|  |
| --- |
|  |
| SharePoint Broken Link Manager™ |
| User Manual |
|  |
|  |
| **11/29/2017** |

|  |
| --- |
|  |

This document will provide users with instructions on how to use QiPoint’s SharePoint Broken Link Manager™ tool.

Table of Contents

[Introduction 5](#_Toc499825215)

[Installation & Minimum Requirements 6](#_Toc499825216)

[IMPORTANT: This program does not need to be installed on a SharePoint Server. 6](#_Toc499825217)

[Minimum Hardware Recommendations 6](#_Toc499825218)

[Supported SharePoint Versions 6](#_Toc499825219)

[.NET Framework Required 6](#_Toc499825220)

[Local Machine & Windows System Permissions Required 8](#_Toc499825221)

[Anti-Virus, Offline Sync Folders (i.e. Google Drive, OneDrive, Drop Box, etc) and Performance 8](#_Toc499825222)

[SharePoint User Permissions Required 9](#_Toc499825223)

[Find/Replace Permissions Required 9](#_Toc499825224)

[Preserve Author 9](#_Toc499825225)

[Office 365 Accounts 9](#_Toc499825226)

[Configuration 10](#_Toc499825227)

[Managed Paths 10](#_Toc499825228)

[Getting the Managed Paths 10](#_Toc499825229)

[Errors if incorrect Managed Paths are specified 10](#_Toc499825230)

[Entering the Managed Paths 11](#_Toc499825231)

[Broken Link Manager Default Options 12](#_Toc499825232)

[What will be scanned 19](#_Toc499825233)

[Outbound Links 19](#_Toc499825234)

[Supported Link Types 19](#_Toc499825235)

[SharePoint Web pages 19](#_Toc499825236)

[JS Links / Display Template URLs 19](#_Toc499825237)

[Navigation Links 20](#_Toc499825238)

[Document ID Links 20](#_Toc499825239)

[Web Parts 20](#_Toc499825240)

[SharePoint Lists & libraries 20](#_Toc499825241)

[InfoPath Forms 20](#_Toc499825242)

[File Contents 21](#_Toc499825243)

[Linked CSS & JavaScript files 22](#_Toc499825244)

[What will NOT be scanned 23](#_Toc499825245)

[Links 23](#_Toc499825246)

[Lists & libraries 23](#_Toc499825247)

[File Contents 23](#_Toc499825248)

[My Sites / Personal Sites 24](#_Toc499825249)

[Navigation Links 24](#_Toc499825250)

[Performing a Scan for Broken Links 25](#_Toc499825251)

[Start a Scan 26](#_Toc499825252)

[Scanning Multiple Sites at Once 28](#_Toc499825253)

[Scan Options 28](#_Toc499825254)

[Job Name 28](#_Toc499825255)

[Site URL to Scan 28](#_Toc499825256)

[Credentials 28](#_Toc499825257)

[Scan Options 29](#_Toc499825258)

[File Extensions 29](#_Toc499825259)

[Inclusions Rules (Optional) 31](#_Toc499825260)

[Exclusion Rules (Optional) 33](#_Toc499825261)

[NOTES on Inclusion/Exclusion Rules: 33](#_Toc499825262)

[Find / Replace (Optional) 34](#_Toc499825263)

[Email (Optional) 36](#_Toc499825264)

[Export Options (Optional) 37](#_Toc499825265)

[What is a Broken Link? 40](#_Toc499825266)

[A broken link is a link that doesn't work, often resulting in an error page. A broken link happens when the link points to a web page that has been deleted or moved. 40](#_Toc499825267)

[Manually Finding and Replacing Links 40](#_Toc499825268)

[Automatically Finding and Replacing Links 41](#_Toc499825269)

[What will be replaced 41](#_Toc499825270)

[Supported Link Types 43](#_Toc499825271)

[Find/Replace Order 43](#_Toc499825272)

[What will NOT be replaced 43](#_Toc499825273)

[Versioning 44](#_Toc499825274)

[NOTE 44](#_Toc499825275)

[Minor (Drafts) and Major (Published Versions) 44](#_Toc499825276)

[Check In & Check Out 45](#_Toc499825277)

[Preserve Author 45](#_Toc499825278)

[Content Approval & Find / Replace 45](#_Toc499825279)

[Bulk Import Find Replace Links 45](#_Toc499825280)

[Find / Replace Options 46](#_Toc499825281)

[Link Report - Finding URLs that were automatically replaced/fixed 48](#_Toc499825282)

[How to find all Links that matched Find & Replace but were not replaced? 49](#_Toc499825283)

[Scheduling Jobs 50](#_Toc499825284)

[Schedule a Job 50](#_Toc499825285)

[List of Scheduled Jobs 52](#_Toc499825286)

[Scheduled Job Logs and Reports 53](#_Toc499825287)

[Batch Processing Multiple Site Collections 54](#_Toc499825288)

[Licensing 55](#_Toc499825289)

[Components 55](#_Toc499825290)

[Changing the license type of a component 55](#_Toc499825291)

[Trial Version 55](#_Toc499825292)

[Free Version 55](#_Toc499825293)

[Obtaining a License 55](#_Toc499825294)

[Transferring a License 56](#_Toc499825295)

[Logs 57](#_Toc499825296)

[How to Interpret the Report 59](#_Toc499825297)

[Example 59](#_Toc499825298)

[Which Links Are Broken? 59](#_Toc499825299)

[Reports 61](#_Toc499825300)

[Opening the Report 61](#_Toc499825301)

[Report Fields 61](#_Toc499825302)

[Report Archive 63](#_Toc499825303)

[Grid View Report 64](#_Toc499825304)

[Customizing the Report 64](#_Toc499825305)

[Filtering 64](#_Toc499825306)

[Advanced Filtering 64](#_Toc499825307)

[Searching within the Report 65](#_Toc499825308)

[Filtering Queries 65](#_Toc499825309)

[Grouping 66](#_Toc499825310)

[Views 67](#_Toc499825311)

[Export 68](#_Toc499825312)

[Best Practices: Improving performance and Large Scan Jobs 70](#_Toc499825313)

[Tips, Comments & Troubleshooting 72](#_Toc499825314)

[Improving time to complete Large Scan Jobs 75](#_Toc499825315)

[Technical Support 79](#_Toc499825316)

[Product Features 80](#_Toc499825317)

# Introduction

This utility is part of the SharePoint Essentials Toolkit TM Suite. This utility is used to manage and help report on broken links found in SharePoint sites.

# Installation & Minimum Requirements

Please see the “SharePoint Essentials Toolkit – User Guide” for installation instructions.

## IMPORTANT: This program does not need to be installed on a SharePoint Server.

## Minimum Hardware Recommendations

**Processor:** Minimum Dual-Core, 3GHz. Recommended Quad-Core 3GHz or higher  
We recommend Quad-Core 3GHz processor or higher when scheduling more than 20 Jobs at the same time.

**RAM:** Minimum 1GB available. Recommended 2GB or higher available

We recommend:

* At least 1GB of available RAM for jobs\* with less than 200GB and less than 10K items to scan
* At least 2GB of available RAM for jobs\* with 200GB to 500GB or with 10-100K items to scan
* At least 4GB of available RAM for jobs\* with 500GB+ or with 100K items or more to scan.

\*Jobs – includes all jobs being run at one time, such as multiple jobs in a schedule for one or more sites

**Hard Disk:** 200MB Available Hard Disk Space (for the application files, logs, temp files and reports).   
Temp files are automatically cleared as needed. This is in addition to the disk space required by the SharePoint Essentials Toolkit.

## Supported SharePoint Versions

Microsoft SharePoint Online / Office 365  
Microsoft SharePoint Server 2016  
Microsoft SharePoint Server 2013  
Microsoft SharePoint Server 2010

## .NET Framework Required

.NET Framework 4.5 or higher is required. If it is not found on the client machine, the user will be prompted to automatically download and install the prerequisite.



## Local Machine & Windows System Permissions Required

Please see the SharePoint Essentials Toolkit User Manual.

## Anti-Virus, Offline Sync Folders (i.e. Google Drive, OneDrive, Drop Box, etc) and Performance

Please see the SharePoint Essentials Toolkit User Manual.

# SharePoint User Permissions Required

Users require specific Read permissions to be able to use the tool, see below for the specific permission levels required:

NOTE: The Visitor SharePoint Group and the READ Permission Level have these permissions already included by default.  
Also, providing the user with a Web Application User Policy of ‘Full Read’ will automatically grant the user all of these read permissions required to use this tool.

**View Items** - View items in lists and documents in document libraries.

**Open Items** - View the source of documents with server-side file handlers.

**View Versions** - View past versions of a list item or document.

**View Application Pages** - View forms, views, and application pages. Enumerate lists.

**Browse Directories** - Enumerate files and folders in a Web site using SharePoint Designer and Web DAV interfaces.

**View Pages** - View pages in a Web site.

**Use Remote Interfaces** - Use SOAP, Web DAV, the Client Object Model or SharePoint Designer interfaces to access the Web site.

**Open -** Allows users to open a Web site, list, or folder in order to access items inside that container.

**NOTE: T**he above permission levels are required for each site you will be running a report for. If the report is being run for a subsite, the account running the tool does not require permission on the root site to build a Broken Link report.

## Find/Replace Permissions Required

The user must have edit rights on the files/items being modified.

### Preserve Author

If you need to “Preserve Author” information such as “Last Modified By”, “Last Modified Date”, the user running the job must have Full Control over the Site/List/List Item being updated to preserve this information.

## Office 365 Accounts

* When scanning Office 365 SharePoint sites, an Organizational account must be used, such as [user@mycompany.com](mailto:user@mycompany.com) or [user@mycompany.onmicrosoft.com](mailto:user@mycompany.onmicrosoft.com). Microsoft accounts (Windows Live IDs) such as [user@live.com](mailto:user@live.com) or [user@hotmail.com](mailto:user@hotmail.com), are not currently supported for authentication.

# Configuration

## Managed Paths

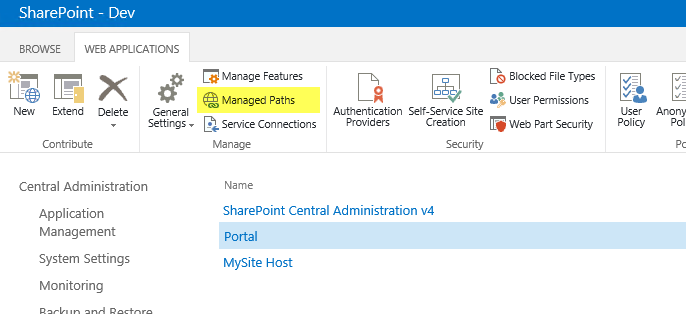
SharePoint uses managed paths to help organize you SharePoint Sites. By default, if you do not have any custom Managed Paths, you do not need to change anything, we have added the default Managed Paths for SharePoint already, such as “/sites/, “/personal/”.

### Getting the Managed Paths

If you are unsure of what the Managed Paths are for your environment, please contact your SharePoint Administrator.

#### Central Administration

You can obtain a list of all Managed Paths from Central Administration under “Application Management” -> “Manage Web Applications” and clicking the Managed Paths button in the ribbon:



#### SharePoint Management Shell (PowerShell)

You can use the command below to retrieve a list of managed paths. Run this command on any SharePoint Server:

Get-SPManagedPath -WebApplication <http://webapplication>

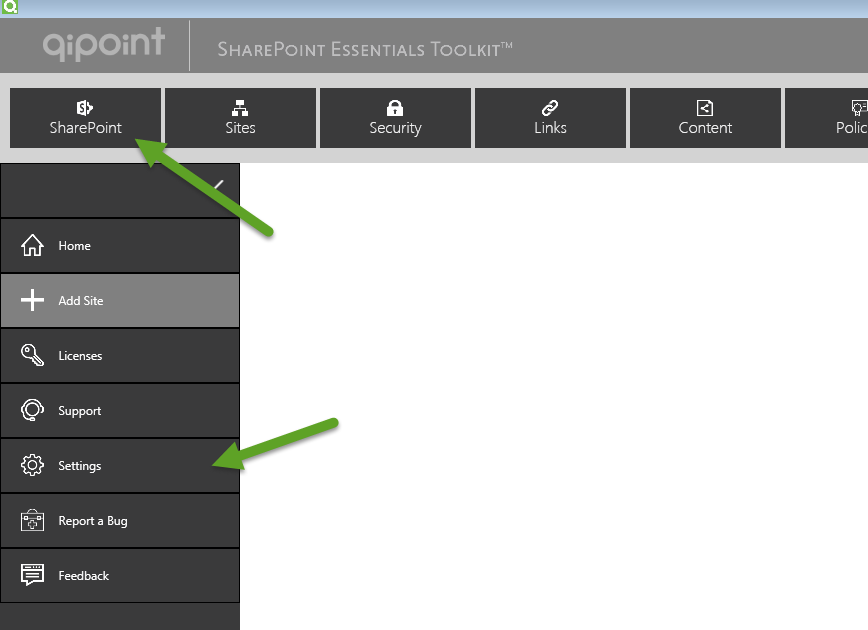
### Errors if incorrect Managed Paths are specified

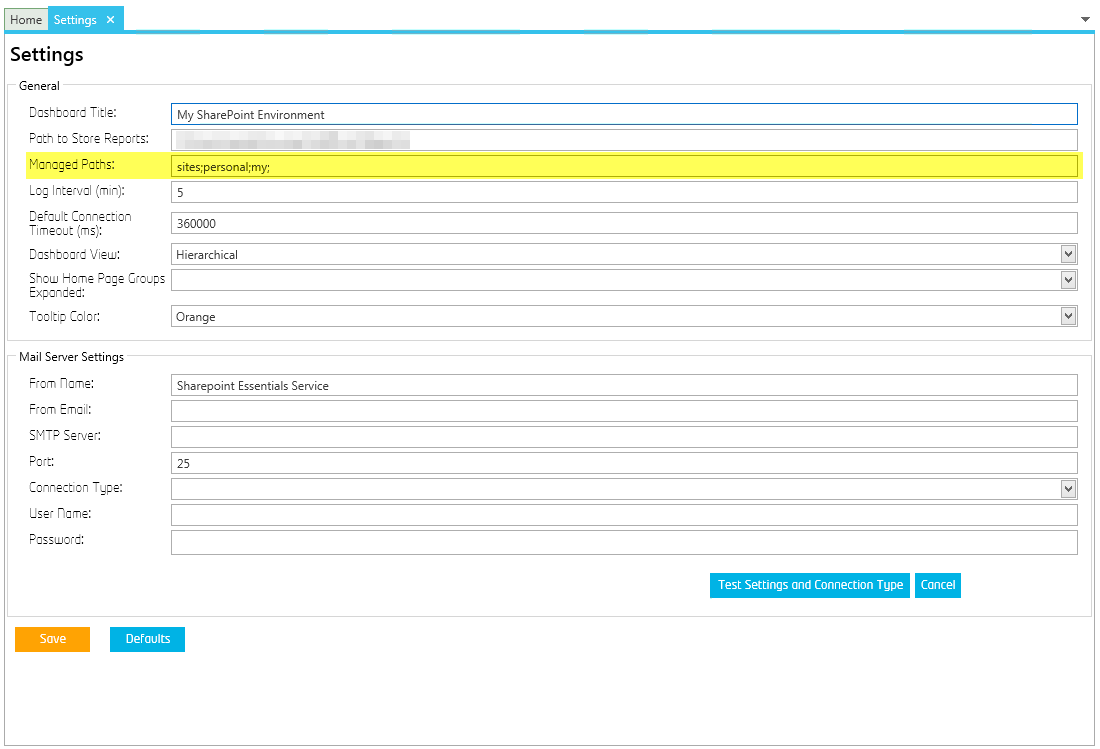
If you do not have the correct SharePoint Managed Paths listed here, if using the SharePoint Broken Link Manager tool, you may get paths with duplicate entries and false negatives in the reports. For example, if you have a Managed Path called ‘Projects’, the utility will not be able to understand the paths correctly when relative and will appear as <http://contoso/projects/projects/tasks>, and giving a false report of that URL being broken.

Currently only the SharePoint Broken Link Manager component relies on this setting to build reports correctly.

### Entering the Managed Paths

Open the SharePoint Essentials Toolkit “Settings” page (see below) to enter all Managed Paths in your SharePoint environment.



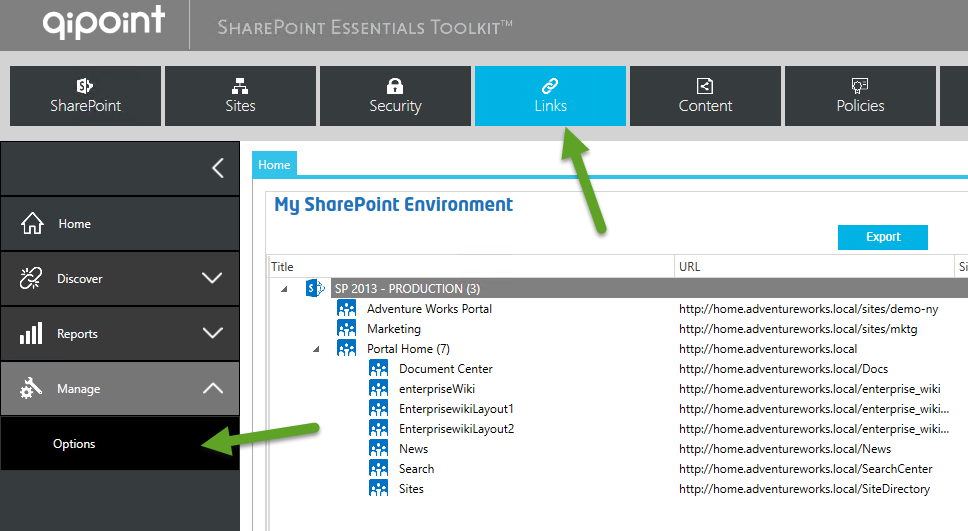


NOTE: The tool does not attempt to get your environment’s Managed Paths automatically because the tool could be run as an account with Read permissions, and therefore may not have access to Central Administration to get this information.

