normalized automatically.

The robots.txt file generated by OrFind must be placed at your domain root. For example, if you Root URL was "www.example.com/products/", the robots.txt file created by OrFind must be placed so that it's accessible from "www.example.com/robots.txt" (not in the "products" sub-folder).

The robots.txt file created by OrFind will only have entries under the "*" user-agent field. If you require different rules for different user-agents, you will need to manually edit your robots.txt file in a text editor (such as Notepad).

4.6 Passwords and Forms

If your website is password protected or contains content hidden behind a form submission, you can configure OrFind to access that content using the "Passwords and Forms" feature. To access the Passwords & Forms window, click **"Project | Passwords & Forms"** from the main menu bar.

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

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

The "Method" field tells OrFind what type of credentials to supply to the remote server. There are two methods for supplying login information to a website, POST and HTTP.

Post Method

This is the most common method of logging into a website. If you enter your login information by

filling out some type of web based form, then use this method. To configure the login credentials click the '**Form Wizard...**' link under **Actions** to launch the Inspyder mini-browser. Use this browser to log into your website. When you complete the login form, OrFind will ask if you want to save the form data, click 'Yes' and OrFind will automatically create the necessary login rule for you.

Prompt for Manual Login

If you don't want OrFind to store your login credentials, you can enable the '**Prompt for Manual Login**' option. This option only works with websites using cookies for authentication (most websites using form/post based authentication also use cookies). When you click the 'Go' button to start crawling your site OrFind will launch the mini-browser to capture your login session on-the-fly. When you close the mini-browser OrFind will resume crawling with your logged in session.

HTTP Method

The second 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 (the domain field is optional). Click 'Add' to save the rule.

To remove a rule, select the rule to remove from the list, then click the **"Delete Page"** link under **"Actions"**.

5   Crawl Results

OrFind's results are displayed in the "Crawl Results" section below the project settings. The Crawl Results contains 3 tabs: Orphan Analysis, Changed Files, and Log.

Orphan Analysis

The orphan analysis tab is broken down into two hierarchical "tree" displays showing orphaned and accessible files respectively.

Orphaned Files

The contents of this area provides a tree view of the orphaned file hierarchy. Each folder in the tree lists the size of all the files and folders contained within. This may include files used by server-side scripts and should be reviewed for false-positive orphans. OrFind is also capable of following content included in style sheets and in style sections of HTML pages.

Any false-positive orphans should be excluded by using the Excluded Pages feature. You can right click on any orphan item to either **Open the URL** in your default browser, add that page or the entire folder to the **Excluded Pages** list or to specify access information for **Protected Pages**.

The check boxes beside each file/folder are used to indicate which files should be removed when the "**Delete**" button is clicked. See Removing Orphans for more information.

Accessible Files

The contents of this area provides a list of the accessible files. It is identical to the orphan files section, but does not have check boxes and these files cannot be deleted through OrFind.

HTML References

This tab provides a list of all file references that were detected in the HTML content of your website. These results are useful if you are trying to determine which pages link to a specific file on your website.

Server-Side References

This tab is similar to the HTML References tab, but provides all the references that were detected in server-side code only (such as through a PHP or ASP "include" directive). If server-side analysis is disabled in Advanced Project Settings, this tab will be hidden.

Changed Files

The contents of this tab provides a list of files that were modified, added or removed, since the last successful crawl of the site. The display is comprised of the filename and the modification type. The first time OrFind is run on a website all files will appear as "Added".

Log

Information about the crawling, backup and deleting processes is here. These results are not included in a report (since they are primarily used for troubleshooting), but right clicking on the results will allow you to save the contents to a text file.

To clear this information in the Crawl Results section, click the "**Clear Results**" button in the lower right hand corner of the OrFind window.

