|  |
| --- |
|  |
| SharePoint Broken Link Manager 2015 - User Guide |
| User Manual |
|  |
|  |
| **6/27/2016** |

|  |
| --- |
|  |

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

Table of Contents

[Introduction 4](#_Toc454819792)

[Installation & Minimum Requirements 5](#_Toc454819793)

[IMPORTANT: This program does not need to be installed on a SharePoint Server. 5](#_Toc454819794)

[Minimum Hardware Recommendations 5](#_Toc454819795)

[Supported SharePoint Versions 5](#_Toc454819796)

[.NET Framework Required 5](#_Toc454819797)

[Local Machine & Windows System Permissions Required 7](#_Toc454819798)

[Anti-Virus, Offline Sync Folders (i.e. Google Drive, OneDrive, Drop Box, etc) and Performance 7](#_Toc454819799)

[SharePoint User Permissions Required 8](#_Toc454819800)

[Find/Replace Permissions Required 8](#_Toc454819801)

[Preserve Author 8](#_Toc454819802)

[Office 365 Accounts 8](#_Toc454819803)

[File Content Scan - Notes on Behavior (Enterprise Edition) 9](#_Toc454819804)

[Broken Link Manager Options 10](#_Toc454819805)

[What will be scanned 15](#_Toc454819806)

[Outbound Links 15](#_Toc454819807)

[Supported Link Types 15](#_Toc454819808)

[SharePoint Web pages 15](#_Toc454819809)

[Document ID Links 15](#_Toc454819810)

[Web Parts 17](#_Toc454819811)

[Lists & libraries 17](#_Toc454819812)

[File Contents (Enterprise version only) 17](#_Toc454819813)

[Linked CSS & JavaScript files 18](#_Toc454819814)

[What will NOT be scanned 19](#_Toc454819815)

[Links 19](#_Toc454819816)

[Lists & libraries 19](#_Toc454819817)

[File Contents (Enterprise version only) 19](#_Toc454819818)

[My Sites / Personal Sites 20](#_Toc454819819)

[Performing a Scan for Broken Links 21](#_Toc454819820)

[Start a Scan 21](#_Toc454819821)

[Scan Options 22](#_Toc454819822)

[*Job Name* 22](#_Toc454819823)

[*Site URL to Scan* 22](#_Toc454819824)

[*Credentials* 22](#_Toc454819825)

[*Use Proxy Server* 22](#_Toc454819826)

[Scan Options 22](#_Toc454819827)

[*File Extensions* 23](#_Toc454819828)

[*Inclusions* 23](#_Toc454819829)

[*Exclusions* 24](#_Toc454819830)

[*Find / Replace* 25](#_Toc454819831)

[*Email* 26](#_Toc454819832)

[Export 27](#_Toc454819833)

[What is a Broken Link? 29](#_Toc454819834)

[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. 29](#_Toc454819835)

[Manually Finding and Replacing Links 29](#_Toc454819836)

[Automatically Finding and Replacing Links (Enterprise version only) 30](#_Toc454819837)

[What will be replaced 30](#_Toc454819838)

[Supported Link Types 31](#_Toc454819839)

[What will NOT be replaced 31](#_Toc454819840)

[Versioning 32](#_Toc454819841)

[Check In & Check Out 32](#_Toc454819842)

[Content Approval & Find / Replace 33](#_Toc454819843)

[Bulk Import Find Replace Links 33](#_Toc454819844)

[Find / Replace Options 34](#_Toc454819845)

[Link Report - Finding URLs that were automatically replaced/fixed 35](#_Toc454819846)

[Scheduling Scan Jobs (Enterprise version only) 36](#_Toc454819847)

[Schedule a Job 36](#_Toc454819848)

[Viewing & Managing Scheduled Jobs 37](#_Toc454819849)

[Batch Processing Multiple Site Collections 38](#_Toc454819850)

[Licensing 39](#_Toc454819851)

[Components 39](#_Toc454819852)

[Changing the license type of a component 39](#_Toc454819853)

[Trial Version 39](#_Toc454819854)

[Free Version 39](#_Toc454819855)

[Obtaining a License 39](#_Toc454819856)

[Transferring a License 40](#_Toc454819857)

[Reports 41](#_Toc454819858)

[Opening the Report 41](#_Toc454819859)

[Grid View Report 41](#_Toc454819860)

[Customizing the Report 41](#_Toc454819861)

[Filtering 41](#_Toc454819862)

[Advanced Filtering 41](#_Toc454819863)

[Searching within the Report 42](#_Toc454819864)

[Filtering Queries 42](#_Toc454819865)

[Grouping 43](#_Toc454819866)

[Views 43](#_Toc454819867)

[Export 44](#_Toc454819868)

[How to Interpret the Report 45](#_Toc454819869)

[*Example* 45](#_Toc454819870)

[Which Links Are Broken? 45](#_Toc454819871)

[Report Fields 46](#_Toc454819872)

[Report Archive 48](#_Toc454819873)

[Best Practices: Improving performance and Large Scan Jobs 49](#_Toc454819874)

[Tips, Comments & Troubleshooting 51](#_Toc454819875)

[Improving time to complete Large Scan Jobs 54](#_Toc454819876)

[Technical Support 58](#_Toc454819877)

[Version Comparison 59](#_Toc454819878)

# 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 one time.

**RAM:** Minimum 4GB. Recommended 8GB or higher  
We recommend at least 4GB of available RAM for sites less than 200GB and less than 10K items to scan, 6GB RAM free for sites with 200GB to 500GB or with 10-100K items to scan, 8+GB available RAM for sites with 500GB+ or with 100K items or more to scan. NOTE: You can reduce the RAM required by breaking down URL scan jobs into separate smaller jobs, such as create a single job per List/Library to run at different time intervals (rather than all Lists/Libraries at once in a single job which would require more memory/RAM utilization).

**Hard Disk:** 500MB 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 Office 365  
Microsoft SharePoint Server 2016  
Microsoft SharePoint Foundation 2013  
Microsoft SharePoint Server 2013  
Microsoft SharePoint Foundation 2010  
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.

## 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.

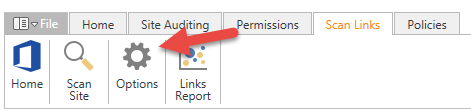
# File Content Scan - Notes on Behavior (Enterprise Edition)

* This utility (only Enterprise Edition) can scan file contents. Files to be scanned are downloaded to a temporary folder on the client machine (C:\Users\UserName\AppData\Local\Temp\QIPoint) and cleaned up after the item file scan completes. File names are changed in this temporary location. We used this physical file location instead of in-memory loading of files in order to minimize the applications memory footprint.
* Sites with very large files may take more time to scan and adequate space must be available on the client machine.
* This temp folder is automatically cleaned up if size reaches 10MB+ (it tests this condition for every file it scans) and clears this directory when the scan is complete. The actual documents living in SharePoint are never modified (unless you are using the Find/Replace mechanism described in a later section).
* This temp folder is also checked for cached files when the program exists, and they are permanently deleted if any are found to ensure disk space is not consumed unnecessarily.

