ScannerVision User Manual

scannervision™

Table of Contents

1 Welcome

Thank you for choosing ScannerVision as your document capture solution.

ScannerVision is an advanced document processing solution which enables you to scan documents via TWAIN or to load existing documents from a hard drive or network share and then to process and ultimately store them in one or more of the many document stores supported by ScannerVision such as Laserfiche, SharePoint and AivikaOne. Integration with third party document management systems is achieved through connectors which are 32 bit Windows DLLs (Dynamic Link Libraries) or .Net assemblies that are loaded by ScannerVision. ScannerVision provides a fully documented Software Development Kit (SDK) which allows independent software vendors and even customers themselves to develop connectors to interface to any external system.

ScannerVision provides powerful metadata support. Metadata is data that provides information about the documents that are processed by ScannerVision and can come from various sources. Some metadata is generated by the ScannerVision Processing Server as it processes documents. Other metadata can be captured by users as they scan documents at the MFP or at their workstation or it can be obtained from the document itself through barcode reading and/or various OCRing functions such as magnetic ink character recognition (MICR), optical mark recognition (OMR), hand writing recognition (ICR) and full document and zonal optical character recognition (OCR).

ScannerVision allows you to apply conditional logic to determine which metadata to use under certain circumstances and where to route documents. What this means is that ScannerVision can make intelligent decisions on what metadata to use or where documents should be stored based on the conditions you specify.

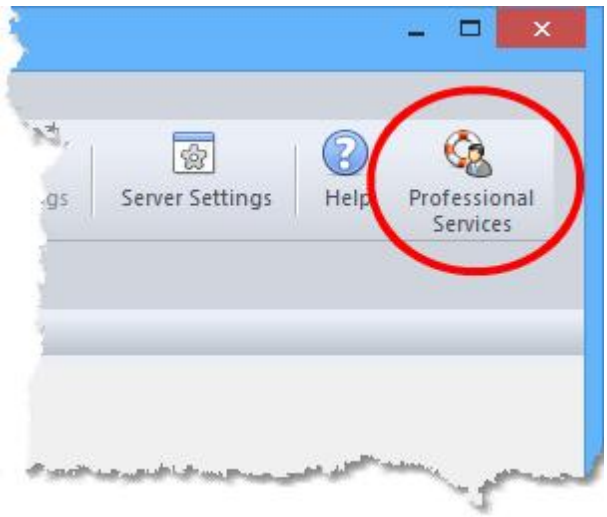
This help manual will guide you through the process of configuring your ScannerVision work flows - called templates. Please read through the manual from start to finish. Some of the concepts will only become clear once you have implemented them in a template. We encourage you to configure test templates as you work through the manual and to try out the many features that ScannerVision offers.