5.1   Removing Orphans

After orphan analysis is complete OrFind can remove the orphaned files from your server or hard drive. To do so, check one or more file items in the Orphans file tree and click the **"Delete"** button. A confirmation window will appear asking you to verify the files that are to be deleted. On this screen you can also specify a backup location for backing up the orphans prior to deleting them.
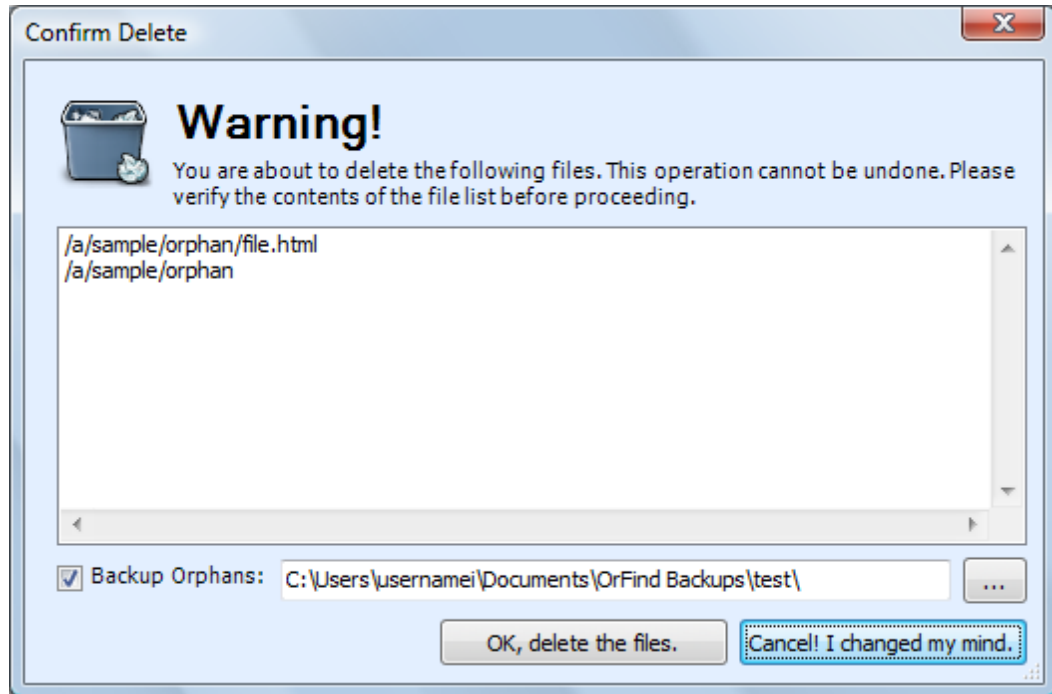
If a file cannot be backed up (due to an FTP error or any other reason) OrFind will not delete the file.

6   Reporting

The contents of the Crawl Results can be exported from OrFind in PDF, Excel or comma delimited text format (CSV). To generate a report of the current results, click the "Report" button from the main toolbar or select "Project | Generate Report..." from the main menu bar. The Report Options window is used to specify which information from the Crawl Results are to be included in the report as well as the data format and filename.
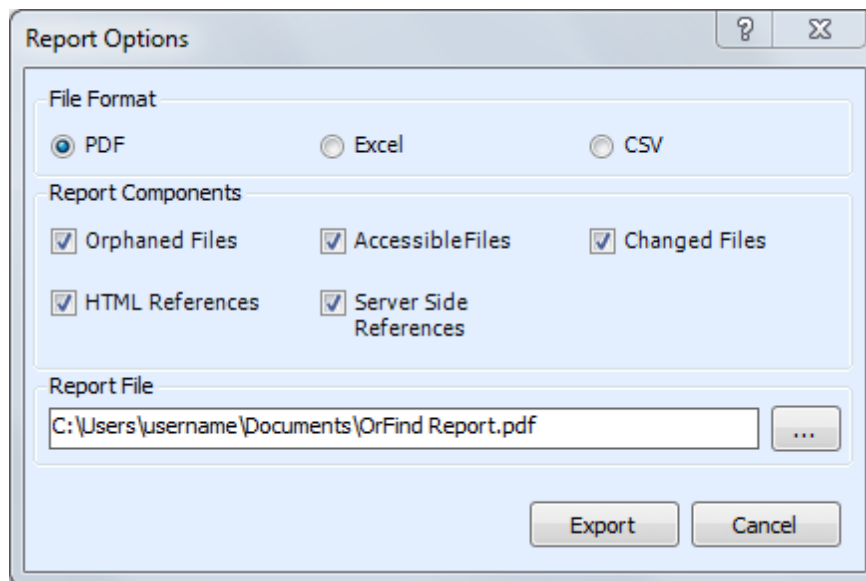
File Format

The File Format option allows you to select the report file type you would like OrFind to create. PDF and Excel formats contain formatting elements, logo image, etc. The CSV format simply contains text that can be opened up in a spreadsheet application for further analysis if desired.

The logo used in the PDF/Excel reports can be customized by replacing the logo image file located in: C:\Program Files\Inspyder Software Inc\Inspyder OrFind\Reports\
An image resolution of 300 DPI or higher is recommended for printing.

Report Components

The Report Components check boxes allow you to select which data should be included in the report. Any component that is unchecked will not be shown in the report. It is not possible to include the **Log** section in the report.

Report File

Use the Browser (...) button to choose an output filename and location.