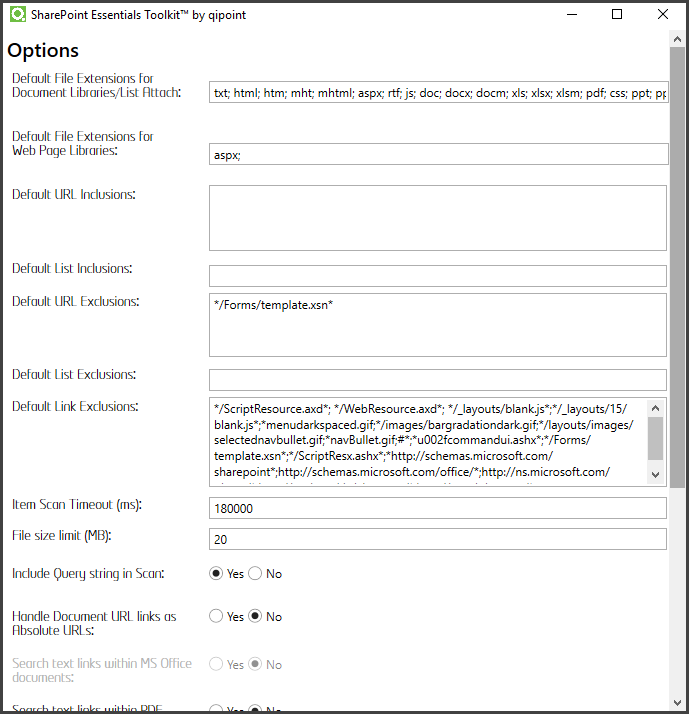
## Broken Link Manager Default Options

You can change the default options for the Broken Link Manager component by following the instructions below. These options will be set as the default values used in all reports that are run.

1. From the Top Navigation, click on the “Links” button
2. From the Left Navigation, click to expand “Manage”, then click the “Options” button



1. The following “Options” window will appear:



1. Complete the form or click OK if the settings are correct. The next section will provide descriptions of the fields.

Default File Extensions for Document Libraries/List Attachments

(Only used in Enterprise Version) These are the default file extensions for the files to include in scanning file contents. The file extensions can then be overwritten in the actual scan job settings prior to a scan. This will determine which files’ contents will be scanned when checking Document Library and List Attachment for broken links.

Default File Extensions for Web Page Libraries

(Only used in Enterprise Version) These are the default file extensions for the files to include in scanning lists based on the Page library template. The file extensions can then be overwritten in the actual scan job settings prior to a scan. This will determine which files’ contents will be scanned when checking Page Libraries (such as ‘Pages’ or ‘SitePages’ Libraries) for broken links.

Default URL Inclusions

(Only used in Enterprise Version) These are the default URLs to include when performing a scan. The URL inclusions can then be overwritten in the actual scan job settings prior to a scan. This applies to the URL of the SharePoint object such as the site, list or SharePoint Item/SPListItem.

* Separate multiple URLs using semi-colons ‘;’
* Only one match needs to be true for the rule to succeed
* URLs are case insensitive
* Wildcards (\*) are supported, for example: ‘http://portal/sites/test1/\*’ will only include URLs that start with ‘http://portal/sites/test1/’. ‘\*/sites/\*’ will only include URLs that have ‘/sites/’ in the URL
* Exclusion rules will be applied at this point
* Once a match is made, all the links within the item matched (such as a file or web page) will be scanned
* A log entry will show the item as being included or ‘matched’

Default List Inclusions

(Only used in Enterprise Version) These are the default Lists to include when performing a scan. Enter the List titles (use display name not internal name).

* Separate multiple list names using semi-colons ‘;’
* Only one match needs to be true for the rule to succeed
* The List inclusions can then be overwritten in the actual scan job settings prior to a scan
* List names are case insensitive.
* Wildcards (\*) are supported, for example: ‘\*Migrated\*’ will only scan lists that have the word ‘Migrated’ in the title, all other lists will be skipped
* Once a match is made, all the items within the list will be scanned (pending other inclusion URL rules)
* Inclusion rules do not apply to links within the page/item/file contents that are matched
* Exclusion rules will be applied at this point
* A log entry will show the item as being included or ‘matched’

Default URL Exclusions

(Only used in Enterprise Version) These are the default URLs to exclude when performing a scan. The URL exclusions can then be overwritten in the actual scan job settings prior to a scan. This applies to the URL of the SharePoint object such as the Site, List or SharePoint Item/SPListItem URL.

* Separate multiple URLs using semi-colons ‘;’
* Only one match needs to be true for the rule to succeed
* URLs are case insensitive
* Wildcards (\*) are supported, for example: ‘\*/test/\*’ will exclude all URLs that contain ‘/test/’ from being scanned. If the exclusion rule does not match the URL of the page/item/file, all links within the matched item will be scanned
* Exclusion rules do not apply to links within the pages/item contents
* Exclusion rules are applied after Inclusion rules
* A log entry will show the object as being skipped

Default List Exclusions

(Only used in Enterprise Version) These are the default Lists to exclude when performing a scan. Enter the List titles (use display name not internal name). The List exclusions can then be overwritten in the actual scan job settings prior to a scan.

* Separate multiple list names using semi-colons ‘;’
* Only one match needs to be true for the rule to succeed
* List name exclusions are case insensitive.
* Wildcards (\*) are supported, for example: ‘\*test\*’ will exclude all lists that have the word ‘test’ in the title, all other lists will be included
* Exclusion rules are applied after Inclusion rules
* A log entry will show the list as being skipped

Default Link Exclusions

(Only used in Enterprise Version) These are the default URL links to exclude when performing a scan. This applies to the file contents only (URL links found within file contents) and not the actual SharePoint Item URL. Enter the URLs to exclude. The Link exclusions can then be overwritten in the actual scan job settings prior to a scan.

* Separate multiple link exclusions using semi-colons ‘;’
* Only one match needs to be true for the rule to succeed
* URLs are case insensitive
* Wildcards (\*) are supported, for example: ‘\*/scripts.aspx’ will exclude all URL links that have ‘/scripts.aspx’ in it, it will be excluded/skipped from the scan.
* Exclusion rules are applied after Inclusion rules
* A log entry will show the item as being skipped

Item Scan Timeout (ms)

This is the duration until the application will stop trying to process a SharePoint Item, Page or Document, such as when parsing an item and waiting to connect to URL addresses found within that item. Normally you should not have to change this value unless you see frequent timeouts in the logs.

File size limit (MB)

(Only used in Enterprise Version) This is the maximum file size of an item to process when scanning links. Increasing this amount will impact the time it takes to parse large files. The timeout setting above still applies. By default, this is set to 20MB and can be increased or decreased as required.

Include Query String in Scan

Select yes to include the query string in scans (default). For example, a hyperlink with <http://someplace.com/default.aspx?productId=123> will be scanned including the query string (“?productId=123”). If ‘No’ is selected here, the URL that would be scanned will be <http://someplace.com/default.aspx> (which excludes the query string). Normally you would include the query string in scans unless they are causing issues when scanning. NOTE: Some web pages require a query string in its URL to return a valid page result. You may want to turn this off in specific link scans when including the query string produces undesired results.

By default, this option is on.

Search text links within MS Office documents

(Only used in Enterprise Version) Select “Yes” to include non-hyperlink URLs in the scan. Non-hyperlink URLs are links within MS Office documents that do not have a ‘link’ when the mouse is hovered over them. If this is set to ‘No’, some links that you see within documents will not be scanned by the tool because they are treated as ‘Text’. Set this option to ‘Yes’ to search for all URLs within files. This may increase the time the job takes to complete as entire document contents are parsed for links vs retrieving only ‘hyperlinks’ in a file.

By default, this option is off.

Search text links within PDF documents

(Only used in Enterprise Version) Select “Yes” to include URLs in the scan that were created using 3rd party tools (when exporting or converting the PDF file). For example, some documents are created using a non-Adobe application and the links are converted using a PDF export tool, the hyperlinks remain in-tact but are not treated as Links from Adobe API (Application Programming Interface). Set this option to ‘Yes’ to search for all URLs within PDF files. This may increase the time the job takes to complete as entire document contents are parsed for links vs retrieving only ‘hyperlinks’ in a file returned by the Adobe API.

By default, this option is off.

Scan linked CSS and JS files

(Only used in Enterprise Version) Select “Yes” to include linked CSS (Cascading Style Sheets) and JS (JavaScript) files that are associated with web pages.

For example, a web page with a tag:

<link rel="stylesheet" type="text/css" href="/\_layouts/1033/styles/Themable/search.css?rev=T%2Bhraxktc1A8EnaG5gGMHQ%3D%3D"/>

When this page is scanned, the CSS file located at “/\_layouts/1033/styles/Themable/search.css” will also be scanned for broken links (such as broken image links). The same concept for JavaScript files in <script /> tags.

This will likely significantly increase the time the job takes to complete as all linked document contents are parsed for these type of links and could result in the same CSS or JS files being scanned multiple times.

This option is useful if you need to see the impact a CSS broken image link or JS inserted link is causing across pages and sites.

By default, this option is off.

Scan list form pages

Turning this option on will perform a scan on list Form pages, such as AllItems.aspx and NewForm.aspx. Sometimes these pages may have been customized and have links that you want to scan within them.

By default, this option is off.

Scan MS Excel Formulas

(Only used in Enterprise Version) If this option is turned on, the tool will retrieve all MS Excel cells that have formulas with links and then report on those links, such as to another workbook in SharePoint to perform a calculation. This will increase the time to scan each MS Excel file.

By default, this option is off.

Debug Mode

Use this option only when troubleshooting invalid responses. If this option is set to Yes, the tool will attempt to include more details in the HTTP Responses in the Active Logs (and exported CSV logs). This will significantly use more system resources and is recommended not to turn this on unless troubleshooting. If this option is set to No, the tool will provide standard details in the Active Logs for HTTP Responses.

By default, this option is off.

# What will be scanned

This tool will scan the following for hyperlinks, as well as perform automatic Find/Replace (Enterprise Edition only):

## Outbound Links

* + Outbound Links to external sites
  + Outbound links to (internal or external) SharePoint artifacts such as files, web pages, list items
  + Outbound links to (internal or external) non-SharePoint content such as web site pages, files within your network, web applications

## Supported Link Types

* + Http, Https
  + FTP
  + File shares and UNC paths
  + Tooltips that contain hyperlinks
  + Descriptions or Alt URLs within hyperlinks

## SharePoint Web pages

* + Includes all links on a web page, including all content areas, web parts, and navigation areas
  + Page libraries and also Form (aspx) pages (Allitems.aspx, All Forms.aspx) within SharePoint Lists and Libraries are scanned for links
  + HTML formatted links matching: (not case sensitive)

‘href’, ‘src’, ‘’source’, ‘background’, ‘title’, ‘\_tooltip’, ‘alt’

Examples include:

<a href=”http://www.somelink.com”>anchor display text</a>

<a href=”http://www.somelink.com” title=”http://www.somelink.com” ></a>

<image source=”http://www.somelink.com” alt=”http://www.somelink.com”></image>

<table background=” http://www.somelink.com” />

* + All non HTML links that are embodied within text content (Example plain text: “You can find the document here at http://somelink.com” (which is not underlined as hyperlink or has a link on mouse over)) will be picked up and scanned. This is an optional setting in the Link Options page.

## JS Links / Display Template URLs

The tool will check web part JS Link sections for URLs. They will also be replaced if Find & Replace is used.

## Navigation Links

The tool will check the Top Navigation / Global Navigation and Quick Launch / Current Navigation for links. They will also be replaced if Find & Replace is used. Both Managed Navigation (using Managed Metadata term set) and Structural Navigation Links are supported. However, Managed Navigation (using Managed Metadata term set) does not work with Find & Replace.

## Document ID Links

The tool will automatically resolve Document ID links using the SharePoint Document ID web service of that Site. The Document ID Service will either REDIRECT a user to the file if a match is found, however, if the link is ‘broken’, the Document ID service shows a message such as “No documents with the ID HCH2YEEPCJYU-5-45 were found in this site collection.” The tool will recognize this and correctly show a 404 ‘Not Found’ in the Broken Link Report. Please ensure you are using the latest version of our tool as this was a new feature as of SET version 4.4.3.0.

## Web Parts

The following web parts will be scanned for broken links:

* Script Editor Web Parts
* Xslt List View Web Parts
* Image Viewer Web Parts
* Page Viewer Web Parts
* Summary Links Web Parts
* Content Editor Web Parts

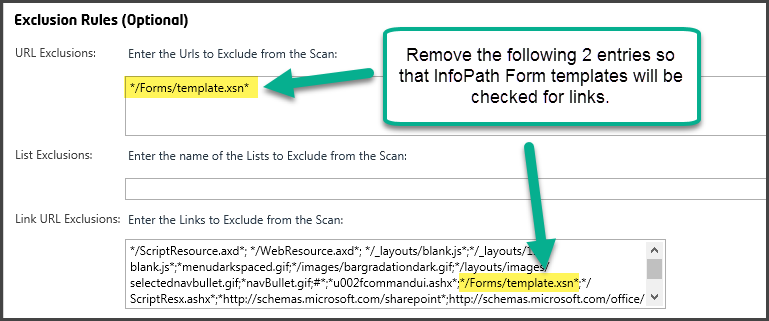
NOTE: Content Editor web parts with linked/embedded files are scanned separately if the file is located in the SharePoint site (in a list or library). Content Editor web parts with links typed directly into the web part area (‘Edit Web Part’) are scanned for links.

## SharePoint Lists & libraries

* + Includes all SharePoint List Definition Types except Invalid Lists and External List types.
  + Includes all column meta-data that may contain links (such as Text, Multi-line Text, Hyper-Link data types).

## InfoPath Forms

* + InfoPath Form fields with URLs will be scanned for links and can be replaced (Enterprise Edition only).
  + InfoPath Form Template data connection reference links are supported, they can also be replaced using our tool (Enterprise Edition only).
  + If you upgraded from SharePoint Essentials Toolkit version 4.5.6.0 or older, in order to fix InfoPath URLs and Data Connection Links, you need to remove the following settings:



## File Contents

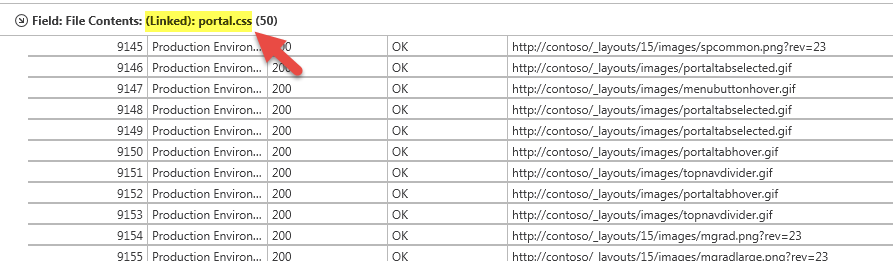
* + Scans Document Library file contents and List Item Attachment file contents
  + Scans Text friendly formats (ASCII) such as xml, xslt, mht, rtf, txt, js, css, html, aspx, php, etc. (Tip: If it is text readable in Notepad, it typically can be scanned).
  + NOTE: CSS Files – when scanning partial/relative paths within CSS files, please note that they are relative to the CSS file directory, not the document. For more information on this, see [http://www.w3.org/TR/REC-CSS1/#URL](http://www.w3.org/TR/REC-CSS1/#url).
  + Scans MS Office Word, Excel, PowerPoint, and Project documents.
  + Microsoft Office ‘97 versions and higher.
  + Scans Microsoft Word for Macintosh 98 to X, including other .doc applications such as OpenOffice and AbiWord.
  + The tool supports MS Office documents that have images that have hyperlinks behind them, these types of links behind images are supported by the tool.
  + MS Excel Formulas can be scanned by the tool, there is an option to enable this feature. If enabled, the formula where the URL is found, will display the Cell Value in the Link Description field of the report in order to help you find the Cell in question (that has the link). The Field Type will be identified as “DocumentFileContentsExcelFormula” so that you can sort, filter or group by this column.
  + MS InfoPath Forms and URLs within the template.xsn files are supported. You may need to remove the default template.xsn exclusions rule to include these in the job. See above “InfoPath Forms” for more details
  + Scans PDF documents (except for PDF files that have been scanned as images).
  + MS Office files that have Embedded Objects (OLE) that are hyperlinks are scanned. However, the embedded objects themselves are not scanned for links.
  + MS Office images with hyperlinks behind them are supported
  + PDF Text Links (non-hyperlinks) - There is an option in the Scan Links (tab) -> Options -> “Include text links within PDF documents”. Setting this option to ON will increase resource utilization and time to run a scan.
    - * This option will parse all text within PDF files.
      * Text Hyperlinks are URLs that do not appear as links in the document – they are not ‘clickable’

Such as <http://somelink> (real link) vs http://somelink (text link)

* + - * These are different than ‘real’ ‘Hyperlink URLs’ (these show a hyperlink tooltip when mouse is hovered over them) which are retrieved programmatically via PDF API (programming interface).
      * Default is OFF

## Linked CSS & JavaScript files

* + Optionally, you may scan linked CSS and JavaScript files for broken links (which are referenced within a web page). You can turn this feature ON/OFF by going to Scan Links (tab) -> Options -> "Scan linked CSS and JS files"
  + If this feature is turned ON, when scanning pages for broken links, if a page has a CSS (Cascading Style Sheet) or JavaScript file referenced (such as in a <script /> <link /> reference tag), its contents will also be scanned for potential broken links, such as broken image links.
  + These types of links can be identified by viewing the column “Field” in the report. Linked content that is scanned will be in the format “File Contents: (Linked) /PublishingImages/companyStyle.css”
  + You can also filter or group these in your report by right clicking on the column “Field” and selecting “Filter Editor”, then add a filter to only include/exclude ‘linked’ file contents, this can help when resolving broken links within CSS & JavaScript files and reviewing the impact they may have on your web pages. You can also group by the Field column to view all links for a ‘linked css’ file. See below:



# What will NOT be scanned

This tool will not scan:

## Links

* + Links that end with ‘**#**’ are skipped
  + Links within ‘**javascript**:’ in HTML tags are skipped, however the JavaScript source file links are checked
  + Links containing “\_vti\_bin/spsdisco.aspx” are skipped
  + Links to email addresses, ‘**mailto:**user@domain.com’ are shown as valid, it will not check the email address or domain
  + Links with ‘appredirect.aspx’ are skipped
  + Links with ‘\_layouts/userdisp.aspx’ are skipped
  + Any links that you add to the “URL Exclusions” list will be skipped
  + MS Word bookmarks are not treated as URL links and are skipped
  + Data Connection links within MS Office documents are not supported
  + Links listed in the “Link URL Exclusions” settings are skipped (see Scan Links (tab) -> Options -> “Default Link Exclusions” to set up the default URL exclusions
  + Inbound links coming from external websites are not scanned. The tool will have no way of knowing what sites on the internet or other networks or servers have links to your site. Only Outbound links will be scanned that are found in your SharePoint environment, this includes all links to external sites or links internal to your environment.