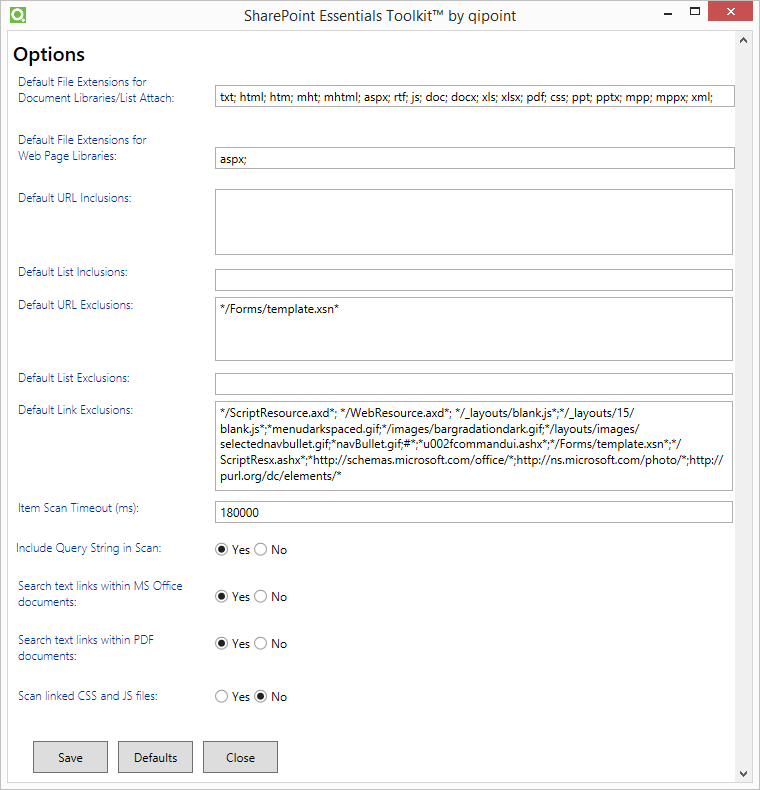
# Broken Link Manager Options

You can change the options for the Broken Link Manager component by following the instructions below:

1. From the ribbon, click on the Scan Links tab
2. Click the “Options” button



1. The following dialog 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.

* 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 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. Normally you should not have to change this value unless you see frequent timeouts in the logs.

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.

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.

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.

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 may 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.

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.

# What will be scanned

This tool will scan:

## 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, OOB navigation, custom navigation, and custom controls with links
  + 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.

## Document ID Links

Links to files using the SharePoint Document ID service will be scanned and shown as ‘200’ or ‘Redirect’. However, the Document ID Service does not send back a code that relates to an HTTP Status Code, so even if the ID is not found by the service, our tool still reports a HTTP Status Code of ‘200’ ‘OK’ or ‘redirect’. This response comes from the SharePoint service as a valid response since the web service is answering the request correctly, for example, the page will return with a message “No documents with the ID HCH2YEEPCJYU-5-45 were found in this site collection.” Is considered a valid response, not a 404 not found since the SharePoint Document ID Service looks up the file and responds with a web page.

## Web Parts

The following web parts will be scanned for broken links:

* Script Editor Web Parts
* Xslt List View Web Parts
* Image 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.

## Lists & libraries

* + Includes all column meta-data that may contain links (such as Text, Multi-line Text, Hyper-Link data types)

## File Contents (Enterprise version only)

* + 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 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
  + Scans PDF documents (except for PDF files that have been scanned as images)
  + MS Office Text Links (non-hyperlinks) - There is an option in the Scan Links (tab) -> Options -> “Include text links within MS Office documents”. Setting this option to ON will increase resource utilization and time to run a scan.
    - * This option will parse all text within MS Office 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)

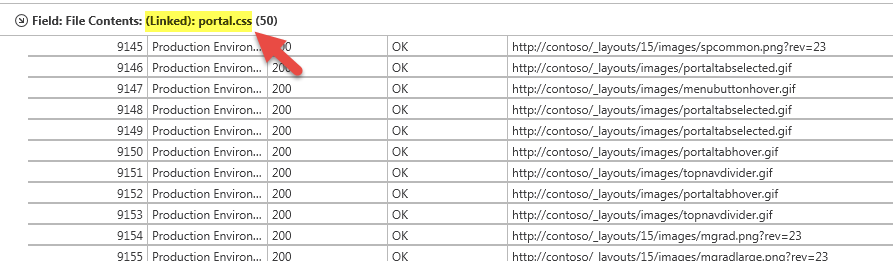
* + - * These are different than ‘real’ ‘Hyperlink URLs’ (these show a hyperlink tooltip when mouse is hovered over them) which are retrieved programmatically via MS Office API (programming interface).
      * Default is OFF
  + PDF Text Links (non-hyperlinks) - There is an option in the Scan Links (tab) -> 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
  + 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
  + InfoPath Form Libraries are not currently supported; this may be added in a future release. However, there is a built-in SharePoint feature to “Re-Link” forms to the template.
  + External Lists are not supported.

## File Contents (Enterprise version only)

* + We cannot list all possible formats out there, but here are some common file types that we know of that are not currently supported by our tool to scan for links or find replace:

NOTE: This relates to scanning within these files, metadata such as SharePoint fields or columns are still scanned and supported for file types below.

* + - One Note file contents
    - 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.
    - Image file contents

## My Sites / Personal Sites

* + Links to My Sites and Personal sites are supported, however scanning links within a personal site (pages or files) is not currently supported.
  + This is due to the limitation of the SharePoint CSOM/API (programming interface) and web services that can fetch this information. Once the SharePoint CSOM/API supports this, we will implement this functionality in the tool.