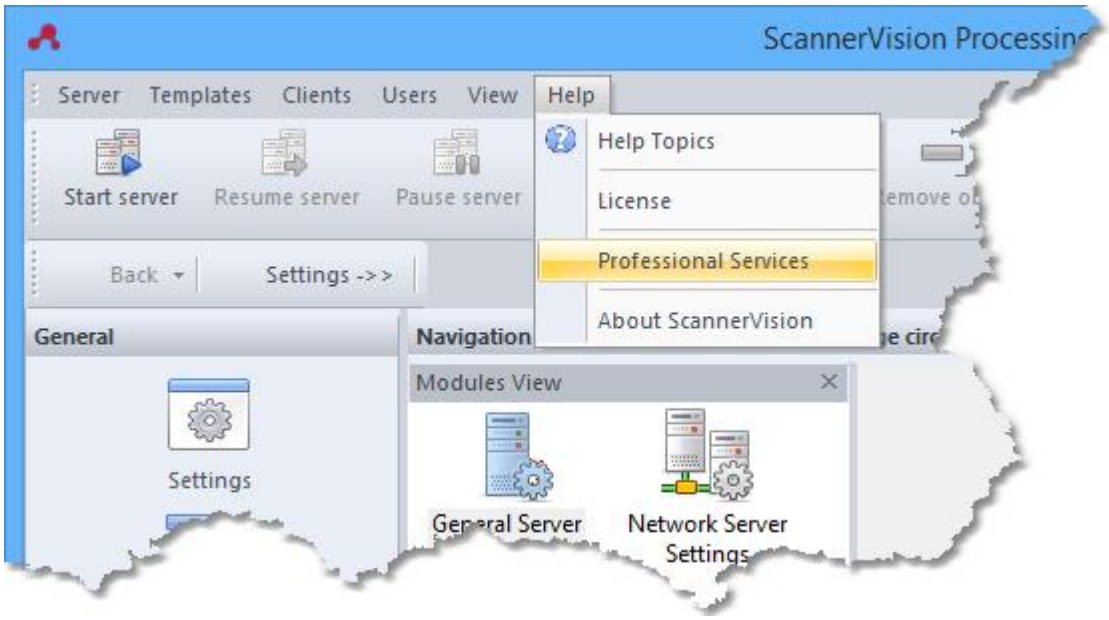
2 Support

If you require any support or information about ScannerVision your first port of call is your reseller. If your reseller is not able to assist you directly, a request for support will be forwarded by them to ScannerVision. If you are not able to contact your reseller you can send an email to

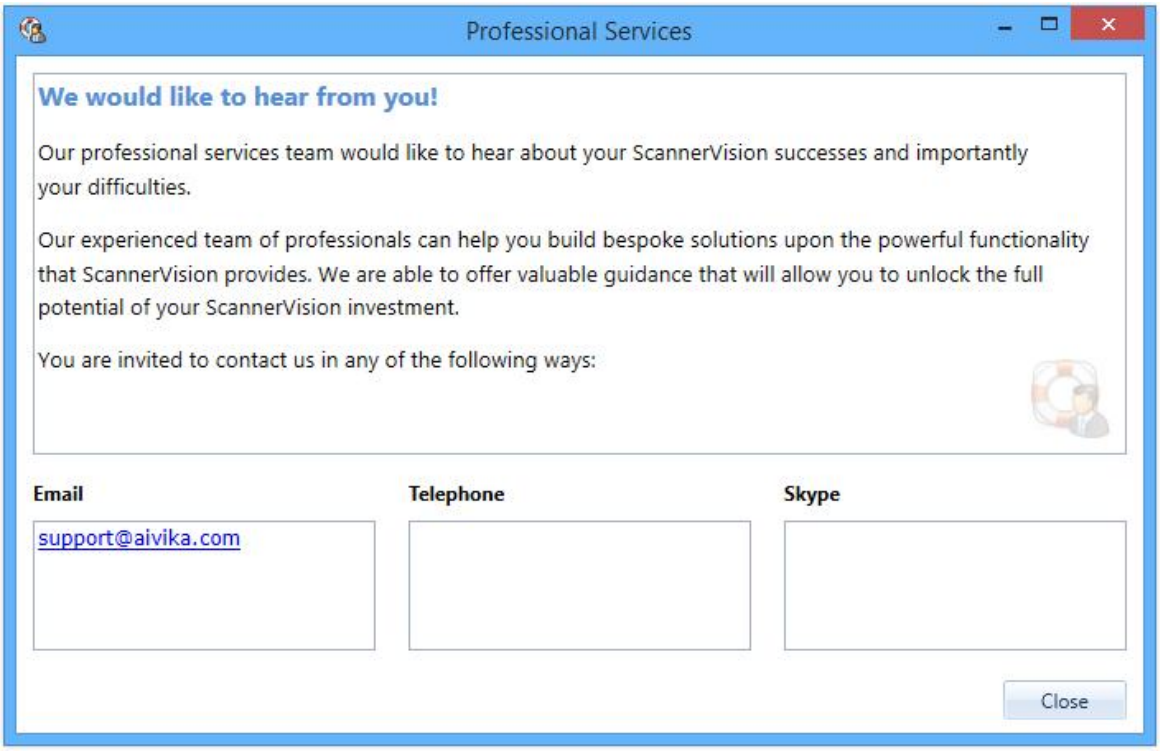
support@scannervision.com. We will then introduce you to another partner in your region who would be able to assist you.