6.1   Report Designer

OrFind includes a Report Designer tool that allows you to make **limited** customizations to the look and feel of the PDF/Excel reports. To access the report designer, click "Settings | Report Designer" from the main menu bar.
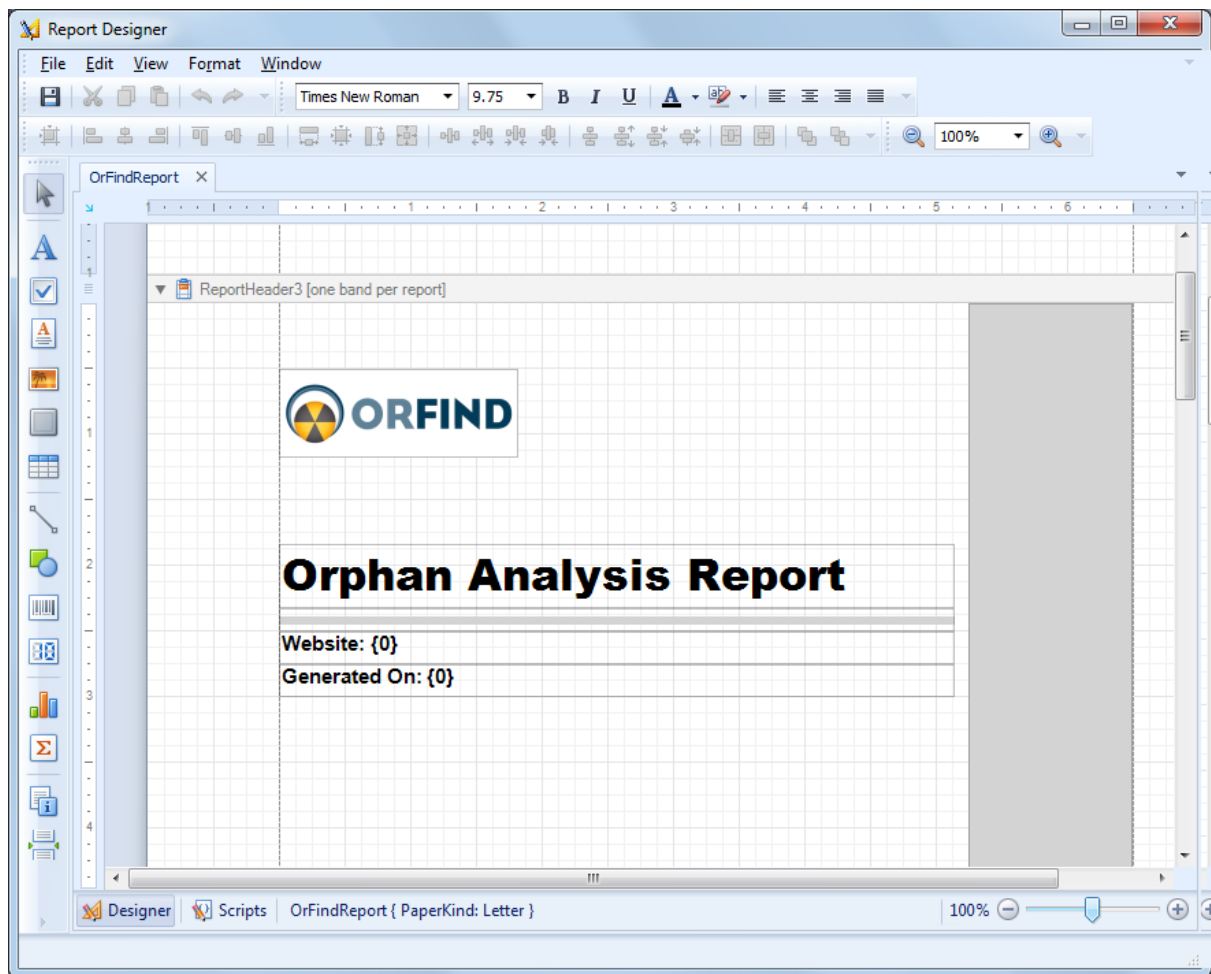
The main purpose of the report designer is to facilitate translation of the report text and provide basic branding. Changes made using the Report Designer affect all reports created with OrFind. The report customizations are **not** project specific.

To begin modifying a report, select the element on the main window. To edit the element text, double click the element and enter new text as necessary.

When you have finished making changes to the report template, click the "Save" button to save your changes. Any new reports created by OrFind will include the changes made in the report designer.

Note: It is possible to delete existing report elements, but it is not possible to restore these deleted elements without reverting all report changes.

To remove all the customizations done in the report designer, select "**Settings | Report | Reset Changes**" from the main menu bar.

7   Defaut Project Settings

Default Project settings specify the initial configuration used for any newly created projects. Options that are unique to every Project (such as Root URL) cannot be defined as a default setting, however common options (such as Crawler Timeout) can be configured here.