## Lists & libraries

* + Does NOT include invalid list types
  + External Lists are not supported
  + Calendar Overlays links are not supported
  + Copy Source fields are not supported by Find & Replace but they will be reported

## File Contents

NOTE: Below items relate to scanning inside of file contents, metadata scanning is working for these types.

* + We cannot list all possible file types out there, but the following file types (that we know of) are not currently supported by our tool to scan for links or find replace:
    - One Note 2016 - limited support for One Note2013 and older
    - Adobe Illustrator (.ai) – not supported
    - Adobe PhotoShop (.psd) – not supported
    - AutoCAD (.dwg, .dxf) – not supported
    - Corel WordPerfect (.wpd) – not supported
    - AutoDesk (.3ds, .prj) – not supported
  + MS Word 97 text based links (links that do not have a hyperlink but just text, such as (1) <http://contoso> vs (2) http://contoso, the (2) is a text based URL that is not ‘clickable’). Actual hyperlinks in MS Word 97 are supported.
  + MS Office OLE / Embedded Objects contents are not scanned for links.
  + Image file contents are not scanned for links.

## My Sites / Personal Sites

* + Links to My Sites and Personal sites are supported, however scanning links within personal site web pages is not fully supported, User Profile metadata properties are not supported. SharePoint Lists & Document Libraries within a My Site are supported.

## Navigation Links

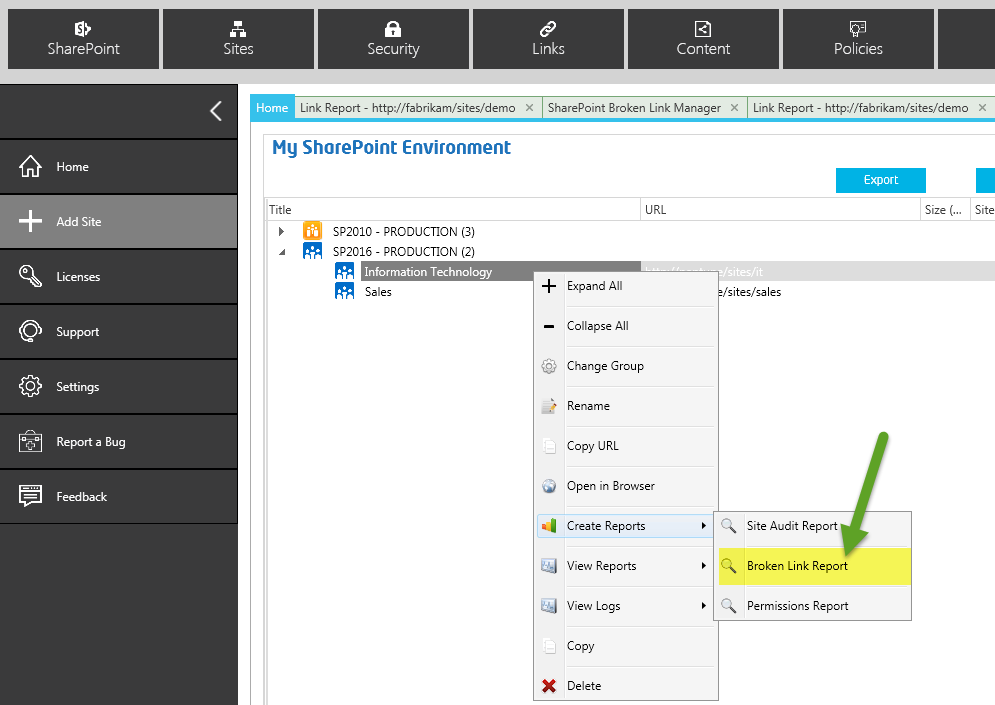
Managed Navigation (using Managed Metadata term set) will show in the reports, however, Find & Replace will not work on links in the navigation that are using Managed Metadata.

# Performing a Scan for Broken Links

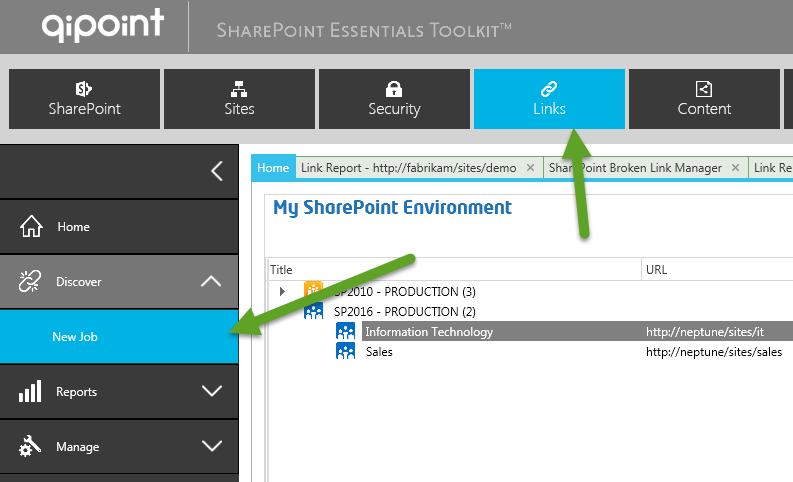


There are 2 ways you can perform a scan for broken links.

1. From the Home Page Dashboard, right click on a site (that you want to build a report for) and click “Create Reports” -> “Broken Link Report”



1. Click on the “Links” button in the Top Navigation, click on the “Discover” button in the Left Navigation to expand it, then click “New Job”



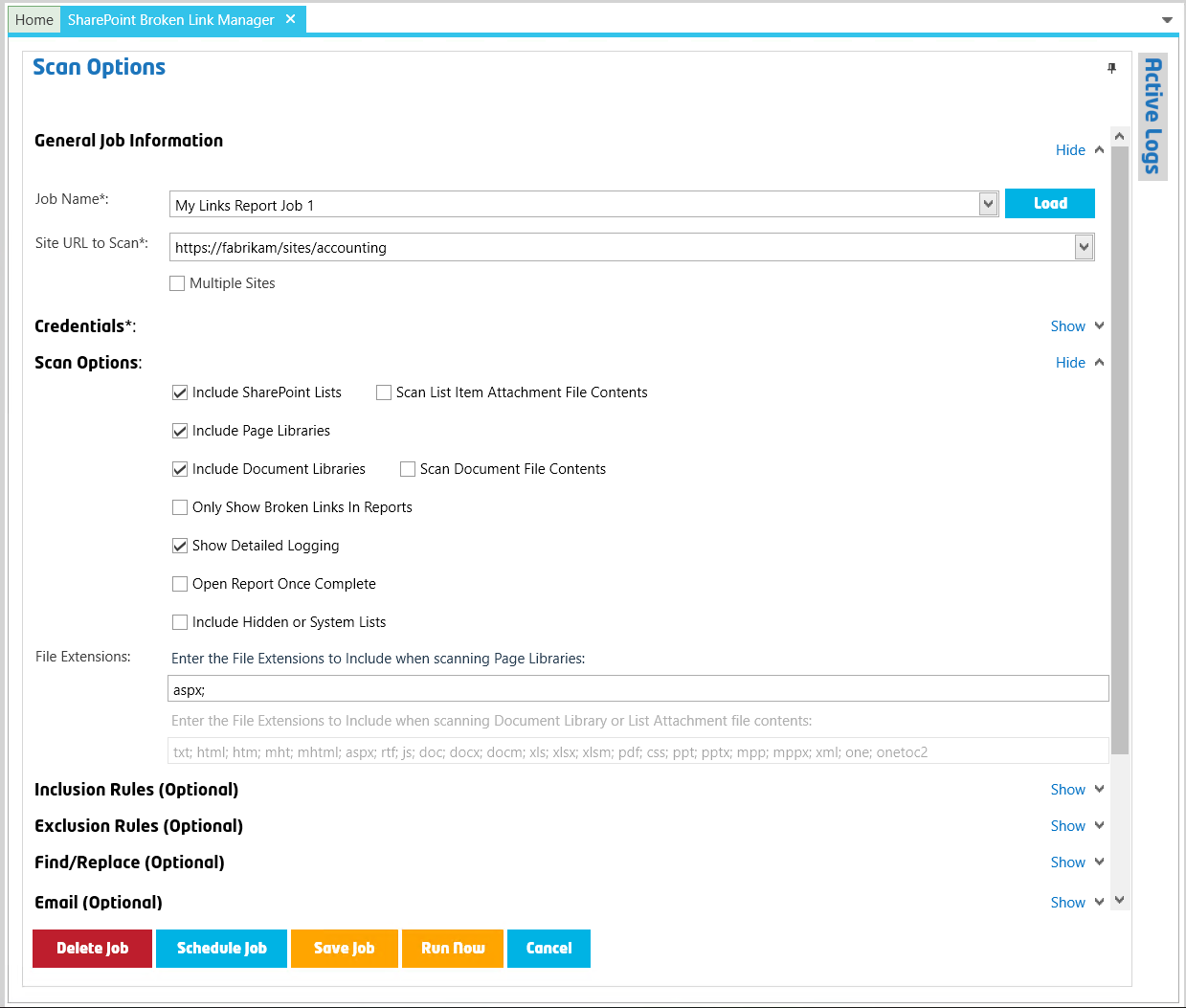
## Start a Scan

1. Open “SharePoint Essentials ToolkitTM”
2. Use one of the options above to open the Broken Link Manager “New Job” window
3. A “New Job” window will appear called “SharePoint Broken Link Manager” (below screenshot)

(NOTE: This page below may appear different depending on the version you are using)

1. Complete the form. Fields with an asterisk \* indicate required fields

See “Scan Options” (next section) for a description of each option



1. Click “Run Now” button to start the job immediately

## Scanning Multiple Sites at Once

To create jobs for multiple sites and subsites, see the “SharePoint Essentials Toolkit User Guide” under the section “Batch Processing Jobs for Multiple Sites”.

## Scan Options

### Job Name

This is the Name of the job. A folder will be created in the report directory for every job. You can use the same Job Name to group reports. The job name will be used as the report name and will be used to help identify the report if scheduled. You must rename the job if you want to save the job for future re-use.

### Site URL to Scan

This is the absolute URL of the site you want to scan for broken links. This can be the path to a Site Collection top level (root) site, or a sub site.

**Example:** <https://companyportal/sites/hr> or <https://companyportal/sites/hr/projects/capital_site/>

Do not include the page path in the URL

Multiple Sites  
Select this option to run a ‘multi-threaded job’. If this option is selected, you will NOT be able to run a manual job using ‘Run Now’, you need to schedule the job as it will be broken down into one job per site. These jobs will run in parallel and can be viewed from the Scheduled Jobs section. Once complete, job reports can be accessed from the Scheduled Jobs page, the Job History page and also from the Home Page Dashboard by right clicking on one or more of the sites.

### Credentials

Select the authentication type and enter the credentials used to access this site.

**Default SharePoint Authentication** – Automatically determines Authentication to connect to the SharePoint site. This will work in most cases when SharePoint is hosted on your internal company network. Enter the custom credentials to use for this site. If your environment is using a custom login screen, such as one that requires Multi-Factor Authentication/a PIN, is using ADFS, or has SSO enabled, use Web Based Authentication below.

**Office 365** – Use this option to connect to Office 365 hosted SharePoint sites. This applies to both Office 365 non-federated environments. If using a Federated/ADFS enabled Office 365 environment or if your Office 365 environment is using a custom login screen, such as one that requires a PIN, use Web Based Authentication below.

**Web Based Authentication** – Use this option to force the tool to prompt you with a pop up window that will display your company login page to provide credentials. If you are using Office 365 with an ADFS server with custom login page, Multi-Factor Authentication/a PIN, is using ADFS, or has SSO enabled, use this authentication type. If you want to schedule jobs, click to “Remember Credentials” and run the job once, it will remember these credentials for the scheduled job.

**Remember Credentials**

Select this option to store the username and password. Passwords are stored using strong ‘salted’ encryption.

### Scan Options

**Include SharePoint Lists** – Check this option to scan all links within Lists (such as Tasks, Issues, custom lists).

**Scan List Item Attachment File Contents** – Scans all links within List Item attachments.

**Include Page Libraries** – Scans all links within lists based on the Page Library template.

**Include Document Libraries** – Scans all links within Document Libraries.

**Scan Document File Contents** – Scans content within each file with a matching extension specified in “File Content Scan”.

**Include System Lists** – Scans all links within Hidden System lists and libraries.

**Only Show Broken Links in Reports** – Check this option to only show broken links, links that are ‘OK’ will not be shown in the report but will be shown in the log file.

**Show Detailed Logging** – Check this option to show details of every URL scanned in the Active Logs. This will affect performance so we recommend after your jobs are working fine, to disable this option unless needed for troubleshooting.

**Open Report Once Complete** – Check this option to automatically open the report once the job is complete. This option is hidden/not available for ‘Multiple Sites’ jobs.

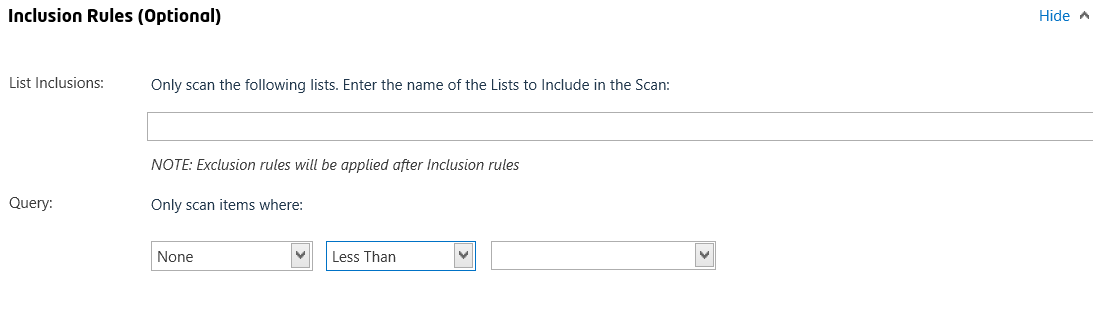
### File Extensions

If the option to scan file contents is enabled (for List Attachments or for Document Library files), you can specify the extensions to include in the scan here. If you selected “Include Page Libraries”, the file extensions textbox will become enabled.

Only files whose extensions match those listed here will have contents scanned for broken links.

The first textbox in this section is for scanning within Web Page Libraries (enter the file extensions to include in these libraries when performing a scan), by default usually most environments only have aspx pages in the Web Page Libraries (such as “Pages” or “Site Pages”). The second textbox in this section is for scanning within Document Libraries (enter the file extensions to include in these libraries when performing a scan).

### Inclusions Rules (Optional)



**URL Inclusions –** Use this section to filter the report by URL(s). You can enter a URL to only scan the SharePoint pages, items, files that match this URL pattern. Enter the URLs of the ‘pattern’ that you want to include in the scan. You can enter an exact URL to match or partial using wildcards (\*). For example: entering [https://sharepoint/sites/hr/ap/\*](https://sharepoint/sites/hr/ap/*) here will include the HR Accounts Payable site (“/ap”) and all contents within it from the scan. Another example: if you want to include all sites and items that have a managed path such as “Clients”, you can enter “\*/clients/\*” which will include all SharePoint objects whose URL path has /clients/ within it. If the SharePoint site, list, library or item does not match the URL specified here, it will be skipped. Once a match is found, all links within the target item will be scanned (inclusion rules do not apply to page/item/file contents). Separate URLs using semi-colons ‘;’.

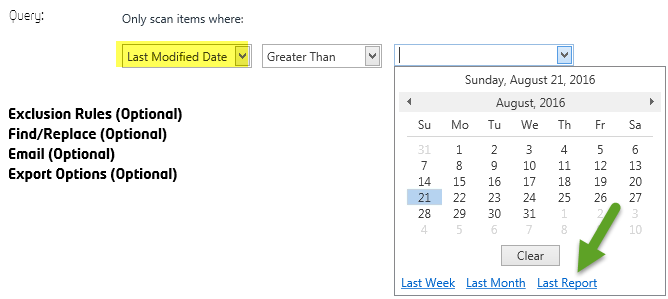
**List Inclusions** – Use this section to filter and only scan specific SharePoint Lists in the job. Use List Title (display name), not the List URL here. The user can specify what Lists to include in the scan by entering the names lists in this section, once you enter a SharePoint List in this section, all other lists not entered here will automatically be excluded.