# Performing a Scan for Broken Links

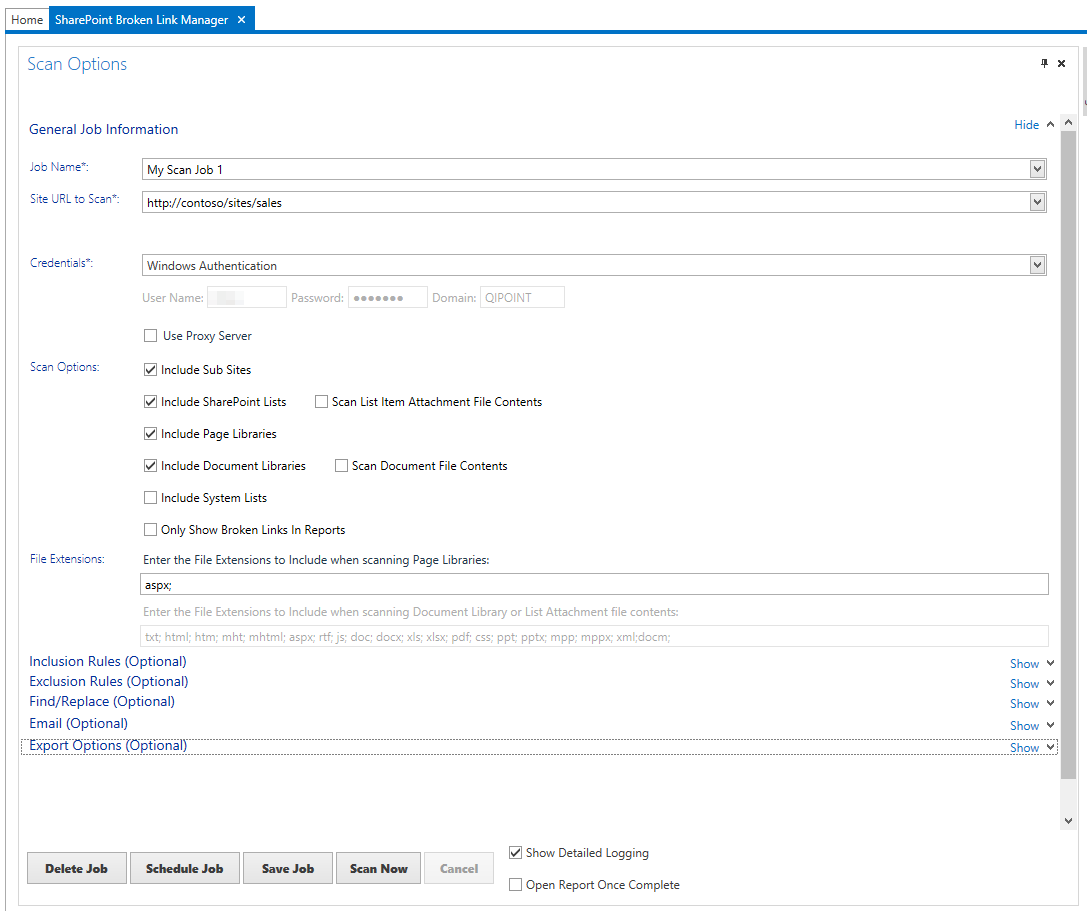
## Start a Scan

1. Open “SharePoint Essentials ToolkitTM”
2. Click on the “Scan Links” tab at the top
3. Click on the “Scan Site” button
4. A new tab will appear called “SharePoint Broken Link Manager” (below screenshot)

(NOTE: This page below may appear different if running Lite or Professional versions)

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

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



1. Click “Scan Now” button to start the scan

## 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

### 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 or a sub site.

**Example:** <https://companyportal/sites/hr>

### Credentials

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

**Windows Authentication** – (Default) Uses Windows Authentication to connect to the SharePoint site. This will work in most cases when SharePoint is hosted on your internal company network. It will use the current user’s login credentials.

**Windows Authentication (Custom)** – Uses Windows 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.

**Claims Based Authentication** – Connects to the SharePoint site using Claims Based authentication. This is the default authentication method for SharePoint 2013. If you are using Azure to host SharePoint, you can use this authentication method.

**Office 365** – Use this option to connect to Office 365 hosted SharePoint sites. This applies to both standard setups and federated setups.

Forms Based Authentication – Use this method if the SharePoint site you are connecting to uses Forms based authentication.

**Web Based Authentication** – Use this option to force the program to provide a pop up window prompt that will display your company login page to provide credentials. If you are using Office 365 with an ADFS server with custom login page, you can use this authentication type.

### Use Proxy Server

Select this option if you use a Proxy Server to access the SharePoint site you are scanning.

### Scan Options

**Include Sub Sites** – Scans all links within this site’s sub sites (of the given URL).

**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”.

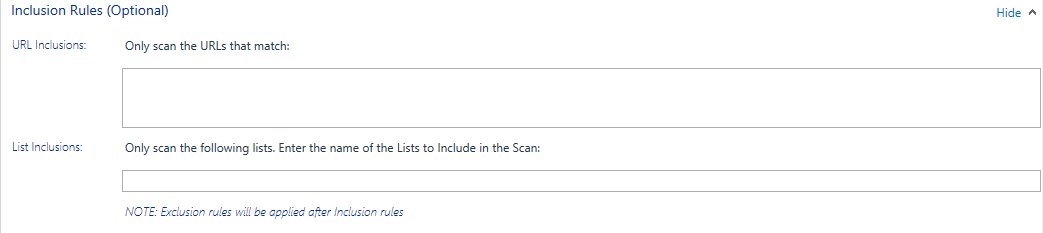
**Include System Lists** – Scans all links within Hidden System lists and libraries.

**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.

### File Extensions

If the user has selected the option to scan file contents, they can specify the extensions to include in the scan here. 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). 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



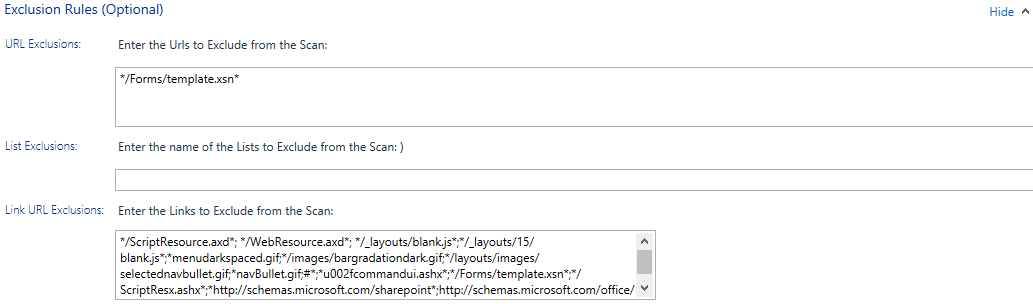
**URL Inclusions –** Enter the URLs of the SharePoint target sites, lists, libraries, items that you want to include in the scan. You can enter an exact URL to match or partial using wildcards (\*). For example: entering [https://sharepoint/sites/hr/ap/\*](https://sharepoint/sites/hr/ap/*) here will 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** - The user can specify what Lists to include in the scan by entering the names lists in this section. 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 ‘;’.