Changes made to the Default Project Settings will also effect **current projects** that use the **same default values**. For example, if you have 10 projects already created in OrFind and you have not changed the "Crawler Delay" setting in any of them (so they are still using the original "0 seconds" value), then those projects are still using the "Crawler Delay" value from the Default Project. Now, if you change the crawler delay in the Default Project from 0 to 1 second, that change will affect all 10 of your current projects.

If you had another project where you had explicitly changed the Crawler Delay, for example, from 0 to 5 seconds, then any change made to the Crawler Delay in the Default Project Settings would not affect this project. This project had the "Crawler Delay" changed within it, so the Default Project Setting value is ignored.

To edit the Default Project Settings click '**Settings | Default Project Settings...**' from the main menu bar. The settings in the Default Project Settings window are nearly identical to those found in **Advanced Project Settings**. For details on what each specific options means, please review the settings in the 'Advanced Project Settings' section of the manual.

8　Project Import and Export

OrFind 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.
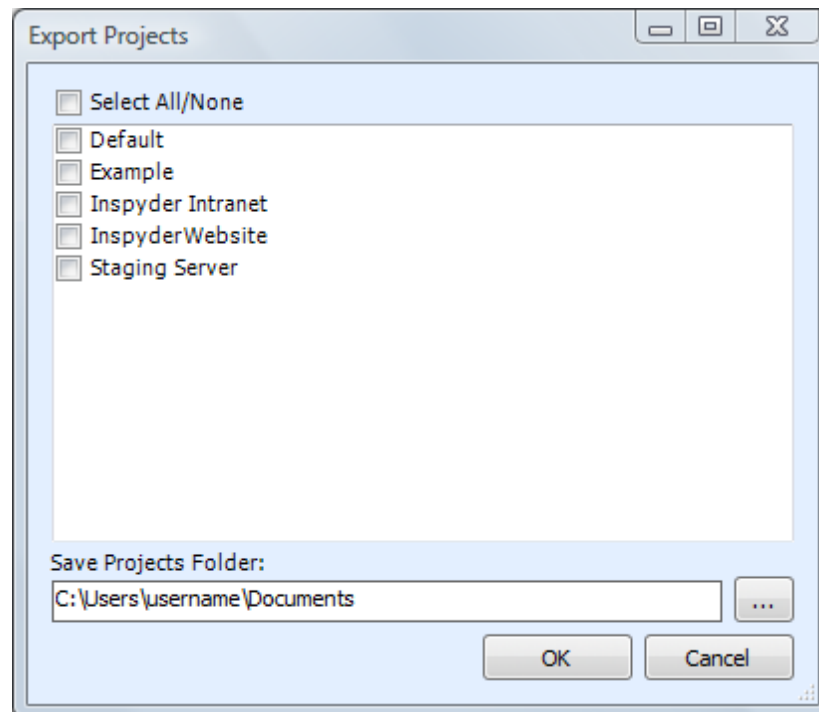
8.1　Exporting Project Files

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".projx

"Project Name" is the name of the project in OrFind. "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 ".projx" file extension.

Exporting projects from OrFind 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 OrFind.

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'.

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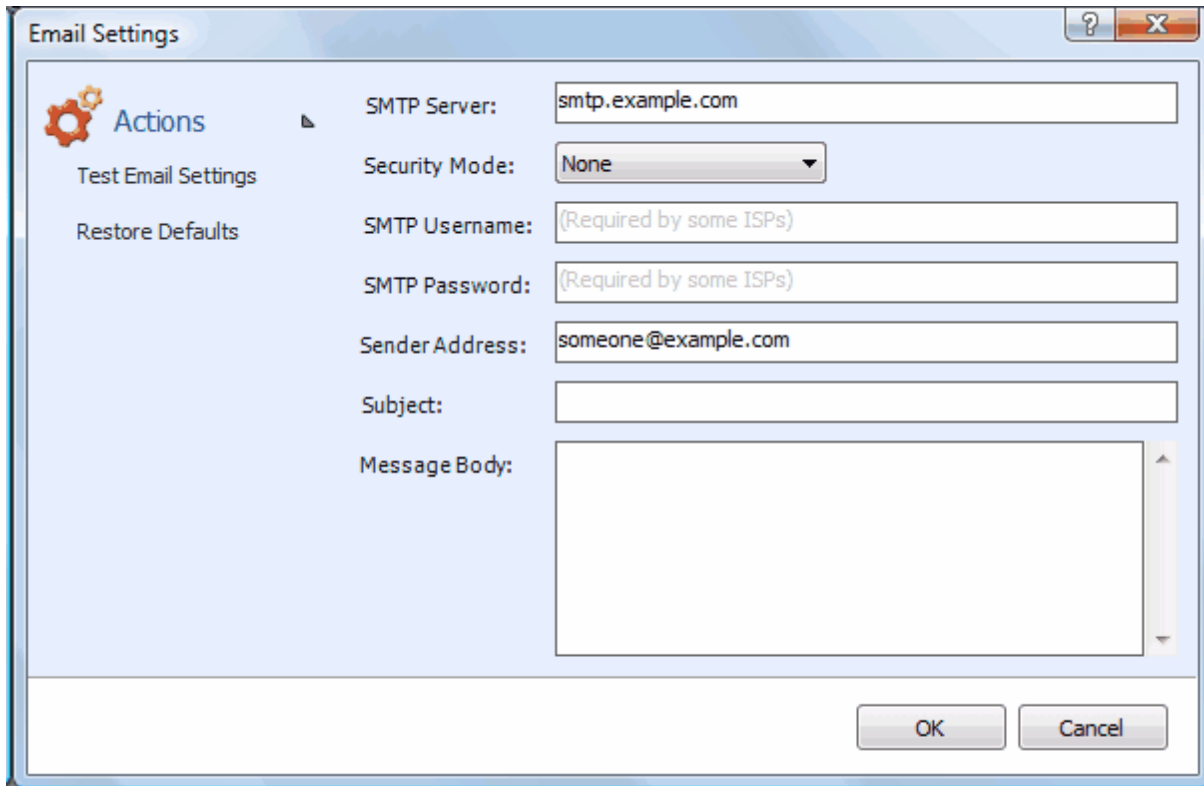
To Import one or more project files into OrFind, select "**File | Import Project Files...**" from the main menu bar. An Open File window will appear that will enable you to select one or more Inspyder project files.