If the name of the list exists in multiple sites, they will all be included if the List title (display name) matches. Wildcards (\*) are supported, so if you want to include all lists and libraries that have the word ‘documents’ in the List Title, you can enter “\*documents\*” and all lists and libraries with ‘documents’ will be included in the scan. These values are not case-sensitive. Once a match is found, all links within the list and its items will be scanned (inclusion rules do not apply to page/item/file contents). Separate Lists using semi-colons ‘;’.

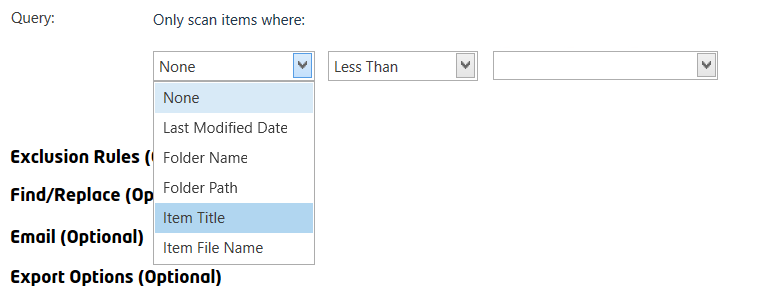
**Query – You can specify to only include items that match criteria matching criteria specified here.**

**You can choose to only do a “Differential Scan” by selecting the options below. It will then only build a report against items and files that have been modified since the last Broken Link Manager Report date.**

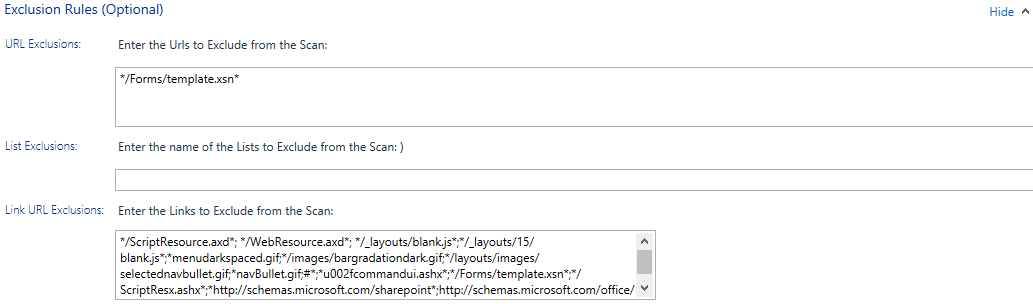
**Choose “Last Report” for a “Differential Scan” or you can specify a custom date here:**



**You can also only scan folders or items that match criteria given here.**



### Exclusion Rules (Optional)



**URL Exclusions -** Enter the URLs of the SharePoint target sites, lists, libraries, items that you want to exclude in the scan. You can enter an exact URL to match or partial using wildcards (\*). For example: entering [https://sharepoint/sites/hr/ap/\*](https://sharepoint/sites/hr/ap/*) here will exclude the HR Accounts Payable site (“/ap”) and all contents within it from the scan. Separate URLs using semi-colons ‘;’.

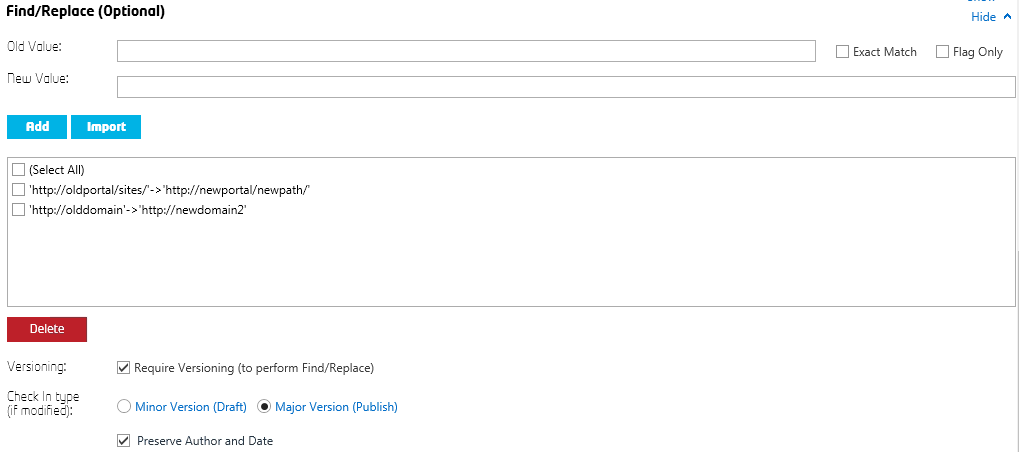
**List Exclusions -** The user can enter the names of lists to exclude in the scan here. If the name of the list exists in multiple sites, they will all be skipped if the List title (display name) matches. Wildcards (\*) are supported, so if you want to include all lists and libraries that have the word ‘documents’ in the List Title, you can enter “\*documents\*” and all lists and libraries with ‘documents’ will be included in the scan. These values are not case-sensitive. Once a match is found, all links within the item will be skipped. Separate Lists using semi-colons ‘;’.

**Link URL Exclusions -** The user can enter URLs to exclude in the scan. This will be applied against the links found within items, documents, files and page contents (it is not applied against the SharePoint site/list/item URLs, use ‘URL Exclusions’ or ‘List Exclusions’ options above for excluding SharePoint content from scans.). Wildcards are supported. These values are not case-sensitive. Separate Link URL Exclusions using semi-colons ‘;’.

### NOTES on Inclusion/Exclusion Rules:

* **Exclusion rules are applied AFTER Inclusion rules (if any).**
* **If the “Minimal Download Strategy” feature is enabled on a site, the URL changes to a path similar to** <http://contoso/_layouts/15/start.aspx#/Shared%20Documents>**. Do not use the path in the address bar if using “Minimal Download Strategy” AND URL Inclusions/Exclusions as this path is virtual and will not be used by the tool. For example, for lists, use the list URL specified in the SharePoint List Settings page when using URL Inclusions/Exclusions for lists, such as** <http://contoso/Shared%20Documents>**.**

### Find / Replace (Optional)



You may use the Find / Replace feature to automatically fix target URLs.

Find / Replace will modify URLs (regardless of HTTP Status Code) that match a specific user-defined value. Find /Replace will search within all content where it performs a scan for broken links, such as SharePoint Page contents, MS Office documents (Enterprise Version only), list item metadata, PDF file contents, text files, html files, etc…

See section (below) “Automatically Finding and Replacing Links” for more information.

**Old Value** - Enter the old URL value you are searching for and would like to replace.

**New Value** - Enter the new URL value you want to use to replace the old value with.

**Flag Only** – This will not cause any changes and will only flag items as matches for a find/replace. This is useful to test the scan before actually performing a find/replace.

**Exact Match** – Checking this on will force the replace to only occur when an exact match is found (case sensitive)

**IMPORTANT**: Exact Match only works with absolute URLs, not relative links.

**Add** – Click this button once you have added an ‘Old Value’ to replace and a ‘New Value’ or selected the ‘Flag Only’ option

**Import** – Click this button to bulk import Find / Replace values from a MS Excel file. First column should be ‘Old Value’, second column should be ‘New Value’.

**Version State** – (Only used if Versioning is turned on). This option only applies to Find and Replace where the file to be updated (with a Find and Replace), is in a SharePoint List that has Versioning enabled. Once a Find / Replace match is found, the utility will automatically check out and check in the file and create a new version for it (if versioning is enabled on the list). This Version State option specifies the type of version that will be created if a Find and Replace occurs.

**Minor Version (Draft)** – If a Find and Replace match is found, the tool will perform the Find and Replace and it will save the updated item as a Minor version.

**Major Version (Publish)** – If a Find and Replace match is found, the tool will perform the Find and Replace and it will save the updated item as a Major version.

**Preserve Version Type** – If a Find and Replace match is found, the tool will check the current version of the item (check the version number and determine if it is a major or minor version, such as 2.3 which would be minor, or 3.0 if it was a major version), perform the Find and Replace and it will save the updated item as the same version type as it was originally. This way, if the item was in Draft state and was in the middle of changes, it will not be published as the Version Type will be kept (such as 2.4 for minor, or increment to 4.0 for major version type).

If the file is checked out to a different user than the tool is running as, the item/file will not be updated and will not be checked in, you will see a warning in the Active Logs.

Example if Preserve Version Type is enabled:

**Original File**

File1.docx v0.3 (Minor)

File2.docx v3.0 (Major)

**After Find Replace**

File1.docx v0.4 (Minor is preserved)

File2.docx v4.0 (Major is preserved)

**Preserve Author and Date** – This will ensure the ‘Modified By’ and ‘Last Modified Date’ field values are not changed by the tool when one or more URLs are replaced in the item.

**IMPORTANT:** The tool creates a new ‘Minor Version’ when it ‘Preserves Author’. This is by design and is a limitation of SharePoint API where we cannot Preserve Author as a Major version and we cannot perform a Find and Replace and Preserve the Author in the same version, it must be done in 2 steps so therefore, you will have 2 new versions after a Find and Replace.

**NOTE:** The job must be run as an account with Full Control on the Site in order to use “Preserve Author” as the SharePoint object model requires Full Control in order to modify this type of information.

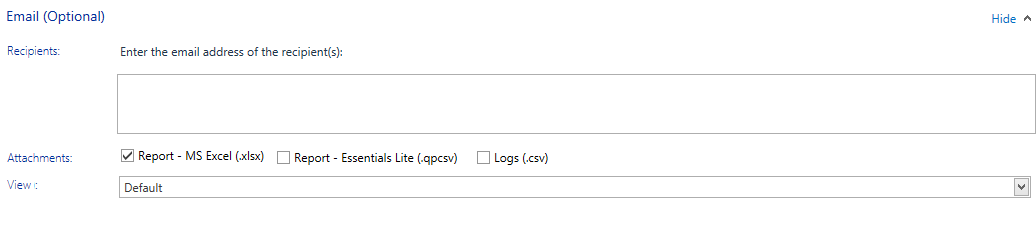
For additional information on Preserve Author, see the [Preserve Author](#_Preserve_Author) section.

**Preserve Content Approval State** – If a Find and Replace is performed and the item has Content Approval enabled on the List, enabling this option will cause the tool to ‘Preserve the Content Approval Status’ of the item. For example, if the list item was in ‘Approved’ state, a change normally will cause this ‘Approved’ item to go back into Draft state. If you enable this option, after a Find and Replace occurs, the tool will set the ‘Approval State’ back to what it was originally (before the Find and Replace occurred).

**NOTE:** The job must be run as an account with Full Control on the Site in order to use “Preserve Content Approval State” as the SharePoint object model requires Full Control in order to modify this type of information.

**IMPORTANT:** The ‘Approval Status’ of the item will be kept however the ‘Approved By’ person will show up as the user in which this job is run. This is a limitation of the SharePoint API as you are not allowed to approve an item as someone else.

### Email (Optional)



You can optionally enter one or more email addresses for the users who you would like the report to be emailed to. Separate multiple email addresses with a semi-colon ‘;’. Once the report is completed, it will attempt to email the recipients found in this list.

TIP: Enter <Contact Email> for the email address to use the ‘Contact Email’ specified for the site in the Home Page Dashboard. For more information, see the “SharePoint Essentials Toolkit User Manual”.

**Attachments**

*Report - MS Excel (.xlsx)*   
Optionally include the MS Excel report as an attachment. This option is set to Yes (checked) by default. A custom View (see below) can be specified below to set custom filters and groupings to the MS Excel file that is emailed.

*Report – Essentials Lite (.qpcsv)*   
Optionally include a QiPoint Essentials Lite report as an attachment. The client must have the QiPoint Essentials Lite Client (ClickOnce application) installed on their computer to open this type of report.

*Logs (.csv)*   
Optionally include the log files as an email attachment (they will be compressed/zipped).

**View** – (Only applies to MS Excel Report, not SharePoint List Exports or QiPoint Essentials Lite Client reports.) Select the view to use when creating the report. These views can be created from the Broken Links Report page (See [Customizing the Report->Views](#_Views) for more information).

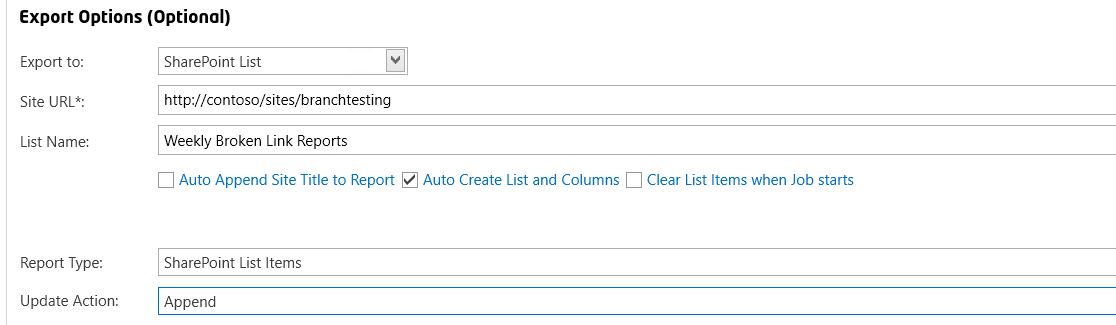
You can select a custom view from the drop-down list to apply it to the MS Excel file that is emailed to user(s). All grouping, sorting and filtering set in these views will be applied to the MS Excel report emailed. If you use the default view, the MS Excel will show a flat list without any grouping or filtering.

### Export Options (Optional)



This allows you to choose how to export the report when the scan job completes. You can use the default ‘MS Excel / CSV’ or select ‘SharePoint List’ to export the report to SharePoint.

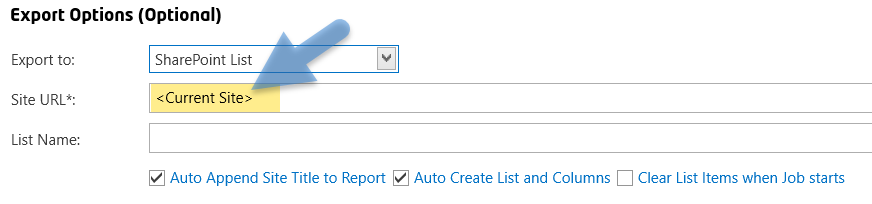
#### Export to SharePoint List



**Site URL**

This is the target Site URL where the report (SharePoint List) will be created.

<Current Site> - You can use this token / keyword to use the currently scanned site. If you have performed a ‘multi-threaded’ job by holding SHIFT or CTRL from Home Page Dashboard, this will export the report to the current site being scanned.



**List Name**

This is the List Title to be given to the List report.

**Auto Create List and Columns**

Check this option to automatically attempt to create the List and Columns. If the list already exists, it will attempt to create the list (only if the list does not exist) and recreate missing columns. If this option is UNCHECKED, it will not attempt to create the list or create the list columns. SharePoint Views that are created for the list are not overwritten when a report is generated/exported.

**Auto Append Site Title to Report**

This will automatically append the Site Title to the end of the List Name. For example, if the List Name is set to ‘Broken Link Report’, and the Site Title is ‘Information Technology’, the SharePoint List report that will be generated will be titled “Broken Link Report - Information Technology”

**Report Type**

You can choose to export the report to a SharePoint List as:

1. SharePoint List Items

NOTE: for reports with many Broken Links, this option may make the list a “Large List” of over 5000 items. If you have many Broken Links and do not want to reach this threshold, please use one of the other options below)

1. MS Excel File – The report will be uploaded as a MS Excel File
2. QPCX File (SharePoint Essentials Lite Client format) – The report will be uploaded as a .qpcx file. This file type can be then opened by any end user who installs our free SharePoint Essentials Lite Client (WPF ClickOnce application).



**NOTE:** If this job is targeting bulk sites (when creating reports against multiple site collections (by holding Shift/Ctrl from home page dashboard to create a report against multiple sites, see ***Batch Processing Jobs for Multiple Sites***): If ‘Auto Append Site Title to Report’ is turned on, the Site Title will be appended to the end of the name of the report automatically and a report will be created for each site scanned (multiple SharePoint Lists, one per site). If ‘Auto Append Site Title to Report’ is turned off, and you are exporting to a SharePoint List, all sites being scanned in the job will be merged into one single report (a single SharePoint List).

**Update Action**

Overwrite: This will first delete all SharePoint Items found (if list and items exist), then it will export the new report items to the list. The list will not be deleted so settings and views for the list remain intact.

Append: This will append report items to the list and if there are existing list items, they will remain intact and will not be overwritten or changed. If the “Report Type” is set to MS Excel or QPCX formats, this option will append the new report by adding a date time stamp to the file name. You can turn off “Append” and turn on the library versioning to always show the most recent Permissions report and have old versions drop off based on the library major versions limit.

Update: (Available for SharePoint List Item Export) This will refresh the list of entries in the SharePoint Report for this site. If you are using a scheduled job for multiple sites that will add SharePoint List Items to the SAME SharePoint ‘Report’ List, you can use this option so that the report data is always ‘Updated’ for the site being reported on.

# What is a Broken Link?

## A broken link is a link that doesn't work, often resulting in an error page. A broken link happens when the link points to a web page that has been deleted or moved.

Each type of broken link can be categorized by a ‘Status Code’. These Status Codes are represented by numbers (“codes”) and can be translated to something ‘human readable’ such as 404 = “Page Not Found” 401 = “Unauthorized” 500 = “Internal Server Error” 403 = “Forbidden”, etc…

These Status Codes are provided from the web server when we try to ‘hit’ the URL, we do not create these codes.