**NOTES:**

* **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 List Settings page when using URL Inclusions/Exclusions for lists, such as** <http://contoso/Shared%20Documents>**.**

### Exclusions

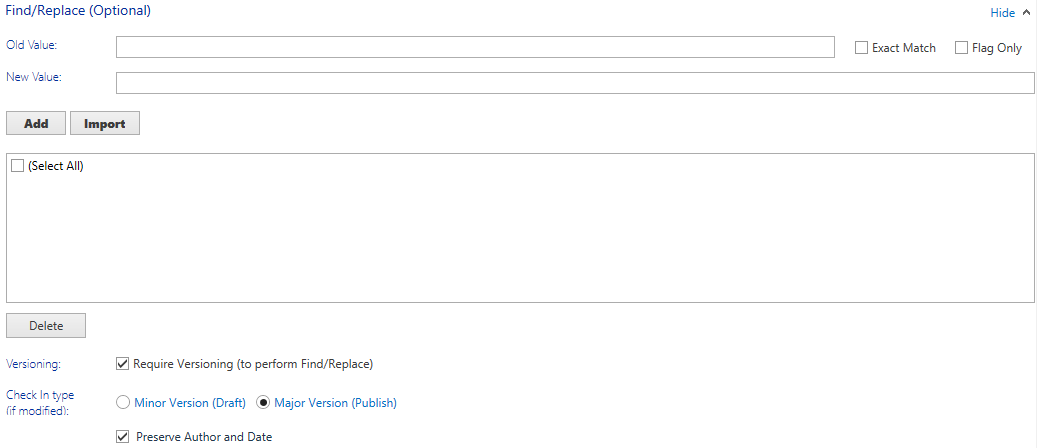


**URL Exclusions -** Enter the URLs of the SharePoint target sites, lists, libraries, items that you want to exclude in the scan. You can enter an exact URL to match or partial using wildcards (\*). For example: entering [https://sharepoint/sites/hr/ap/\*](https://sharepoint/sites/hr/ap/*) here will 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 ‘;’.

### Find / Replace



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 (Enterprise Version only), 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)

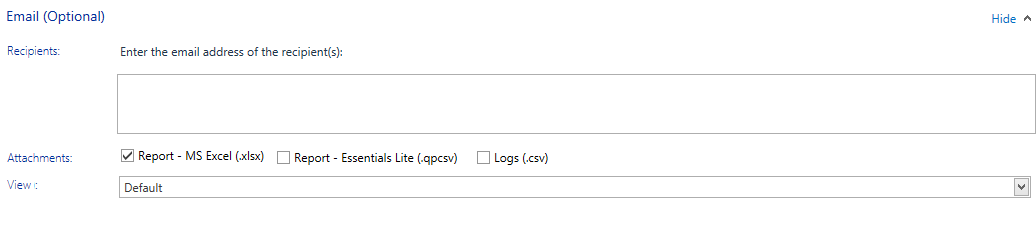
**Add** – Click this button once you have added an ‘Old Value’ to replace and a ‘New Value’ or selected the ‘Flag Only’ option

**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’.

**Check In type** – (Only used if Versioning is turned on). Once a Find / Replace match is found, the utility will automatically check out and check in the file. This option specifies the type of check in that will occur. 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.

**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.

### Email



Enter the email addresses for the users who you would like the report to be emailed to. Separate multiple email addresses with a semi-colon ‘;’.

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 for more information).

### Export



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.

**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”

**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 Multiple Site Collections***): 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, 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.

# 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 (Enterprise version only)

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’.

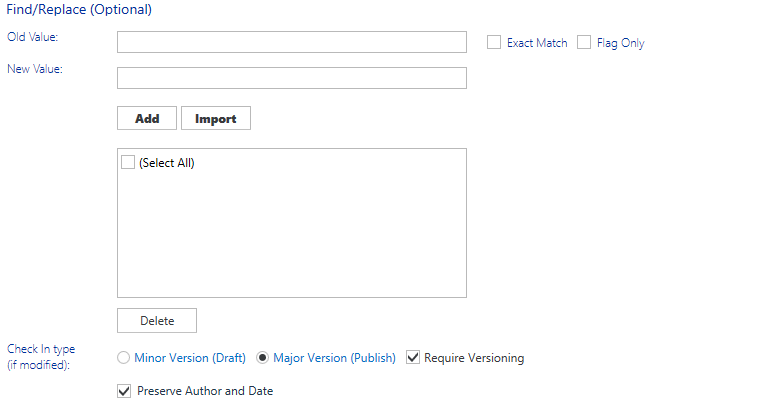
For example: If you enter

**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:



## What will be replaced

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 Project 97 and later
  + Txt, RTF, HTML, and other ASCII encoded file types
* 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
  + 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
  + Tooltips that contain hyperlinks
  + Descriptions or Alt URLs within hyperlinks

## What will NOT be replaced

* Text based links in 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.
* URLs in My Sites will not be scanned or replaced by this tool. Currently My Sites are not 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 InfoPath forms are not supported
* 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.

## 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.

## 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.**

Preserve Author: 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.

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, doing this automatically puts a date time stamp and sets Last Modified Info as current user, SharePoint object model does not allow the tool to change Check In as a different user.

## 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

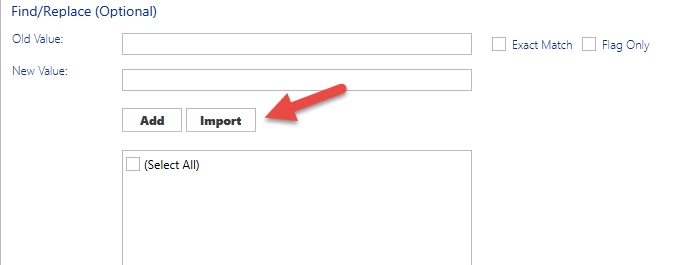
You may import from a MS Excel file a list of URL replacements to use. This can be handy if you have multiple sites and want to have Site Owners send you MS Excel files with fixes to URLs.

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.



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, select this checkbox to preserve the ‘Last Modified’ author and date for the item\item.