It is possible to import a project file from an Inspyder application other than OrFind. 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 OrFind project back into OrFind, then all settings will be used.

9   Email Settings

9.1   Server Settings

Configuring your email server settings (SMTP) allows OrFind to send automated emails. To access your mail server settings, click "**Settings | Email Settings...**" from the main menu bar.

Settings

SMTP Server - This is the hostname of your outgoing (SMTP) mail server. If you are unsure what to use, check with your Internet service provider (ISP) or your email provider (such as Google if you use Gmail).

If your provider uses a non-standard port (25 by default, or 465 if SSL is enabled) you can append it to the hostname by separating it with a colon (for example: "mail.example.com:1234").

Security Mode - If your provider uses SSL (Implicit SSL) or TLS (Explicit SSL) select that option here. SSL mode uses port 465 by default. "None" and TLS modes us port 25.

SMTP Username and Password - This is your SMTP username and password. Use these fields if your service provider requires SMTP authentication. If no username and password are supplied OrFind will not attempt to use SMTP authentication.

Sender Address - This is the email address that OrFind will use as the "From:" address for any emails.

Subject and Body - See the Message Body section in the manual for details.

Actions

Test Email Settings - This action tests your configuration by sending an email to the email address specified in the **Sender Address** field. If you receive this message then your email settings are correct. If an error occurs or you do not receive the email, verify your settings and try again.

Restore Defaults - This action resets the **Subject** and **Message Body** fields to their default values.

9.2 Message Body

The OrFind email subject line and message body can be customized in the Subject and Body fields respectively. By using special tags inside the message body, it can be dynamically customized based on the current project settings and the results of the crawl.

The following tags can be used:

#Orphans#   The number of orphaned files detected
#Accessible# The number of accessible files detected
#Changes#   The number of files that have changed (added, removed or updated) since the last crawl
#Date#   The date the report was generated
#Time#   The time the report was generated
#URL#   The Root URL used
#Project#   The name of the Project

Keep in mind that the report with the full crawl results listing can be included with the email as an attachment.