A list of all possible ‘HTTP Status Codes’ can be found here:

<http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html>

<http://en.wikipedia.org/wiki/List_of_HTTP_status_codes>

## Manually Finding and Replacing Links

When a broken link is found, it will be listed in the report with a HTTP Status code, such as 404 (not found).

When a broken link is found, the user may right click on the item in grid view report and click “Open Item” to open the SharePoint item to edit/fix. The user can also click “Open Link” which will open the link in the default web browser to verify the issue.

**TIP:** Group report items by “HTTP Status Code” and “Item URL” to easily see the links that are broken.

# Automatically Finding and Replacing Links

You may use the Find / Replace feature to automatically fix incorrect URLs.

The tool will Find & Replace URLs that match the given text value you enter, regardless of the URL HTTP Status Code, another words, the link could be a valid one or a broken link, if it matches the input string you enter in ‘Old Value’, it will get replaced with the updated string/text entered in ‘New Value’.

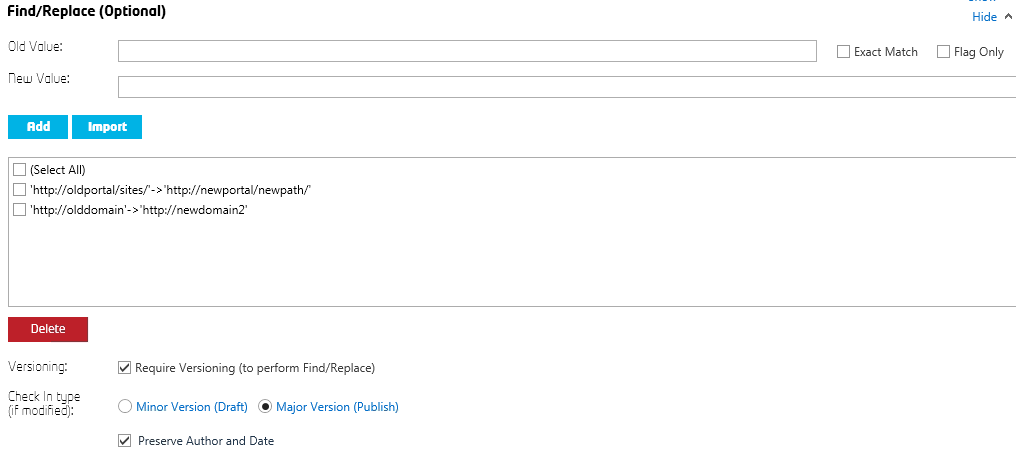
For example: If you enter

**Old Value:** <http://oldportalurl/sites/oldsitename>

**New Value:** <http://mynewportalurl/sites/newsitename>

And the tool finds a URL in a SharePoint page <http://oldportalurl/sites/oldsitename/Lists/Tasks/> it will update this link to become <http://mynewportalurl/sites/newsitename/Lists/Tasks/>

Below is a screenshot of the Find Replace section in the tool, in this example below, I added 2 URLs to be replaced by the tool:



## What will be replaced

You can perform a Find/Replace by exact or partial match:

**Partial Match:** You can enter the text you want to match by putting that value in the “Old Value” text box, such as <http://oldportal>, or “oldserver”, or “/sites/” (if you want to update a managed path). The tool will find all matches contained anywhere in the URL and make all replacements for all occurrences of that text found in the URL. This is NOT case sensitive.

**Exact Match:** You can use an “Exact Match” to perform a find replace where the “Old Value” must be exactly equal to the URL found, otherwise it is not replaced. This IS case sensitive.

NOTE: Exact Match only works with absolute URLs, not relative links.

URLs found in the following locations that match ‘Old Value’ will be replaced:

* SharePoint Lists
* SharePoint Document Libraries
* SharePoint Page Libraries
* Web Pages (such as in Wiki content zones, metadata, and web parts (see below for supported Web Parts))
* Document File Contents are supported for Find/Replace:
  + PDF (except ‘image-only’ PDFs such as hard copy documents scanned and did not use OCR to enable reading/searching of text data)
  + MS Word 97 and later
  + MS Excel 97 and later
  + MS PowerPoint 97 and later
  + MS OneNote 2003, 2007, 2010, 2013 (2016 not yet supported)
  + Txt, RTF, HTML, and other ASCII encoded file types
* Links behind friendly text such as “[Click Here](http://someurlpath.com/)” will be reported and replaced. “Click Here” would appear in the reports under the “HREF Display Text” column
* URLs within link tags, such as the ‘href’ and ‘source’ tags, are replaced (if match is found)
* Links within href display text are replaced (if match is found) (which are often the same as the URL text above)
* Link Description fields shown as ‘Title’ in anchor tags are replaced (if match is found) (which are often the same as the URL text above)
* OOTB Web Parts:
  + The following OOTB web parts are supported and URLs within them will get replaced:
    - Script Editor Web Part
    - Xslt List View Web Part
    - Image Web Part
    - Page Viewer Web Part
    - Summary Links Web Part
    - Content Editor Web Part
  + Web Part Properties such as Title URL and JSLink URLs are supported by the tool and will be parsed for broken links.
  + NOTE: Content Editor web parts with linked/embedded files (such as when pointing the Content Editor web part to an html file (in web part properties) are scanned separately if the file is located in the SharePoint site (in a list or library). Content Editor web parts with links typed directly into the web part area (‘Edit Web Part’) are scanned for links and replaced.

## Supported Link Types

* + Http, Https
  + FTP
  + File shares and UNC paths (such as “\\server\fileshare\folder1”)
  + Tooltips that contain hyperlinks
  + Descriptions or Alt URLs within hyperlinks

## Find/Replace Order

When you have multiple Find/Replace strings entered in the tool, the tool will attempt to match each string (“Old Value”), one at a time, in order from how you see it in the tool from Top going Down. If a match is found, it will attempt a replace (unless Flag Only is set). The tool will continue to go down this list of “Old Values” to match the string to the URL links found, regardless if there was already a match.

The product was designed this way in case you need to further manipulate a URL, you are able to add replacements for an absolute URL, which can get checked first, and then relative URLs such as a managed path, to get replaced after. As you may have instances of both in your environment.

We suggest adding in order of absolute URLs first and then relative URLs after so that you do not get unexpected results.

## What will NOT be replaced

* Text based links in MS Office files (such as URLs that are present in a web page or document (as text), but do not show a **hyperlink tooltip** when you hover mouse over them) are not supported by Find/Replace.

**Example of ‘text based’ link:**



* URLs in User Profile pages (metadata), such as Skills, Description, will not be scanned or replaced by this tool. Currently User Profile metadata is not supported, but SharePoint Libraries in My Sites (and OneDrive) are supported.
* Items that are checked out to another user will not be updated and therefore any matching URLs within them will not be replaced. There will be an error in the logs that shows the file is checked out to another user and will not be updated. If the file is checked out to the current logged in user (running the scan), the file will get updated (Find / Replace will be executed) and the file will be checked-in appropriately depending on the ‘Check In type’ setting. **It is recommended to have all files Checked-In prior to running the Find/Replace Job.**
* Find/Replace URLs in MS OneNote are not fully supported (Links are still reported within these file types and some types of hyperlinks (in certain OneNote sections) are able to be replaced)
* Find/Replace URLs in External Lists are not supported
* Web parts that have a URL which are not hyperlinks (such as plain text URLs that do not show a **hyperlink tooltip** when the mouse is hovered over them) are not supported.
* Find/Replace URLs in Quick Launch or Global (Top) Navigation. These links are stored in the SharePoint content database and the tool does not attempt to replace these, they will be listed in the report, but cannot be replaced by the tool automatically. However, the tool will help identify these and would need to then be fixed manually, such as in the Web Browser, CSS, JavaScript, custom Layouts file, or in Master Page.

## Versioning

If versioning is enabled on the library, a new version is created for the file automatically with a comment indicating the tool has performed a find/replace. NOTE: See Preserve Author section on how to keep “Last Modified By” and “Last Modified” metadata after a Find/Replace.

The tool will NOT attempt to perform a Find/Replace on any items in the version history except the most recent version of the file.

### NOTE

If Minor versions are enabled on the list, and you choose to Check-In Find/Replace changes as a Major (Published) version, and also set ‘Preserve Author’, the version will be incremented once as the Major version, such as version 3.0, this new Major version will have the Find/Replace changes; then it will check-in the item and set the original Author info, this then shows as an extra minor version (in order to display the last modified info), such as version 3.1, this is by design.

If you choose to Check-In Find/Replace changes as a Minor (Draft) version, and also set ‘Preserve Author’, the version will be incremented once as the Minor version, such as version 0.2 (if the original version is 0.1), this new minor version will have the Find/Replace changes; then it will check-in the item and set the original Author info, this then shows as an extra minor version (in order to display the last modified info), such as version 0.3, this is by design.

Minor (Drafts) and Major (Published Versions)  
When performing a Find & Replace, the tool will modify the most recent ‘Published’ Major or Minor version it has access to. If the tool account used has permission to Edit a draft version, and the Draft version is the most recent version of the file, the tool will perform the Find & Replace on the draft version and will Check-In the Item as a Major or Minor version depending on the setting used in the tool. If the Item has only a Major version as the most recent version, it will perform a Find & Replace on that Major (Published) version and Check-In accordingly.

### Check In & Check Out

* As mentioned in ‘What will NOT be replaced” (when using Find/Replace): Items that are checked out to another user will not be updated and therefore any matching URLs within them will not be replaced. There will be an error in the logs that shows the file is checked out to another user and will not be updated. If the file is checked out to the current logged in user (running the scan), the file will get updated (Find / Replace will be executed) and the file will be checked-in appropriately depending on the ‘Check In type’ setting.
* It is recommended to have all files Checked-In prior to running the Find/Replace Job in order to do all replacements found.
* If versioning is enabled on the library, and it has Major and Minor versions enabled, you have the option to Check-In as Minor (Draft) or as a Major (Published) version.

## Preserve Author

If you enable “Preserve Author” in the tool settings, the Last Modified By and Last Modified Date will be kept for files that are updated by Find/Replace.

#### Limitations

Preserve Author will only work if the account used to run the tool has Full Control on the site.

The item will retain the version for the previous Author (if versioning is enabled), however, the tool is forced to perform a Check Out and Check In to perform the update (Find/Replace), then Check Out and Check In to set/preserve the original Author info.

## Content Approval & Find / Replace

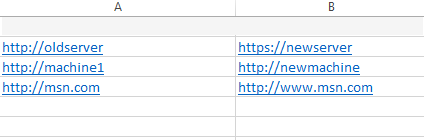
If Content Approval is turned on in the list, if the item is updated (during Find/Replace), the Approval Status will be set to “Pending” again. An Approver will need to approve the document again, there will be a comment when it is checked-in “Automatic Check In – QiPoint Broken Link Manager Find / Replace”

## Bulk Import Find Replace Links

You may import from a MS Excel file a list of URL replacements to use. This can be handy if you have multiple sites and want to have Site Owners send you MS Excel files with fixes to URLs.

To do this,

1. Create a MS Excel file that has two columns, ‘Old Value’ and ‘New Value’. The ‘Old Value’ will store the old value to be replaced, the ‘New Value’ will be the new URL to replace the ‘old value’ with. Enter all of the old URLs and New URLs that will be replaced. You do not need to use column headers in the Excel file when importing.



1. Click “Import” from the Find / Replace section.



1. A “File Upload” pop up window will appear, browse and select the MS Excel file and click “Upload”.

The entries will now be imported and listed in the Find/Replace list box.

## Find / Replace Options

#### Old Value

Enter the old URL value you are searching for and would like to replace. This can be a complete absolute URL or partial. For example: <http://oldserver> will match <http://olderserver> and <http://oldserver/sites/test1>. A ‘New Value’ of <http://newserver> would replace <http://olderserver> with <http://newserver>, and <http://oldserver/sites/test1> with <http://newserver/sites/test1>. You can also replace partial strings/text within the URL, for example, an ‘Old Value’ of ‘/oldpath/site1/site2/’ and ‘New Value’ of ‘/newpath/site1/’ would replace ‘http://www.company.com/oldpath/site1/site2/lists/docs/’ with ‘http://www.company.com/newpath/site1/lists/docs/’

#### Flag Only

This will not cause any changes and will only flag items as matches for a find/replace. This is useful to test the scan before actually performing a find/replace.

There are 2 places in the report that identify a link that is found and is to be replaced once executed:

1. The flagged items will appear in the Broken Links report under HTTP Status Code with an asterisk \* appended, such as 200\*
2. The flagged items will appear in the Broken Links report under New Link URL and New HREF Display Value fields as: (it will display the new URL value that will be used once executed)

*\*\*Match found - Flag Only:*[*http://newserver/somepath*](http://newserver/somepath)

#### Exact Match

Checking this on will force the replace to only occur when an exact match is found (case-sensitive)

#### New Value

Enter the new URL value you want to use to replace the old value with.

#### Add

Click this button to add the replacement strings (old and new values) to the application logic.

#### Import

Click this button to bulk import the replacement strings (old and new values) to the application logic. This expects a MS Excel file to import with first column ‘Old Value’ and second column ‘New Value’.

#### Delete

Select the replacement string in the list box above and click Delete to delete the selected replacement strings.

#### Check in Type

When a URL is matched with an ‘Old Value’ string, the document/item to be ‘fixed’ will be checked out automatically, updated to fix/replace the link, then automatically checked back in. Select the type of check in you want to occur for each document where a match is found and fixed. If you want to review documents before they are published, select “Minor Version (Draft)”.

#### Require Versioning

Check this ‘on’ to require versioning on the list in order to perform a find/replace. If this is checked and a list does not have versioning enabled, the find replace will not be executed. Otherwise, the find replace will execute whenever a match is found regardless of versioning setting on the list.

#### Preserve Author and Date

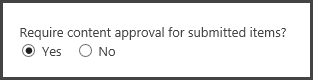
When a URL is matched and the document\item is to be updated (Find/Replace performed), select this checkbox to preserve the ‘Last Modified’ author and date for the item\item.

NOTES: The user account used to run the Job must have Full Control on the List/Item in order for “Preserve Author” to work.

Limitations: If a Library has “Require documents to be checked out before they can be edited?” set to “Yes” – Preserve Author will not be kept in published version. Please see “[Check In and Check Out](#_Check_In_&)” section for more details.

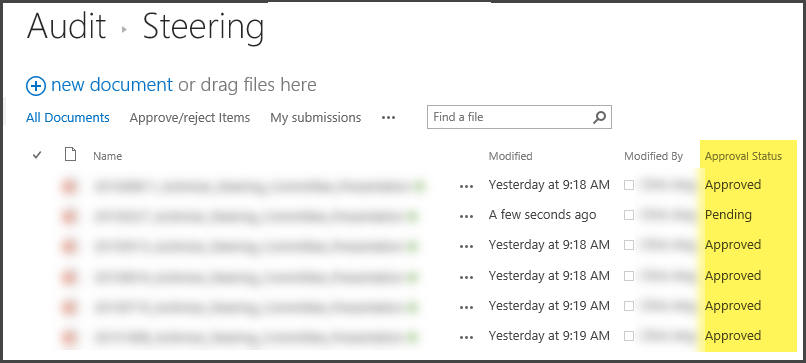
#### Preserve Content Approval State

This only applies to SharePoint Lists that have ‘Content Approval’ turned on.



If Content Approval is turned on, and when a URL is matched and the document\item is to be updated (Find/Replace performed), select this checkbox to preserve the ‘Approval Status’ for the item\item.

IMPORTANT: Preserve Author will not be kept for these items. Version history will be kept. The state is updated back to original after the Find/Replace is performed. Author cannot be preserved due to a limitation by SharePoint API as you cannot ‘Approve’ as someone else.

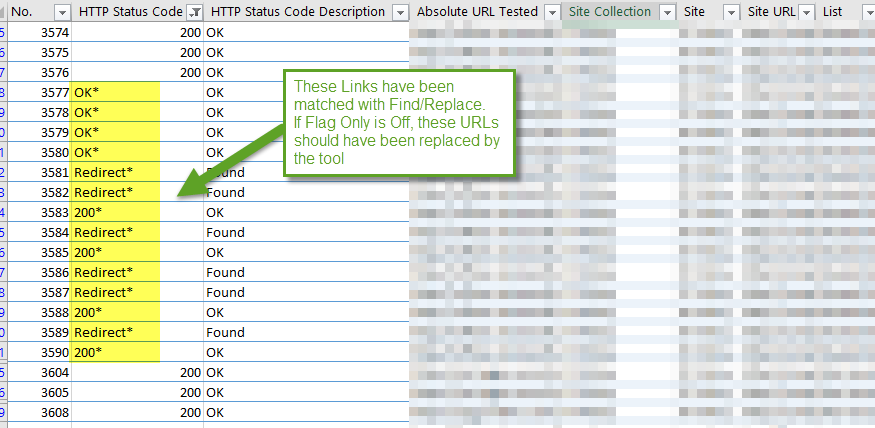


NOTES: The user account used to run the Job must have Full Control on the List/Item in order for “Preserve Content Approval State” to work.