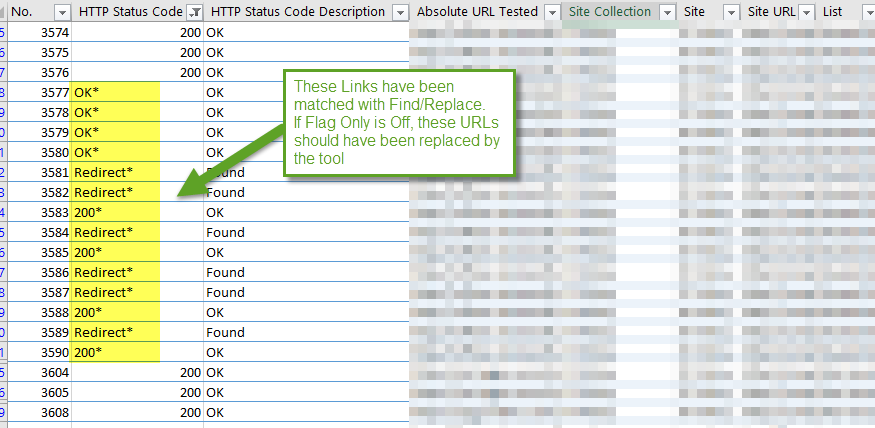
NOTES: The user account used to run the Job must have Full Control on the List/Item in order for “Preserve Author” to work.

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.

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

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.

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\*



# Scheduling Scan Jobs (Enterprise version only)

This product includes a built in scheduler (Enterprise version only) 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).

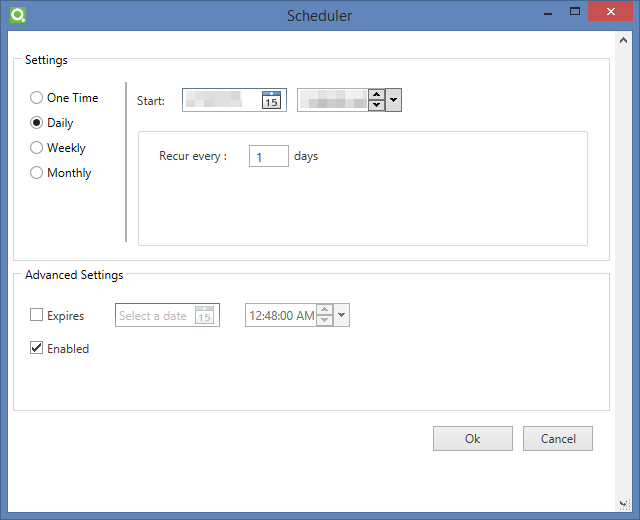
**TIP:** Ensure the “QIPoint Essentials Service” (Windows Service) is running before trying to schedule a job. Ensure the user account used to run the service has ‘Modify’ permission on folder C:\ProgramData\QIPoint (NOTE: This is a hidden path; you can paste C:\ProgramData into the address bar in Windows Explorer to access it) this is where the scheduler job information is stored. By default, all users should have read and edit rights as this is set when the program is installed.

## 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.



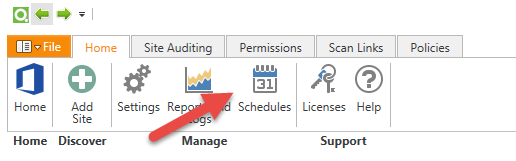
NOTE: You must use “Windows Authentication (Custom)” instead of “Windows Authentication” in order to save the credentials with the scheduled job (as the scheduled job will run as a service). Passwords are encrypted.

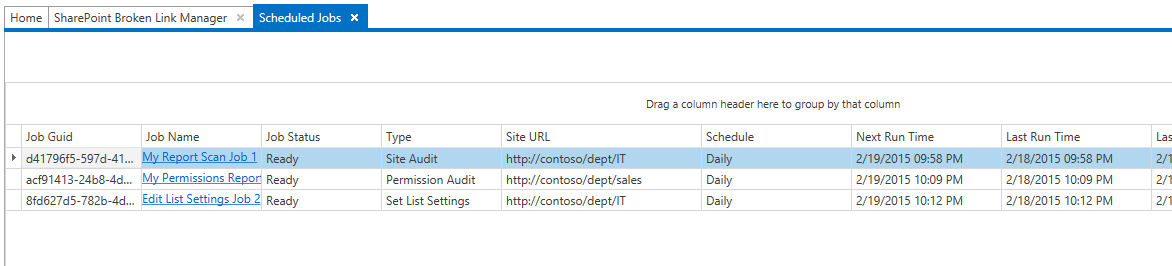


Select the schedule and specific options and click OK to save the scheduled job.

## Viewing & Managing Scheduled Jobs

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





From here you can view reports, logs, execute the job, modify or delete scheduled jobs.

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).

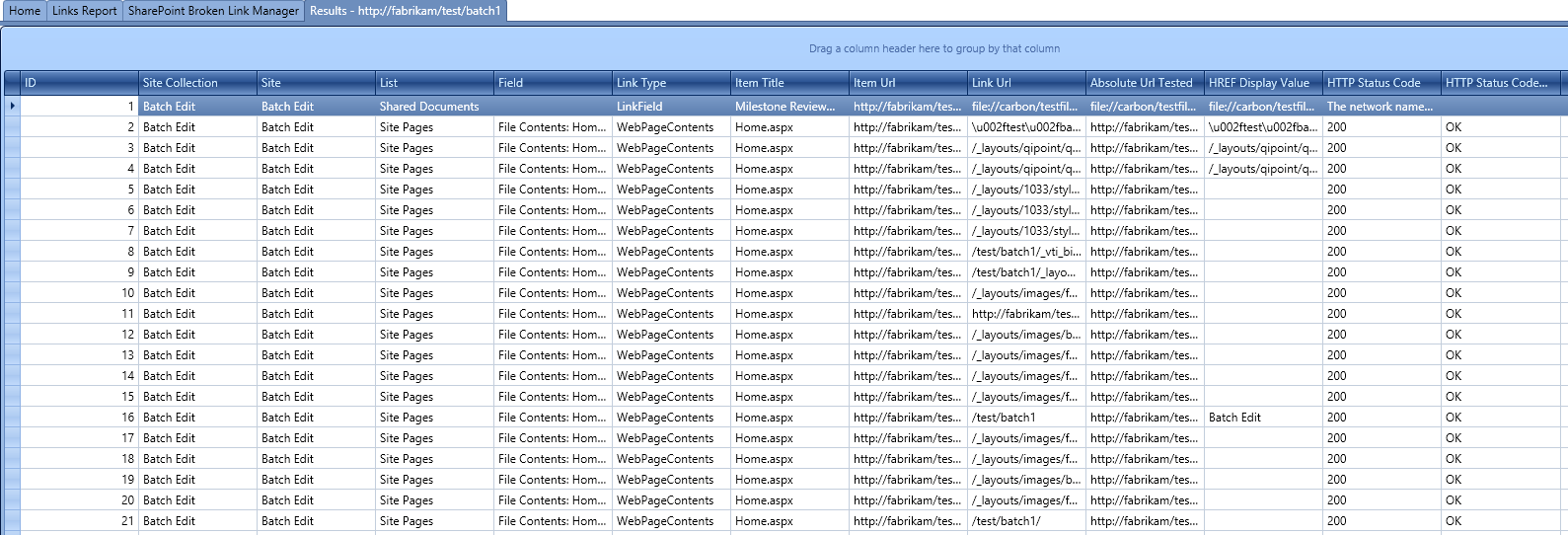
# Reports

## 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 hyperlink 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 CSV/PDF/Excel format
  + From the home page, you can right click on a site that you wish to preview, and click “View->Broken Link Report”
  + To view all reports (in CSV format): from the ribbon, click on the “Home” tab, then click on “Reports and Logs”, windows explorer will open with the CSV reports and log files
    - Alternatively, you can manually navigate to this path. Open Windows Explorer and navigate to the path where the reports are stored (check “Settings”), the default location is in the user’s My Documents folder under “QIPoint\Reports”