10   Scheduler

OrFind has the ability to run automatically in an unattended mode. To configure OrFind for scheduled operation simply select "**Settings | Scheduler**" from the main menu bar.
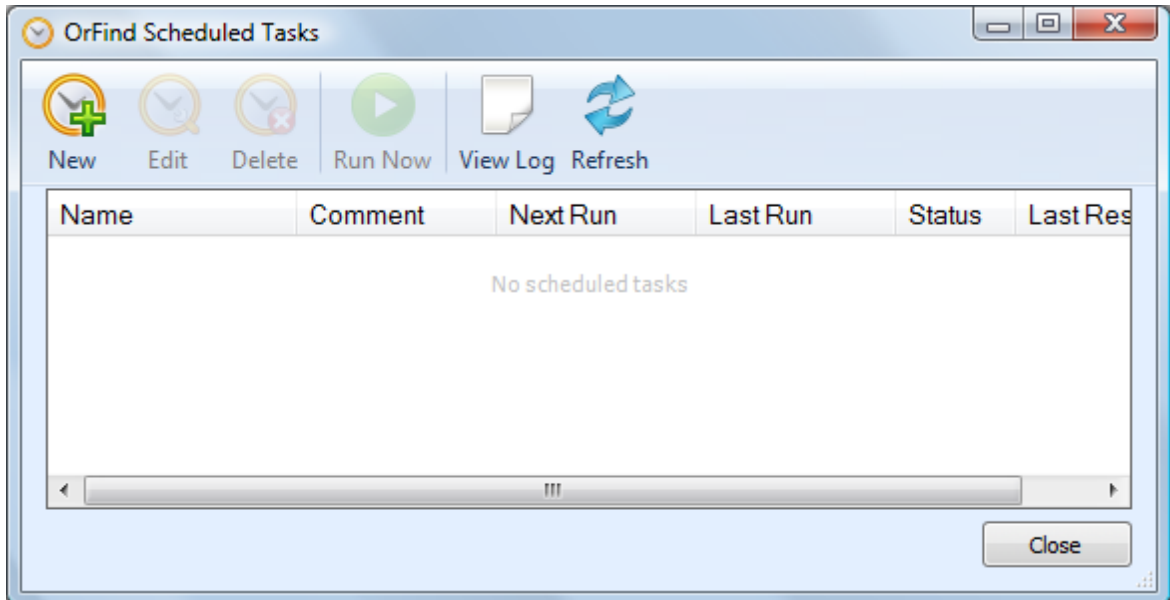
10.1   Scheduling Basics

Scheduling OrFind to run automatically is done by creating one or more tasks in the OrFind Task Scheduler. Each scheduled task contains the settings that will be used when running OrFind automatically. Basic settings include:
• The OrFind Project to load before crawling
• The date and time the task will run next
• How often the task should be re-run

Tasks are run by the Windows Task Scheduler, so OrFind does not need to be open for the scheduled tasks to run, but your computer must be turned on (it will not wakeup automatically).

Tasks configured within OrFind are visible in the Windows Task scheduler. For scheduling options more advanced than those provided by OrFind, please see the Windows Task Scheduler documentation included with Windows.

Name	Comment	Next Run	Last Run	Status	Last Result
Sample Task	Nightly Crawl	1/4/2010 12:00:00 PM	N/A	Ready	1

Any scheduled tasks will appear in the scheduler window. The list of scheduled tasks shows the

following information:

Name - The user-friendly name of the scheduled task

Comment - Any comments that were entered when this task was created

Next Run - The next time the task will automatically run

Last Run - The last time the task ran automatically

Status - Indicates if the Task is "Ready" (ready to run), "Running" (is currently running) or "Disabled" (will never run)

Last Result - The result code of the last run. If this number if not zero than an error may have occurred (the Scheduler Log may have more details on any errors that occurred).

10.2   Adding/Editing a Task

To add a new task to the OrFind Task Scheduler, click the "**Add**" button. The scheduling wizard will launch, which will walk you through the steps to creating your new task. When you are finished, click "Close" to save your task to the Windows Task Scheduler. The new task will appear in the list of scheduled tasks.

You can edit a task at any time by highlighting the task item in the list and clicking the "**Edit**" button. This will launch the same scheduling wizard as before with your corresponding settings already populated. Please note, if you have made changes to your scheduled tasks outside of OrFind then those changes will be lost when you edit your task with the scheduling wizard.

It is possible to force a task to run immediately by selecting the Task item, then clicking the "**Run**" button.

Tasks can also be set to "**Disabled**" by right clicking and then un-checking the "**Task Active**" item. When a task is disabled its settings are still saved by the scheduler, but it will not run automatically.

10.3   Deleting a Task

To remove a task from the scheduler, select the task item from the list of tasks and click the '**Delete**' button. This will permanently remove the task from the Windows Task Scheduler, but will not remove any projects, reports or other data associated with that task.

10.4   Scheduler Log

The scheduler log provides a mechanism to review what happened while a scheduled Task was running. To view the log click the '**View Log**' button.
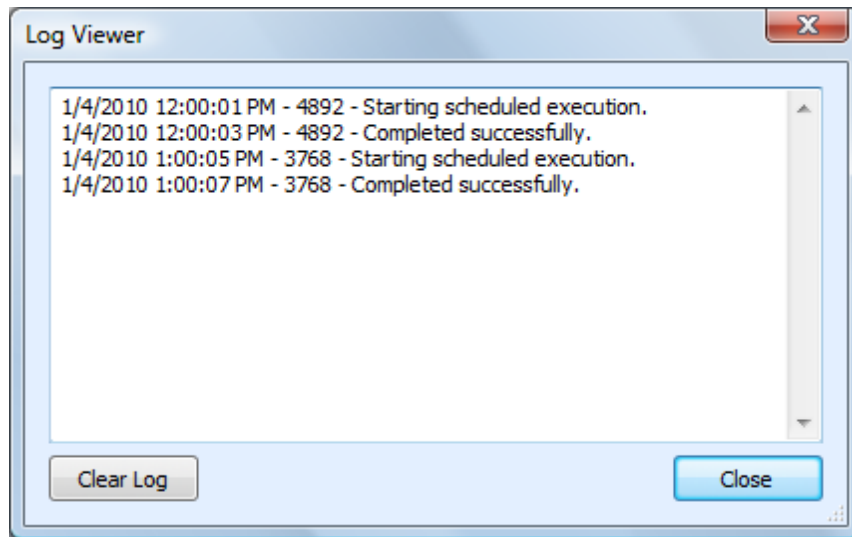
Each line in the log file provides the following information:
- Date/Time the event occurred
- Process ID of the running Task (this is useful for tracing events if you have multiple scheduled tasks running at the same time)
- The log message