# Link Report - Finding URLs that were automatically replaced/fixed

All URLs that have been matched with a ‘Find/Replace’ ‘Old Value’, will have an HTTP Status Code with \* appended to the end of it, such as 200\*, or 404\*

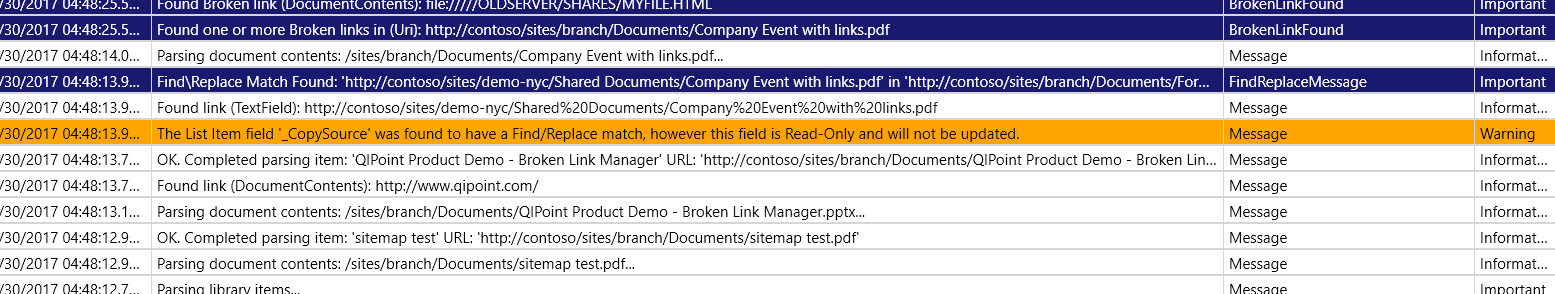
To review all documents that had links updated, the “Link Report” has a column labelled “New Link URL”, this field will be populated when a link is replaced, you can group, sort or filter by this column to quickly find updated documents that had URLs fixed/replaced.



## How to find all Links that matched Find & Replace but were not replaced?

Some links cannot be replaced by the tool, such as Copy Source field links, as this field is a read only field and cannot be modified by the tool.

Links such as these which might be matched by the Find & Replace string, but not replaced, will usually show a warning or error in the logs such as below:



Also, in the report, all links that were matched but not replaced will have an asterisk next to the HTTP Status Code, and will have a blank value for the ‘New Link URL’ ‘New HREF Display Value’ and ‘New Link Description’ columns, such as below:



# Scheduling Jobs

This product includes a built-in scheduler which allows you to scan sites on a schedule. Scheduled jobs run faster than jobs manually run as they do not have to write logs to the UI (user interface).

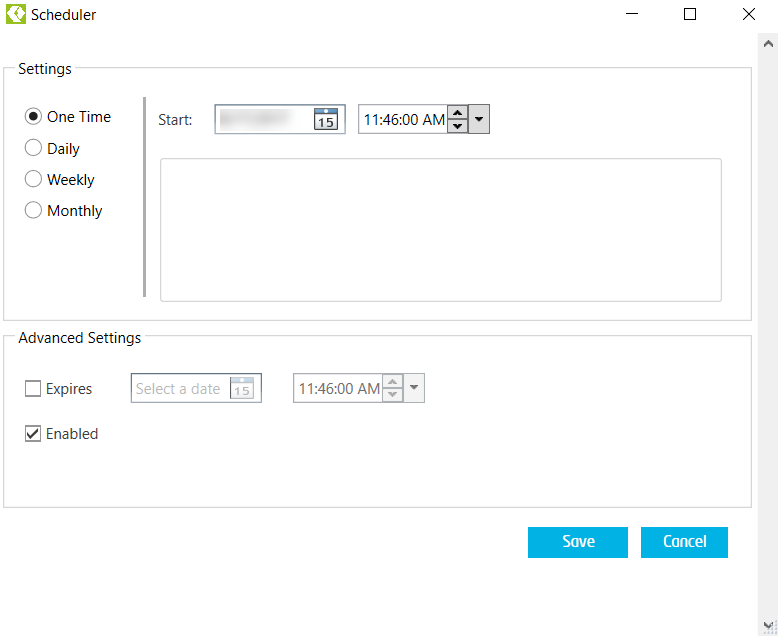
**TIP:** Ensure the “QiPoint Essentials Service” (Windows Service) is running before trying to schedule a job. Ensure the user account used to run the service has ‘Modify’ permission to the Report Path (Home->Settings). By default, this service uses ‘Local System’ which may not have rights to create files (reports) in a network file share for example.



## Schedule a Job

Once you complete the Scan Options page, click the “Schedule Job” button once ready to schedule this job. A window will appear to set the date, time and recurrence of this scan job.

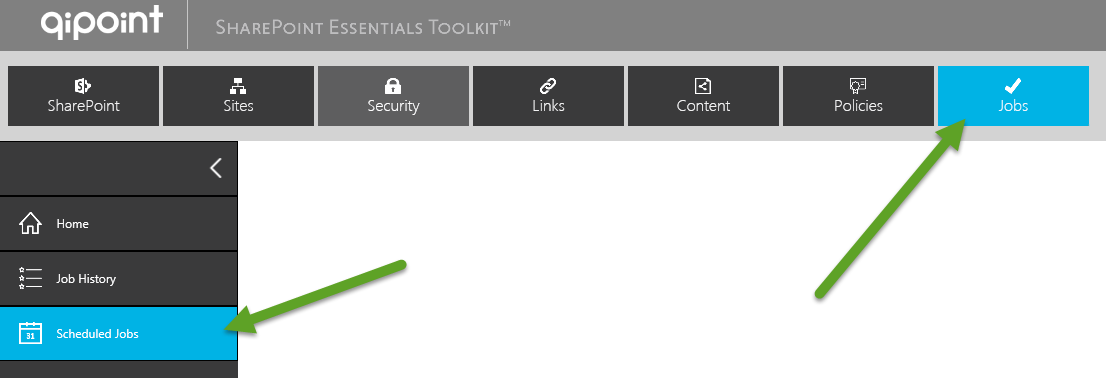


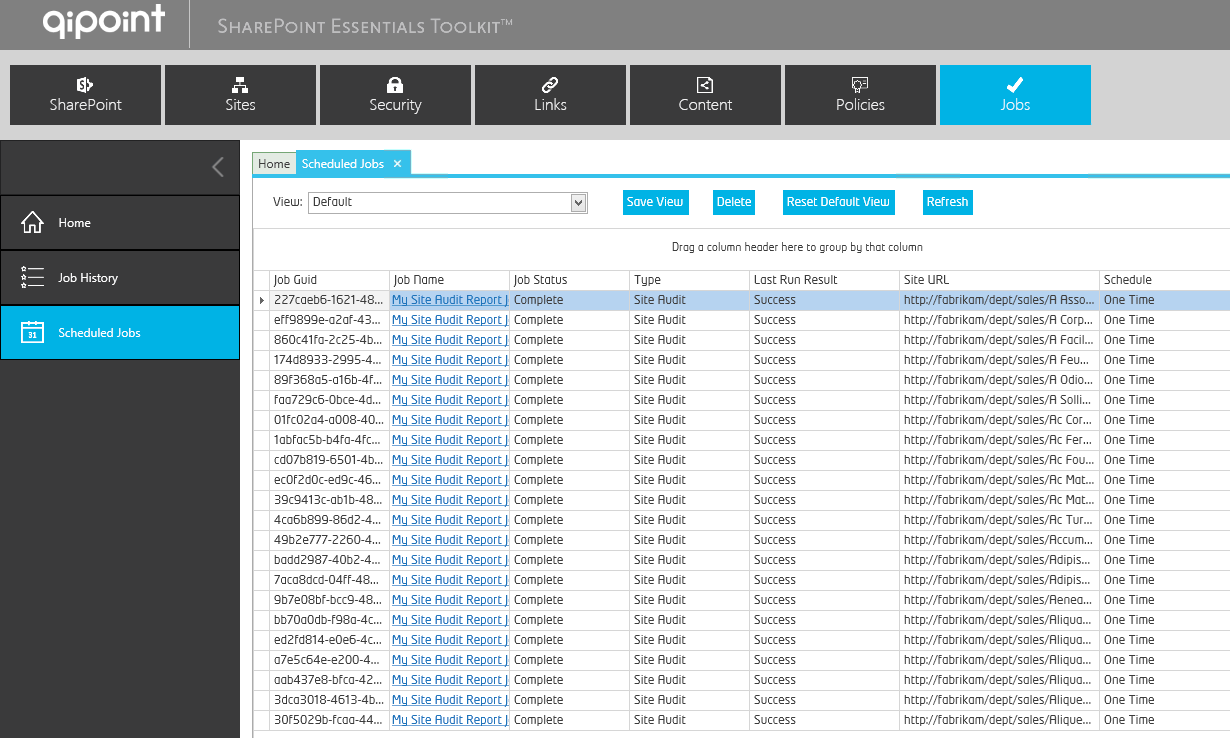


Select the schedule and specific options and click OK to save the scheduled job.

## List of Scheduled Jobs

You can view all scheduled jobs and their status’ by clicking on “Scheduled Jobs” from the Jobs section.





From here you can view reports, logs, execute the job, modify or delete scheduled jobs (by right clicking the job).

## Scheduled Job Logs and Reports

After a Scheduled Job is complete, you can right click it to view the logs or the actual reports. You can also access these logs and reports from the Home Page Dashboard, and also from the Job History section of the tool.

NOTE: Only applicable reports and logs will be shown.



# Batch Processing Multiple Site Collections

You can create a job to report against multiple site collections from the Home Page Dashboard by selecting multiple sites and right clicking. See SharePoint Essentials Toolkit (User Manual) -> Batch Processing Reports for Multiple Sites for more information.

# Licensing

## Components

The SharePoint Essentials ToolkitTM by QiPoint may have multiple components activated to provide more functionality.

To view what components are installed, on the ribbon, click Home -> Licenses.

## Changing the license type of a component

1. Start “SharePoint Essentials ToolkitTM”
2. From the “Home” Ribbon tab, click on “Licenses”
3. Click on “Manage” beside the product you want to changing licenses for
4. Select the **Version** of the application component you would like to activate, example, change to “Enterprise” for the Enterprise version.
5. If you have a paid license key, enter it in the “License Key” text box
6. Click “Apply”
7. Close any open tabs in order for the new licensing options to become available (or hidden if downgrading the license). The new options will appear once the tab is reloaded. You do not need to shut down the application or restart your machine.

## Trial Version

These are the limitations of the trial version

* + Up to 5 Site Collections
  + Can scan up to 100 List Items per list
  + Limited to 100 URLs per list item
  + Limited to 50 Lists

## Free Version

We no longer offer a free version of this product. You must obtain the professional or enterprise version to use the features of this tool.

## Obtaining a License

There are three ways you can contact us to purchase a license

1. Visit us at our [Online Store](http://www.qipoint.com/product/category&path=0)
2. Call us at 1-855-747-6468
3. Email us at [sales@qipoint.com](mailto:sales@qipoint.com)

You may place your order and receive a license key with a PO# or Credit Card (VISA, M/C, Diners Club, AMEX, PayPal)

## Transferring a License

You may need to transfer the license if the machine was lost, stolen, re-formatted, or if you would like to assign a license to another user. Contact us to obtain a new license activation key at [support@qipoint.com](mailto:support@qipoint.com).

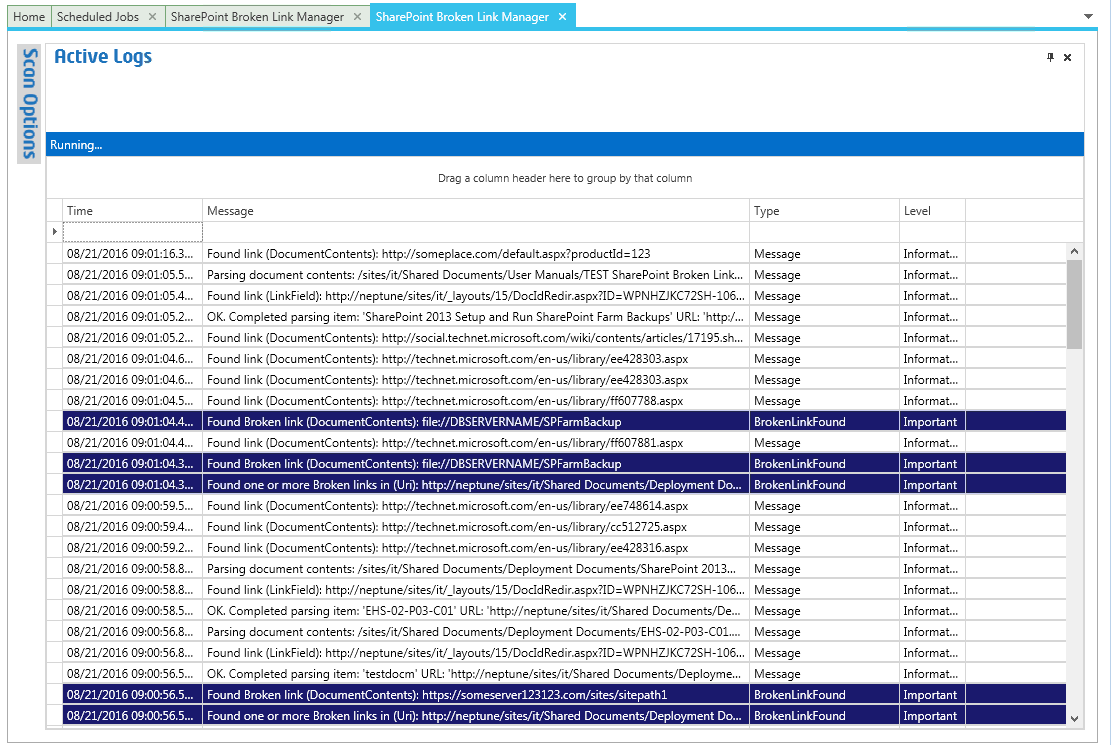
# Logs

As the Job is running, logs will be generated in real-time so that you can view the status and progress.

Log files and reports are stored under the path located in the Settings page (Home -> Settings -> Path to Store Reports).

Normally you can view any log by navigating to the Home Page Dashboard, and right clicking the site you ran the job, and then clicking “View Logs”. You can also view logs from the “Job->Job History” page, or if the job was scheduled, the logs are also accessible from the “Jobs->Scheduled Jobs” window.

Below is an example of the Active Logs page when a job is run manually (not in a schedule). Broken Links will show up with a blue background when detected. A similar log window will appear when viewing jobs from the “Job History” window or the “Scheduled Jobs” pages.



You can export job logs to CSV format by clicking the Export Logs button.



# How to Interpret the Report

### Example

Here is an example of a page with 2 broken links:



In the grid and in exported reports you will see the 2 broken links found, showing the location of each.

### Which Links Are Broken?

The column labelled “HTTP Status Code” is the column which will tell you if the link is broken or not. “HTTP Status Code Description” will also give a description of the response from the server when resolving a URL. Generally, an HTTP Status Code of 200 means the link is valid and reachable, anything else could mean a broken link (but not always the case), see here for more information on the HTTP Status Codes and how to interpret them:

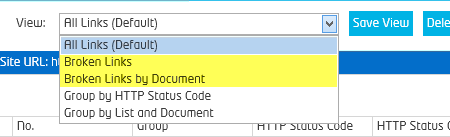
<http://en.wikipedia.org/wiki/List_of_HTTP_status_codes>

<http://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html>

See the ‘[Report Fields](#_Report_Fields)’ section below, that section will provide information about column in the report and how to validate and find the links in your environment.

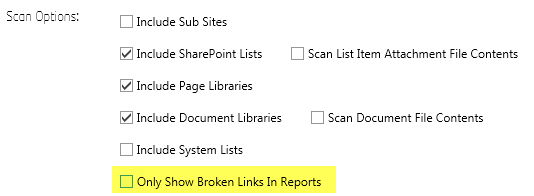
#### Views

You can use one of the built-in views to view only the Broken Links for the site:



#### Only Show Broken Links in Report

There is an option in the Scan Options page to only show broken links in the report. You can select this option to filter all valid links. However, sometimes it is useful to see all links to help with validation.



# Reports

## Opening the Report

Once a scan is complete you can open the report in several ways:

* + When a scan is complete, you will have a button on the top right (above the “Active Log”) that says “Open Report”
  + Once you Click “Open Report” and the Report opens, you can click “Export” to view the report in SharePoint/CSV/PDF/Excel/QPCX format
  + From the Home Page Dashboard, you can right click on a site that you wish to preview, and click “View Reports -> Broken Link Report”
  + From the “Job->Job History” page you can find previously run Broken Link Jobs and their reports and logs.
  + For scheduled jobs, you can also find the reports and logs under the “Job->Scheduled Jobs” page.
  + To view all report files (CSV and QPCX formats available)
    - Open Windows Explorer and navigate to the path where the reports are stored, the default location is in the user’s My Documents folder under “QiPoint\Reports”. You can change the report path under “Home -> Settings -> Path to Store Reports”

## Report Fields

#### *No.*

Line number. This is used as a reference to help identify items that need to be responded to as they are reviewed. This is not related to the actual link and is for reference purposes only.

#### *Group*

This is the group name that was given to the site during the scan. This does not affect SharePoint data; it is only used for reporting purposes. For example, you can set the name of sites to “Production” or “Test” for different environments.

#### *HTTP Status Code*

This is the code response from the server when testing for the broken link. Such as 404. If Find/Replace was used, this value will be prefixed with a ‘\*’ (such as \*404 or \*200) to identify this record as matching the Find/Replace value.

#### *HTTP Status Code Description*

This is the descriptive response from the server when testing for the broken link. Such as “Not Found”

#### *Absolute URL Tested*

The actual URL that the utility scanned to check for a broken link. If the link detected was a relative path, such as a web page with “/sites/site1/Lists/MyList”, the tool will automatically try to determine the absolute URL that the link would normally resolve to if the user clicks on it. If the relative link above was in <http://contoso/>, the Absolute URL Tested would be <http://contoso/sites/site1/Lists/MyList>. Another example of a relative link is ‘../../../Style Library/custom1.css’, this will translate to a path that will be calculated using our proprietary link algorithm to determine the actual SharePoint path that would be used to reach the link. The absolute URL tested may be something like this ‘<http://contoso/Style%20Library/custom1.css>’.