## Grid View Report



## Customizing the Report

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

### Filtering

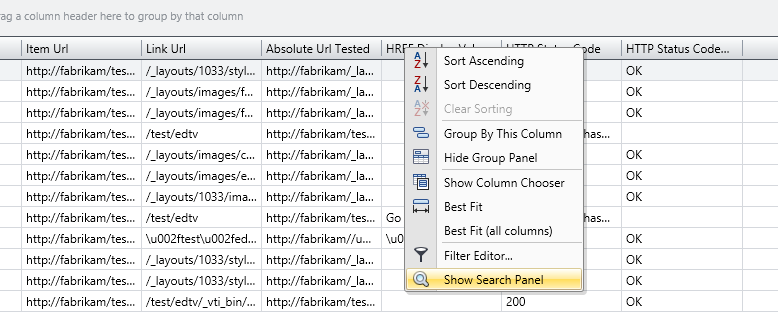
To filter the data based on a field, put your mouse over the column header that you want to filter by. You should see a filter icon, select it and choose the values you only want to have shown in the grid

### 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

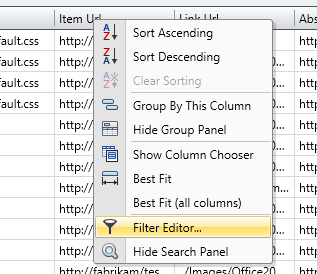
### Searching within the Report

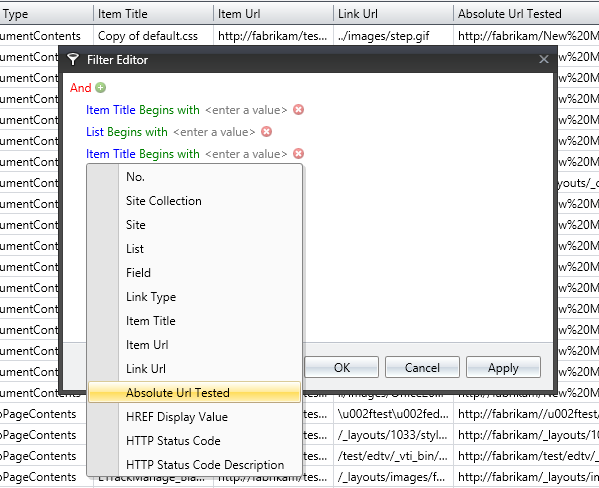
Right Click on any column header and select “Show Search Panel”. The search panel will appear and you can enter any value to search the grid, results will be highlighted



### Filtering Queries

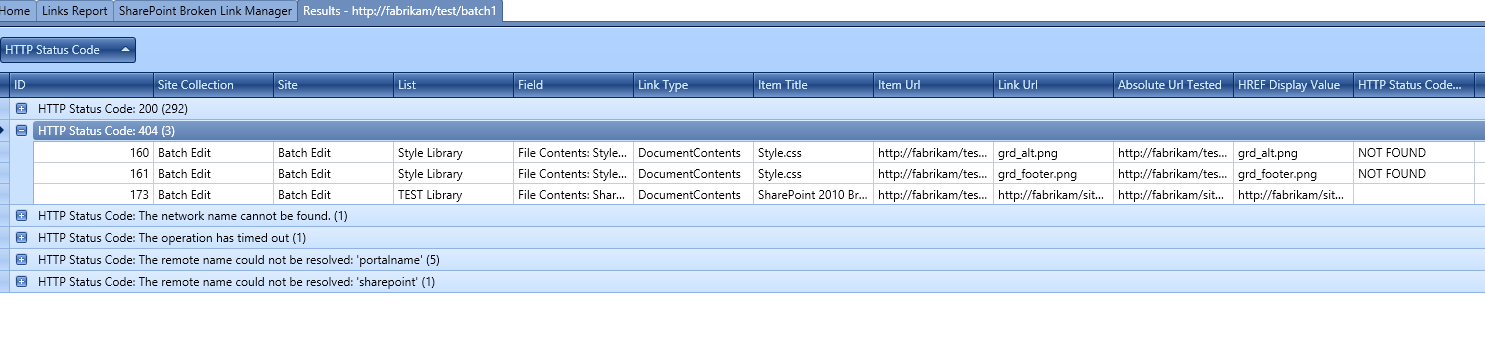
You may also filter the report by using queries. Right click the column header and select “Filter Editor” to add criteria to filter the results





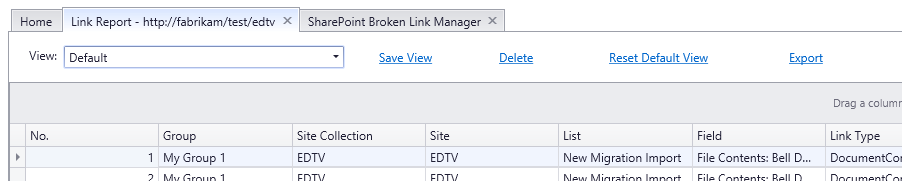
### Grouping

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



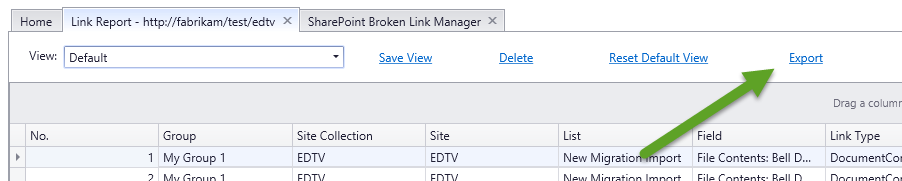
### Views

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 CSV/MS Excel 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 or MS Word.