To remove all entries from the log, click the "**Clear Log**" button.

10.5   Command Line Interface

If you wish to run OrFind from the Windows command line interface, the following parameters can be used. These parameters are also used by the Task Scheduler when scheduling tasks. If you want to manually adjust your scheduled tasks in the Windows Task Schedule refer to these parameters.

The following are the command line arguments:

-orphans	Selects Orphaned Files content from Scan Results to be included in output
-accessible	Selects Accessible Files content from Scan Results to be included in output
-changes	Selects Changed Files Content from Scan Results to be included in output
-ifchanged	This option indicates that the report is only to be generated if file changes are detected.
-output=	Specified the output file for the results (i.e. C:\MyFile.html)
-outtype=	Specifies output format; CSV, PDF or Excel (i.e. -outtype=PDF)
-project=	The project name to be analyzed (i.e. "-project=Inspyder Local")
-email=	Specifies that an email notification should be sent to the following address when this scheduled task completes. (i.e. "-email=someone@example.com")
-attachreport	Indicates that the report file should be attached to the email

To satisfy the Task Scheduler's parsing algorithm please ensure that each argument is placed within double-quotes (i.e. "-output=foo.html"). This is usually only required if the argument contains spaces.
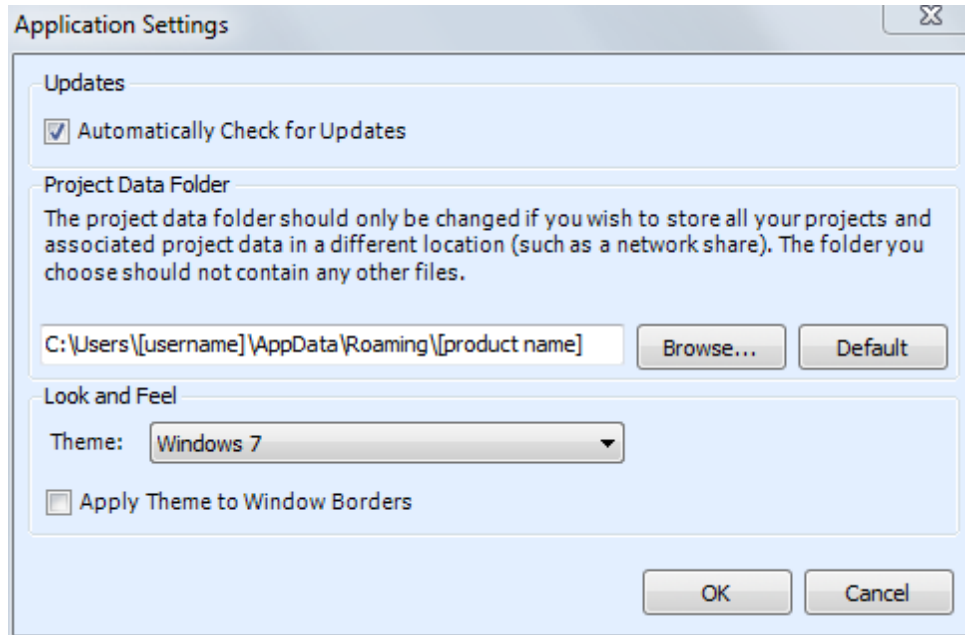
10.6   Result Codes

These are the codes returned to the Task Scheduler after the OrFind 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   Report Generation Error
       Solution: Try running the report in interactive mode, and check for errors. The report file may have been in use or an incorrect output path may have been selected.
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   Emailing Error
       Solution: Check to make sure your SMTP settings are correct in Email Settings and try again.

11 Application Settings

The Application Settings window allows you to change application wide settings that do not affect individual projects. You access this window from the main menu, under "Settings | Application Settings...".

Updates - If this option is enabled, OrFind will automatically check for a new version each time it runs. Unselect this option if you do not wish to see the "A new version is available" message when you start OrFind.