#### *Link URL*

The ‘actual’ URL of the link that was found. This is the actual link that you can search for within a page, item or file that you are validating.

#### *HREF Display Value*

This is the text that is displayed to the user in the broken link, it would link to the ‘broken link’. In web pages, this is the text within an Anchor tag, such as <a href=<http://contoso>>HREF Display Value</a>.

#### *Link Description*

In web pages, the ‘Title’ field in the Anchor tag is this value, such as <a href=<http://contoso>>HREF Display Value</a>. In MS Office documents, this is the “Screen Tip Text” value.

#### Site Collection

The title of the Site Collection that the scanned link belongs to.

#### *Site*

The title of the Site that was scanned.

#### *List*

The title of the SharePoint List that the link was found in.

#### *List URL*

The URL of the SharePoint List that the link was found in.

#### *Field*

The field name that the broken link was found in, for example, you may have a custom column called “Company Web Site” or “Description”.

#### *Link Type*

The type of broken link. Such as within “Document Contents”, “HTML field”, “Note Field”, “Content Editor Web Part”, etc…

#### *Item Title*

The title of the item that the broken link was found in.

#### *Item URL*

The item URL that the broken link was found in.

#### *New Link URL*

If Find/Replace was used, and a match was found, this is the new value for the Link Url that was updated in the document. If Find/Replace was used, and it was set to “Flag Only”, the flagged match will appear in this column and no data will be modified.

#### *New HREF Display*

If Find/Replace was used, and a match was found, this is the new value for the “HREF Display Value”. This would occur if the URL and HREF Display description are both URL values that match Find/Replace criteria. If Find/Replace was used, and it was set to “Flag Only”, the flagged match will appear in this column and no data will be modified.

#### *New Link Description*

If Find/Replace was used, and a match was found, this is the new value for “Link Description”. This would occur if the URL and HREF Display description are both URL values that match Find/Replace criteria. If Find/Replace was used, and it was set to “Flag Only”, the flagged match will appear in this column and no data will be modified.

Site Owner  
If the Site has the default “Owners” group, this column will display the names of the members of this group.

Author  
This is the author of the item (where the URL was found).

Last Modified  
This is the date and time the item (where the URL was found) was last modified.

#### Last Modified By

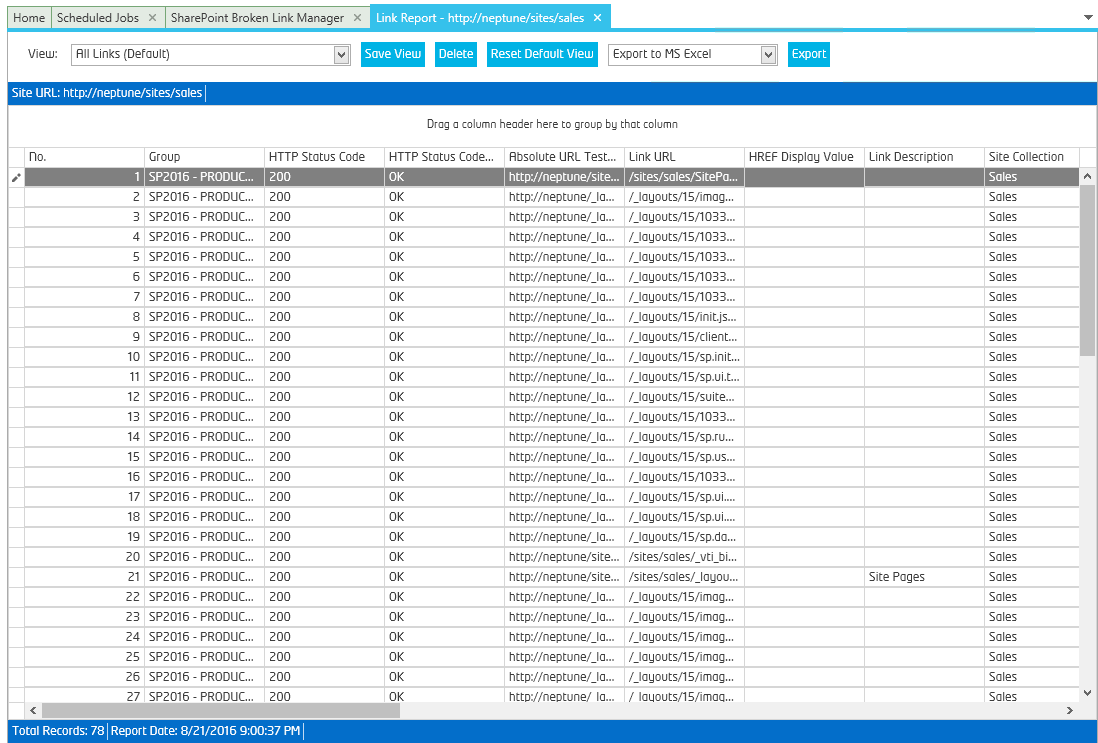
This is the user who last modified the item (where the URL was found).

## Report Archive

When a report is created and there is a report that already exists for that scan (this occurs if it is the same site URL that is scanned and on the same day), the old report (and the related log file) will be renamed and moved to a folder called “Archive” in the same directory. The new report will then be created.

## Grid View Report

Reports are viewed in a Grid View. This Grid View allows the user to drag and drop columns, sort, and apply advanced or simple filtering to help to generate a report most useful for them. For example, they can group to view all broken links within a specific document, or search and sort to find common broken links across lists and sites.



## Customizing the Report

You can filter, sort and group the results from the scan job.

### Filtering

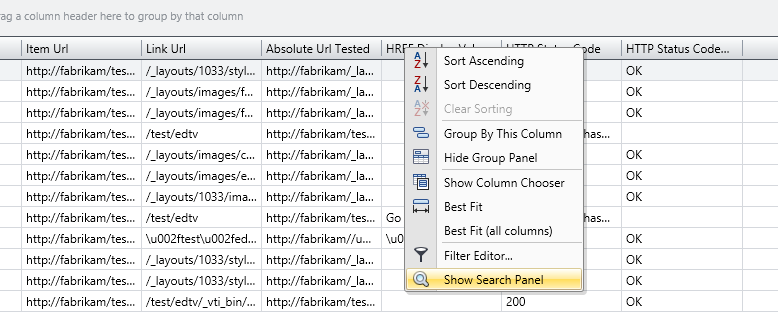
To filter the data based on a field, put your mouse over the column header that you want to filter by. You should see a filter icon, select it and choose the values you only want to have shown in the grid

### Advanced Filtering

To filter the data based on a field using advanced criteria, right click on the column header and select “Filter Editor”. You can now add criteria to filter the results

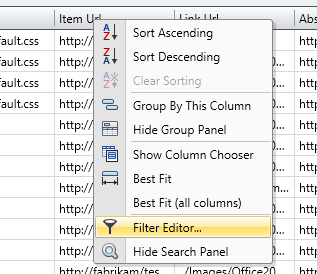
### Searching within the Report

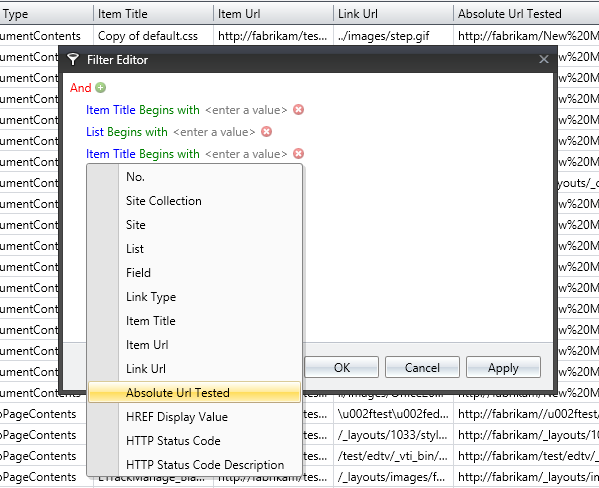
Right Click on any column header and select “Show Search Panel”. The search panel will appear and you can enter any value to search the grid, results will be highlighted



### Filtering Queries

You may also filter the report by using queries. Right click the column header and select “Filter Editor” to add criteria to filter the results

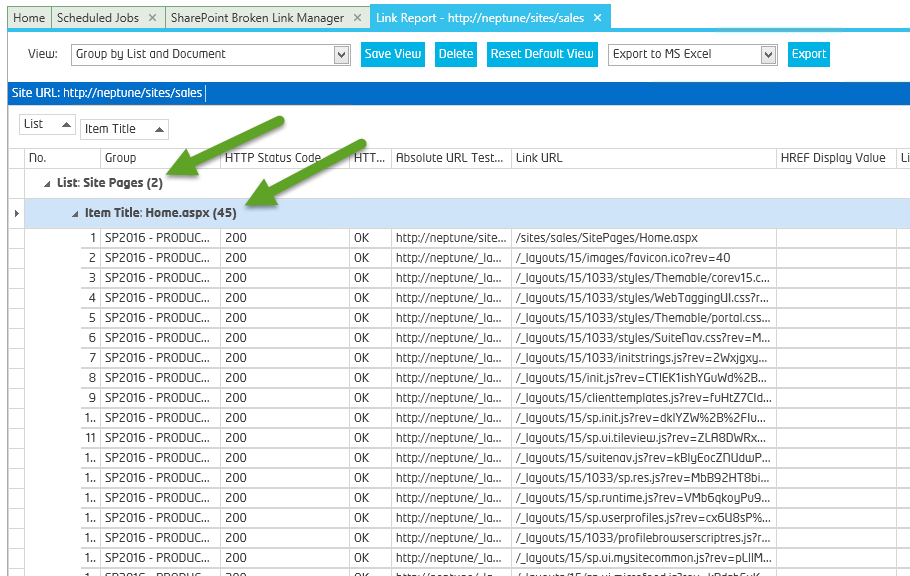




### Grouping

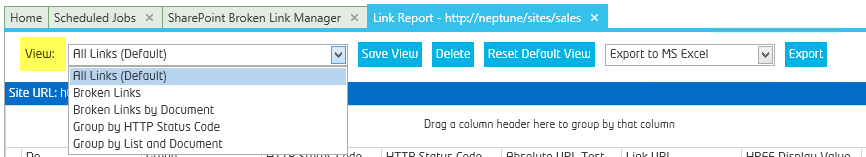
To group the results, drag the column header of the field you want to group by to the top of the results panel. You can also drag multiple panels and rearrange the group ordering by dragging columns left and right of each other

Below I grouped by List and Item, so here I can see all links found under the “Home.aspx” page:



### Views

The user can save views or modify the default view by clicking “Save View”. A view can be deleted by clicking “Delete”, the user will be prompted to delete the current view. The Default view cannot be deleted. “Reset Default View” will reset the Default view to the ‘out of the box’ view.

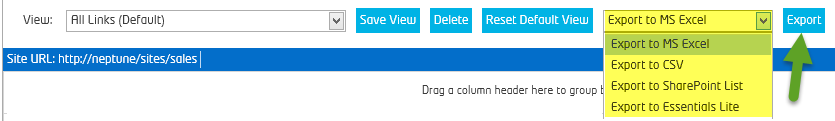


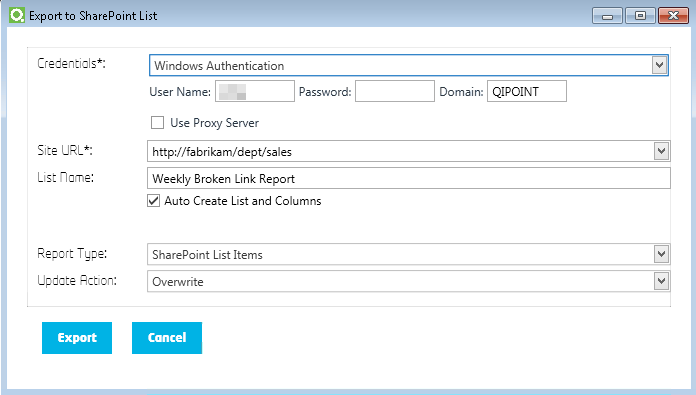
### Export

A report can be exported to the following formats directly:

1. MS Excel
2. CSV
3. SharePoint List
   1. As SharePoint List Items
   2. Uploaded as MS Excel File
   3. Uploaded as QPCX File (SharePoint Essentials Lite Client)
4. QPCX (SharePoint Essentials Lite Client)

To perform the export, select the option in the drop down and then by clicking the “Export” link. Alternatively, multiple rows can be highlighted and ‘Copied’ by right clicking and selecting “Copy”. Then the selected rows can be pasted in MS Excel, an Email, MS Word or any compatible program.

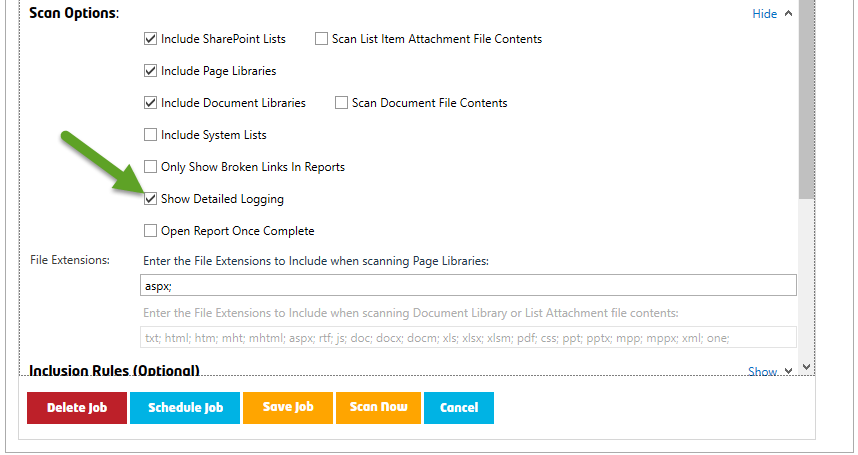




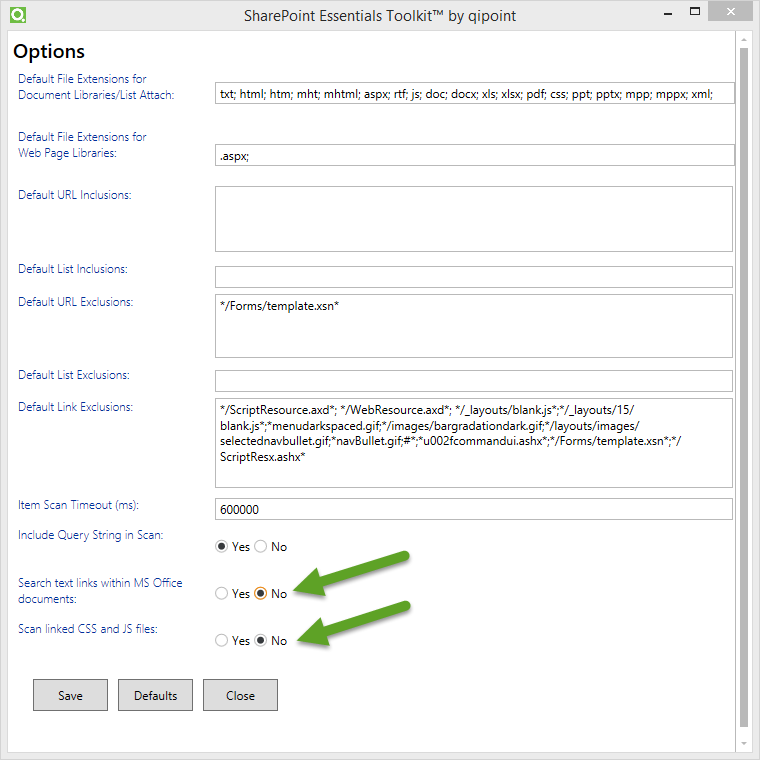
# Best Practices: Improving performance and Large Scan Jobs

For scanning environments with over 500,000 links or over 200GB data in a scan job/project, we recommend the following:

* See SharePoint Essentials Toolkit documentation for “Import Multiple Sites to Dashboard” in order to bulk import sites to the home page dashboard.
* Turn OFF/uncheck “Show Detailed Logging” – this prevents logging of ‘verbose’ information to UI and log file.



* Split scan job into multiple jobs that can run in parallel by going to the Home Page Dashboard, hold SHIFT and multi-select the sites and subsites you want to build a report on, right click and select Create Reports->Broken Link Report. This will allow the tool to run these jobs asynchronously and in multiple threads (vs. a single thread when you create a job and select ‘Include Sub Sites’ for one site collection with multiple sites, these will run sequentially, which is slower).
* Run the ‘split scan jobs’ on different machines (requires additional user/machine licenses). This will offload some resources that can be accumulated from a single scan job. Example, run half on machine 1, other half on machine 2. They can be scheduled jobs with email alerts.
* Run the ‘split scan jobs’ on the same machine but at different time intervals using the Scheduled Job feature or manually running them separately. For example, if you have 3 large libraries in a single site to scan, instead of scanning all libraries in a single job, create 3 separate jobs for each library and schedule them to run at different times. This will reduce the RAM required per job.
* If the machine has anti-virus enabled, please see Anti-Virus and Performance section to set directory exclusions.
* Use Exclusion Rules to exclude repetitive URLs and lists that are not required.
* Turn OFF/uncheck “Include System Lists”.
* Turn OFF/Set to No “Search text links within MS Office documents” in Scan Options page.
* Turn OFF/ Set to No “Scan linked CSS and JS files” in Scan Options page.