## 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’ below, they provide information about each link found.

## 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).

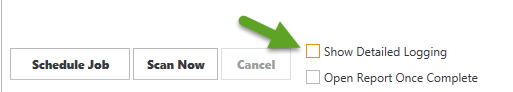
## Report Archive

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.

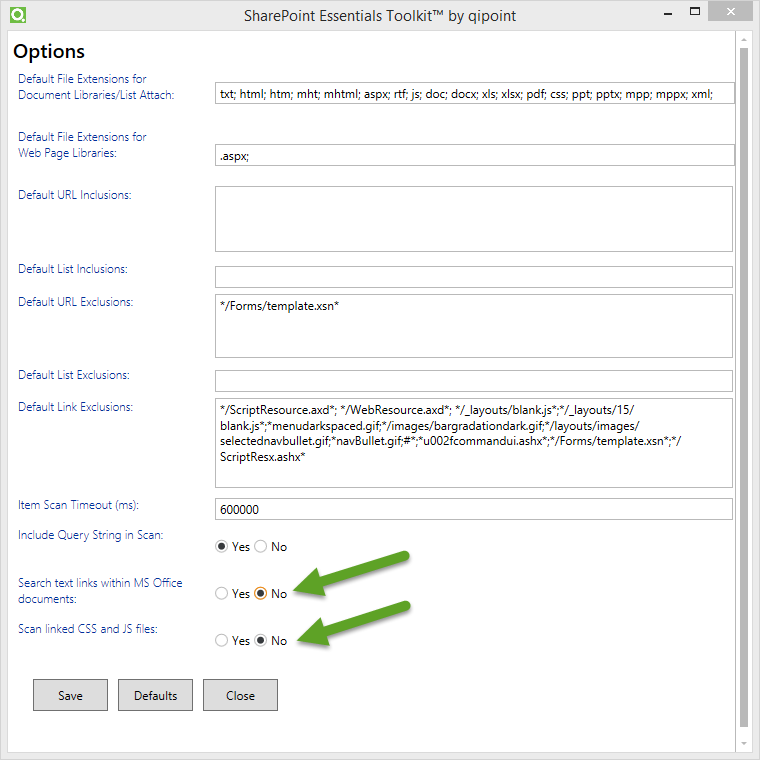
# Best Practices: Improving performance and Large Scan Jobs

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>

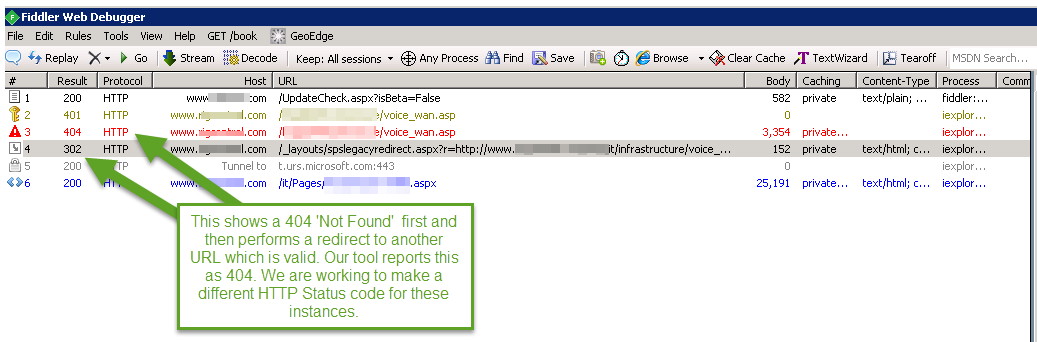
<http://blogs.msdn.com/b/sudeepg/archive/2009/11/10/content-deployment-the-remote-server-returned-an-error-404-not-found.aspx>

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

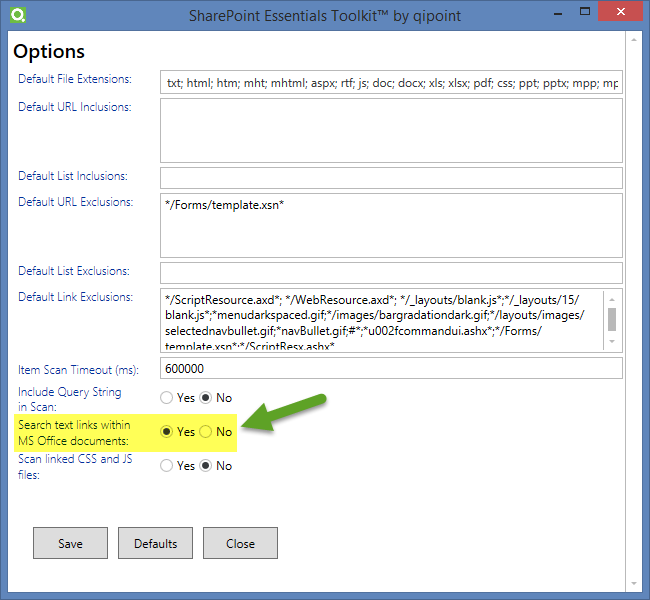
“There was a problem enumerating sub sites. Details: Access denied. You do not have permission to perform this action or access this resource.”

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:



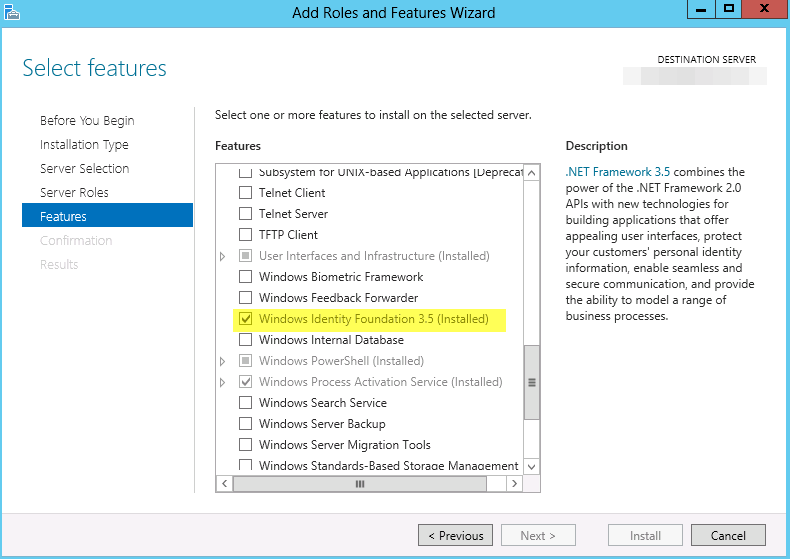
* If you have links in MS Office documents that you believe are not being scanned, be sure to enable ‘Search text links within MS Office documents’ is ON.