Project Data Folder - This setting allows you to change where OrFind stores it's project and data files. This is useful if you wish to change the default location to a shared network location, or to a folder that is easier for you to backup. Note that you must restart OrFind for any changes to take effect. If you change the Project Data Folder your existing projects will not be moved (you must move them manually). Finally, the folder you select as the Project Data Folder should not be used for anything else. It should only contain projects and files created by OrFind.

Look and Feel - These options allow you change the theme of OrFind. "Apply Theme to Window Borders" allows you to over-ride the Windows border with a style that matches the current theme.

12 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.

13   License Agreement

This legal document is an agreement between you, as licensee and Inspyder Software Inc. ("Inspyder"), as licenser. 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), DVD(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 license.

GRANT OF LICENSE: The Licenser grants to You a non-transferable non-exclusive license to use the Product on a single computer. You may physically transfer the Product to another computer provided the Product is used only on a single computer. Installing the Product in a shared environment such as a Terminal Server or otherwise making it accessible to more than the single computer, even if only one user is using the Program at any given time, violates this license agreement. 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 Licenser'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 OrFind to provide an exclusive service such as using OrFind 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 Licenser's entire liability and Your exclusive remedy shall be: (a) the replacement of any media not meeting Licenser's 'Limited Warranty' and which is returned to Licenser 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
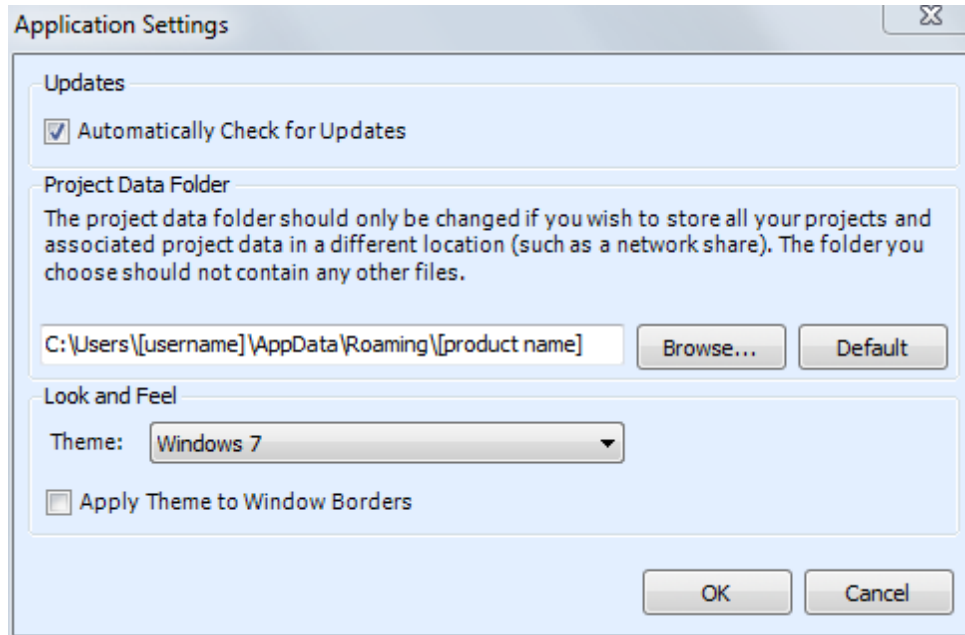
Product and Your money will be refunded. In no event will Licenser nor anyone involved in the creation, production or distribution of the Product be liable to You for any direct, indirect, consequential or incidental damages including any lost profits, lost savings, lost business revenue or other commercial or economic loss arising out of the use or inability to use the Product even if Licenser or any authorized representative of Licenser 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 © 2015 by Inspyder Software Inc.

14   Application Settings

The Application Settings window allows you to change application wide settings that do not affect individual projects. You access this window from the main menu, under "Settings | Application Settings...".

Updates - If this option is enabled, OrFind will automatically check for a new version each time it runs. Unselect this option if you do not wish to see the "A new version is available" message when you start OrFind.

Project Data Folder - This setting allows you to change where OrFind stores it's project and data files. This is useful if you wish to change the default location to a shared network location, or to a folder that is easier for you to backup. Note that you must restart OrFind for any changes to take effect. If you change the Project Data Folder your existing projects will not be moved (you must move them manually). Finally, the folder you select as the Project Data Folder should not be used for anything else. It should only contain projects and files created by OrFind.

Look and Feel - These options allow you change the theme of OrFind. "Apply Theme to Window Borders" allows you to over-ride the Windows border with a style that matches the current theme.

Inspyder Software Inc.
200-3342 Mainway,
Burlington, Ontario,
Canada, L7M 1A7