# Tips, Comments & Troubleshooting

* The utility can be paused to view the active logs and re-started as needed
* If you encounter “The remote name could not be resolved”, you can follow these commands to get your machine to find the site again:
  + From the Start menu
  + Click Run and type “cmd” (or search for ‘Command Prompt’ from the apps page)
  + In the Command prompt window, type “ipconfig /flushdns” and press enter
  + Then, type “ipconfig /registerdns” and press enter
  + Wait a minute and you should be able to browse and resolve the remote name (web site) correctly if it exists
  + Close the Command prompt window
* If you encounter a path that appears to show duplicate entries in the ‘Absolute URL Tested’ column, such as showing <http://contoso/projects/projects/tasks>, be sure to enter all custom SharePoint Managed Paths in the Settings page. See SharePoint Essentials Toolkit User Manual for more detailed information.
* If you encounter an error with “Details: The remote server returned an error: (404) Not Found.”, this can be caused by not setting the SharePoint Managed Paths in Settings page, or this may be related to an IIS setting that is complaining the content/url or query string is too large to be processed. You may need to increase the ‘MaxAllowedContentLength’ in IIS (Internet Information Services).

**References:**

<http://www.iis.net/learn/manage/configuring-security/use-request-filtering>

<http://www.iis.net/learn/troubleshoot/using-failed-request-tracing/troubleshooting-failed-requests-using-tracing-in-iis>

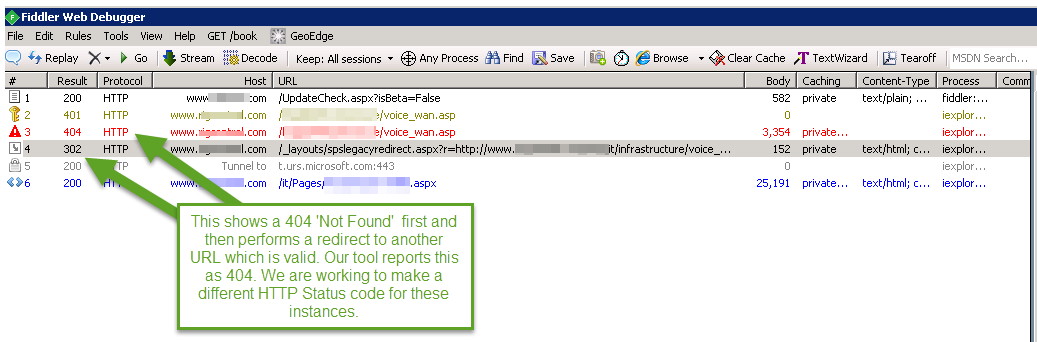
<http://blogs.msdn.com/b/sudeepg/archive/2009/11/10/content-deployment-the-remote-server-returned-an-error-404-not-found.aspx>

* Users with READ rights get an error when trying to retrieve sub site data:

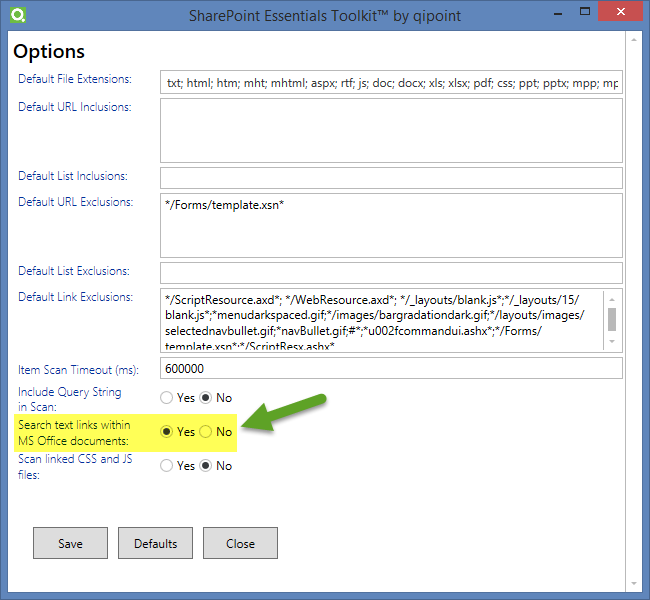
“There was a problem enumerating sub sites. Details: Access denied. You do not have permission to perform this action or access this resource.”

Users require “Browse Directories - Enumerate files and folders in a Web site using SharePoint Designer and Web DAV interfaces.” rights. You can add this permission from Permission Levels in Site Settings -> Site Permissions.

* False negatives: If you have links that are reported as broken (Http Status Code 404) but when you open the site in Internet Explorer or other web browser and the web page loads correctly, check the network traffic using tools such as Fiddler, we have found some pages may be incorrectly being resolved and redirected and actually reporting a 404 first and then redirects to a valid page. This may be an issue with DNS or the Http redirect code being used. We are working on providing a more descriptive status code for these situations to help identify these issues of page URL resolution/redirects. See below for example of this using Fiddler:



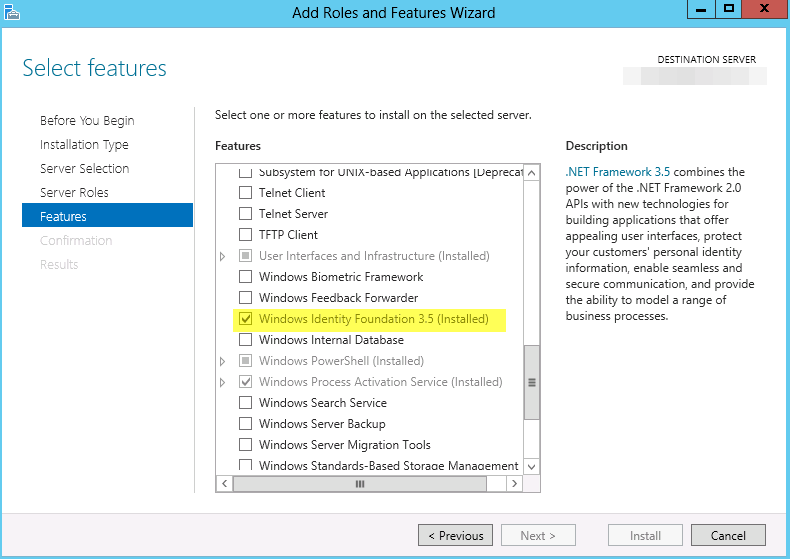
* If you have links in MS Office documents that you believe are not being scanned, be sure to enable ‘Search text links within MS Office documents’ is ON.



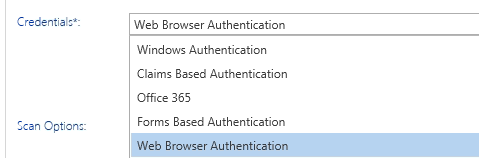
If this is not set to ‘Yes’, only ‘Hyperlink’ links will get scanned by the tool. A ‘Hyperlink’ link can be identified by hovering your mouse over the link, it should show you the URL in a Tool Tip and a ‘Ctrl+Click to follow link” above it.

* For long running scans with more than 50,000 items, it is recommended to turn of the detailed logging to improve performance, this can be found at the bottom of the Scan for Broken Links page. Also please review the Minimum Hardware Requirements in this document.
* If you encounter “Could not load file or assembly 'Microsoft.IdentityModel, Version=3.5.0.0, Culture=neutral, PublicKeyToken=31bf3856ad364e35' or one of its dependencies. The system cannot find the file specified.” a .NET framework component could be missing. Please download and install the Windows Identity Foundation component at <http://www.microsoft.com/en-us/download/details.aspx?id=17331>

If using Windows Server 2012, you can install the Windows Identity Framework using PowerShell (Add-WindowsFeature windows-identity-foundation), or Server Manager UI.



* If you receive errors such as “Forbidden” or “The communication object, System.ServiceModel.Channels.ServiceChannel, cannot be used for communication because it is in the Faulted state”. To resolve this issue, if using Office 365, use the tenant admin user email to login OR use Web Browser Authentication from drop down.



* If you receive errors of missing features when running a scan, or ‘<columnname> not found’, this may be due to faulty features or artifacts installed on the farm/site. The tool tries to retrieve information from a list/library and items in order to parse the contents within them. When a feature is missing or corrupted, the tool may report an error. This response is from the SharePoint API. To resolve this, find the missing/corrupt features and reinstall them or remove them from the site/farm.
* Links that require authentication - 401 Unauthorized – currently the utility uses the same authentication as the credentials added to scan the SharePoint site. So if the link that is scanned requires a different authentication/user name and password, the utility will give a 401 unauthorized status message. You may add exclusions to jobs to exclude sites from appearing as ‘Unauthorized’, however at this time the tool only supports one authentication set for all URLs.
* If scanning a site that is not from SharePoint or is not in the same web application, but it uses the same authentication provider and credentials as current user specified in tool, it may show as a Redirect HTTP Status Code, this is because a redirect occurs to authenticate the user by passing security tokens across http, treat these ‘Redirects’ as valid links, otherwise the HTTP Status Code would be different such as 404 unauthorized or 404 File not found.
* If you encounter something like “"Error parsing webpart 'Get started with your site' for page 'http://contoso/…/SitePages/Home.aspx'. Details: Exception of type 'Microsoft.SharePoint.SoapServer.SoapServerException' was thrown". Be sure to have set the SharePoint Managed Path(s) in the Settings page.

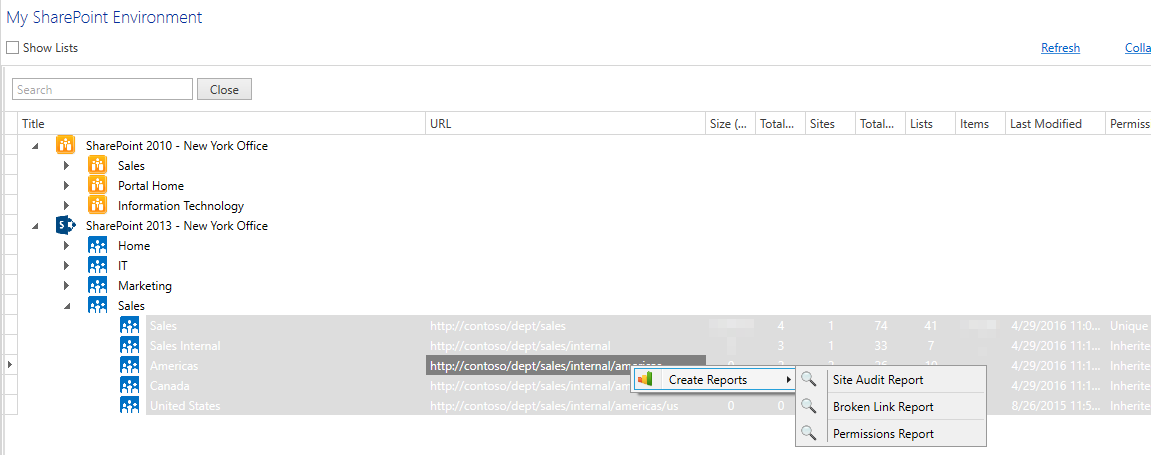
## Improving time to complete Large Scan Jobs

The scan job takes forever to complete! How can I make it finish faster?

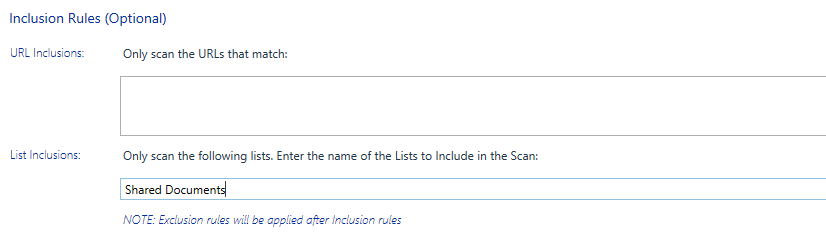
You can improve scan job times by following the tips below:

1. We recommend to break large site collections, or sites with large lists into multiple jobs.

Hold SHIFT or CTRL and select all sites (you want to build a report on) from the Home Page Dashboard, right click, and click “Create Reports -> Broken Links” – this will allow you create multiple jobs (multi-threaded) at once.



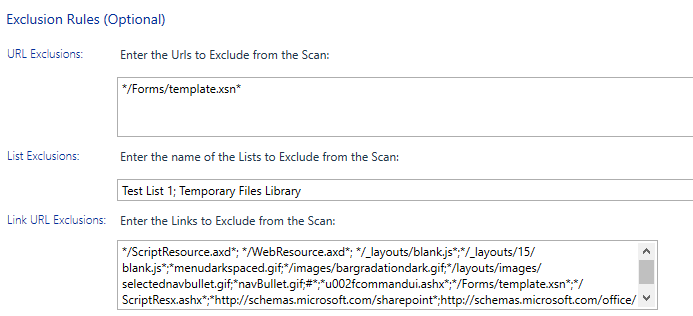
1. If you have a single site with many lists or one or more large lists (greater than 10K items), do not run this site in a single job; Instead, execute or schedule this job and use “Inclusion Rules -> List Inclusions” to separate each large list into a single job; or multiple lists if you have many, into a few jobs - such as a site with 50 lists, separate into 10 per job. Again, you can create a Job Template (click “Save Job”) to make this easier.



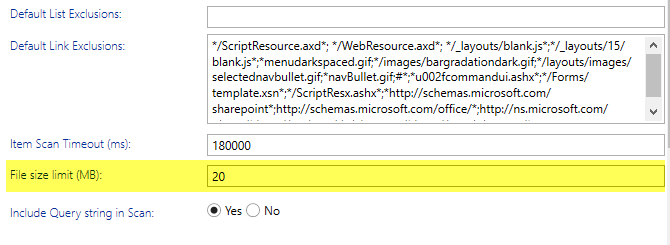
Then, you can then schedule these jobs to run in parallel (at the same Start Time).

***NOTE:*** *This may require you to have more CPU or RAM or to run on a different machine to get the best results.*

1. Exclude content you do not need in the report. Set the “Exclusion Rules -> List Exclusions” to skip lists that are not needed in the report. Add any SharePoint URLs (such as folders or sections of a site) and Link URLs (links found within pages and files – such as navigation menu items) that you do not need in the report, or that you know are valid.



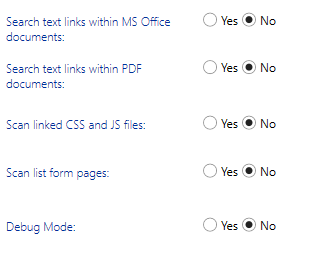
1. Disable “Show Detailed Logging”. Detailed Logging should be used when you are first setting up jobs to ensure they are working correctly and capturing the URLs and links you expect. This option is also useful when troubleshooting issues. Once jobs are tested and configured to run as you need them, uncheck “Show Detailed Logging” to minimize the report time spent on writing each link found and other details.
2. If you are using “Scan List Attachment File Contents” or “Scan Document File Contents” - Specify a lower “File Size Limit (MB)” in the Scan Options page (“Scan Links (tab) -> File Size Limit (MB)”) – such as 10MB. This setting is used to skip large files when scanning file contents. The default is set to 20MB.



1. Turn off: Scanning MS Office text links, PDF text links, Linked CSS and JS files, Debug Mode:

Go to the "Scan Links (tab) -> Options" page, make sure the following are turned off these 'optional' settings: (they are OFF by default)

"Search text links within MS Office documents"  
"Search text links within PDF documents"  
"Scan linked CSS and JS files"  
"Scan list form pages"  
"Debug Mode"



1. Ensure the folders below are not being scanned (excluded) by anti-virus ‘on-access’ scanning and also excluded from any offline syncing such as OneDrive, Google Drive or Drop Box.

(also exclude subfolders of directories below)

* C:\Users\UserName\AppData\Local\Temp\QiPoint
* C:\ProgramData\QiPoint
* <My Documents>\QiPoint

# Technical Support

If you need technical assistance, not to worry! We offer several ways to get in touch with our support team.

**Email:** [support@qipoint.com](mailto:support@qipoint.com) (Average response time is 3-8hrs)

**Phone:** 917-633-5998 opt. 1

**Online Support Ticket System:** <http://support@qipoint.com>

We are here to help!

# Product Features

|  |  |
| --- | --- |
| Feature | Enterprise |
| Scan SharePoint 2010 |  |
| Scan SharePoint 2013 |  |
| Scan SharePoint 2016 |  |
| Scan Office 365 |  |
| Scan SharePoint Sites |  |
| Scan Web Page meta data |  |
| Scan Specific Site for an individual Report |  |
| Searching within Report |  |
| Querying results in the Report |  |
| Report in customizable grid |  |
| Scan Page Libraries and ASPX Page Content |  |
| Scan Content Editor Web Parts |  |
| Scan Navigational links in pages |  |
| Scan Custom Navigation in pages |  |
| Scan Custom Web Parts in pages |  |
| Scan List Item meta data |  |
| Scan Document meta data |  |
| Export to CSV / MS Excel |  |
| Grouping of Scan Results in grid |  |
| Detailed Report on broken links |  |
| Scan Entire Site Collection in single job |  |
| Exclusion Rules (SharePoint URLs) |  |
| Exclusion Rules (SharePoint Lists) |  |
| Exclusion Rules (URLs within items/pages/files) |  |
| Inclusion Rules (SharePoint URLs) |  |
| Inclusion Rules (SharePoint Lists) |  |
| Scan Document Library File contents |  |
| Scan List Item File Attachment contents |  |
| Scan within All ASCII, UTF-8, MIME, TXT File Types |  |
| Scan within MS Word Document Contents |  |
| Scan within MS Excel Document Contents |  |
| Scan within MS Project Document Contents |  |
| Scan within MS OneNote Document Contents |  |
| Scan within PDF Document Contents |  |
| Scan within InfoPath Form Contents |  |
| Email results automatically to users |  |
| Automatic Find \ Replace URLs |  |
| Schedule Scan Jobs & Report Generation |  |
| Export Reports to SharePoint List |  |