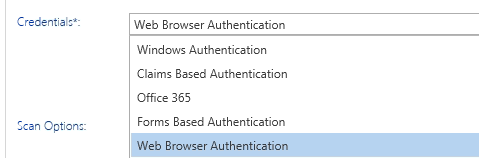
If this is not set to ‘Yes’, only ‘Hyperlink’ links will get scanned by the tool. A ‘Hyperlink’ link can be identified by hovering your mouse over the link, it should show you the URL in a Tool Tip and a ‘Ctrl+Click to follow link” above it.

* 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.



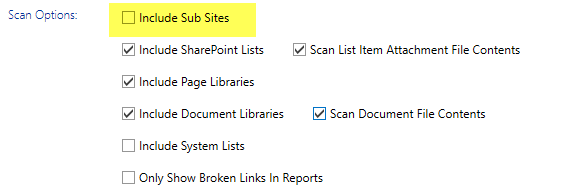
* If you receive errors of missing features when running a scan, or ‘<columnname> not found’, this may be due to faulty features or artifacts installed on the farm/site. The tool tries to retrieve information from a list/library and items in order to parse the contents within them. When a feature is 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

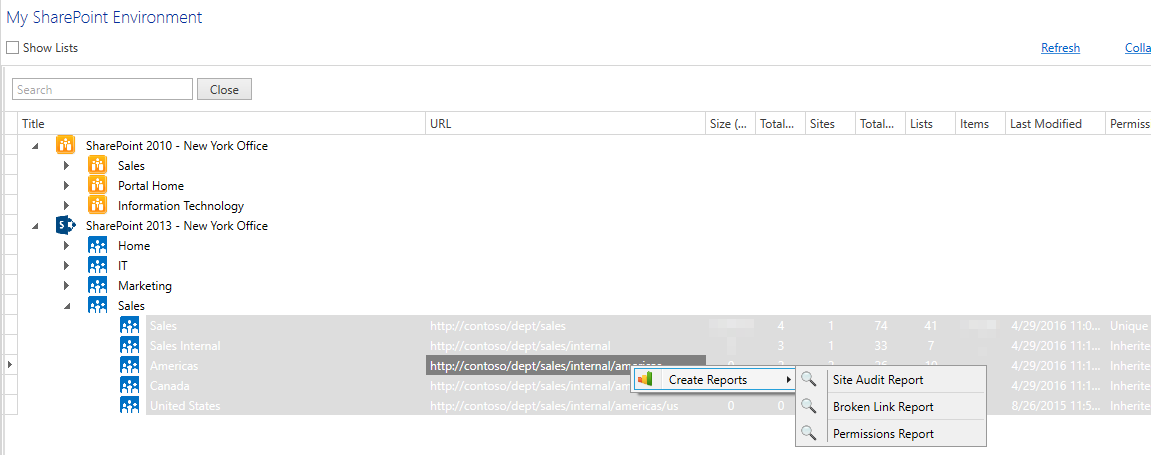
The scan job takes forever to complete! How can I make it finish faster?

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

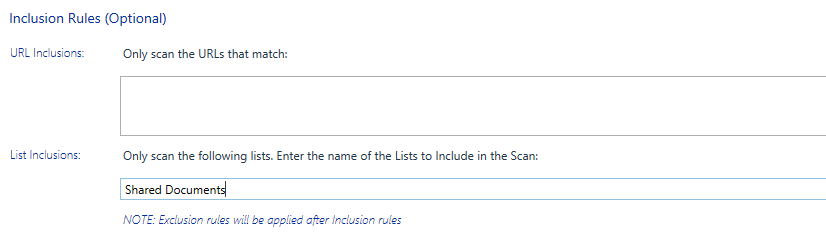
1. Scan Jobs are run in a single thread to ensure logs are captured correctly and aid in validating links and making sure content is scanned correctly. However, if you have a site collection with many sub sites or a single site with many lists or large lists, this may not be the most efficient approach. We recommend to break large site collections, or sites with large lists into multiple jobs.
   1. To do this, uncheck “Include Sub Sites” in the Scan Options page and schedule each site individually.



* 1. You can save the Job Template to make this easier, or you can also 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.



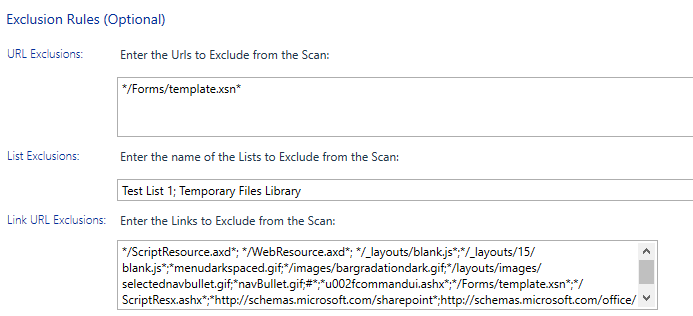
1. If you have a single site with many lists or one or more large lists (greater than 10K items), do not run this site in a single job; Instead, execute or schedule this job and use “Inclusion Rules -> List Inclusions” to separate each large list into a single job; or multiple lists if you have many, 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.



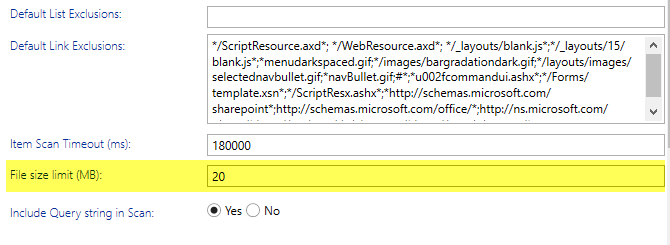
Then, you can then schedule these jobs to run in parallel (at the same Start Time).

***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.



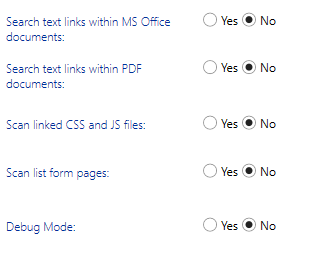
1. Disable “Show Detailed Logging”. Detailed Logging should be used when you are first setting up jobs to ensure they are working correctly and capturing the URLs and links you expect. This option is also useful when troubleshooting issues. Once jobs are tested and configured to run as you need them, 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!

# Version Comparison

|  |  |  |
| --- | --- | --- |
| Feature | Professional | Enterprise |
| Scan SharePoint 2010 |  |  |
| Scan SharePoint 2013 |  |  |
| 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 |  |  |
| Email results automatically to users |  |  |
| Automatic Find \ Replace URLs |  |  |
| Schedule Scan Jobs & Report Generation |  |  |
| Export Reports to SharePoint List |  |  |