To make use of the services of ScannerVision Professionals to help with the design and implementation of customer specific integration solutions, you can click on the "Professional Services" button or in the Help menu shown here:





During the installation of ScannerVision your reseller will configure the information you see in the professional services window to be applicable to your region. An example of the default professional services window is shown below:



3 Declaration

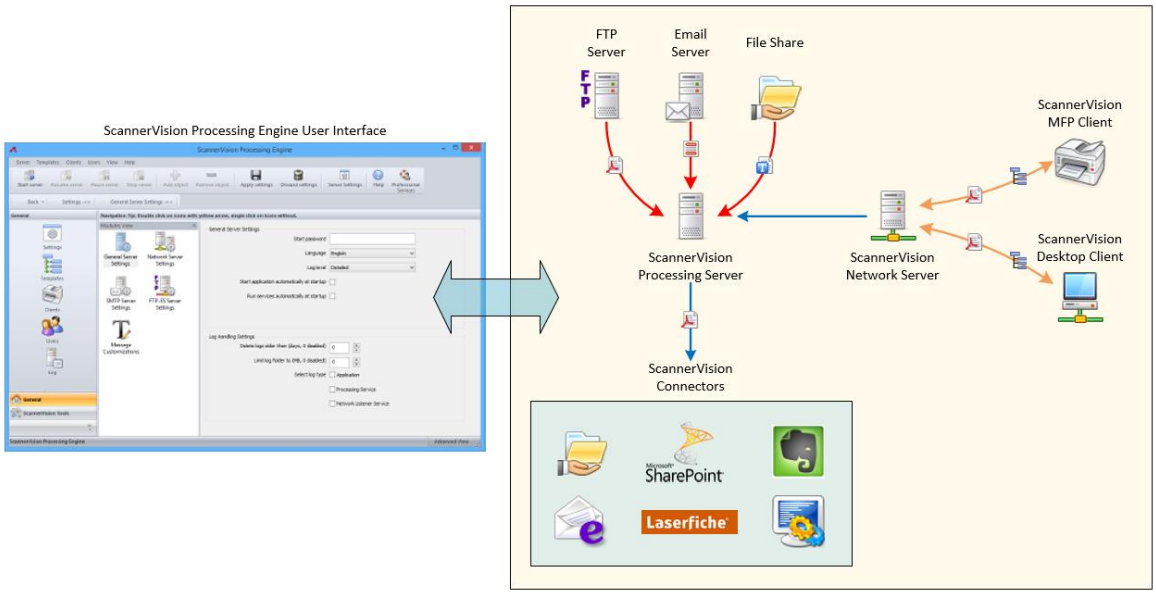
This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit. (http://www.openssl.org/)

4 Overview

In this section we present and overview of the primary role players in the ScannerVision system and how they participate in the three-tiered document work flow of Capture, Process and Store. These role players are:

- ScannerVision Processing Engine User Interface which is responsible for the configuration of the ScannerVision system from the management of services, clients and users to the configuration of templates.
- ScannerVision Processing Server which is responsible for the capturing, processing and storing of documents.
- ScannerVision Networking Server which is responsible for user authentication, serving templates to clients and resolving pick list data.
- Desktop & MFP Clients which are responsible for capturing documents and metadata and uploading them to the network server which in turn hands them over to the ScannerVision Processing Server for processing.

Below is a graphic representation of the ScannerVision ecosystem.



In ScannerVision you have control over WHERE documents come from by setting up Clients or configuring capture sources in a template, WHO are allowed to process documents by setting up Users and WHAT is done with the documents once captured by setting up Templates.

The WHERE

Documents reach the ScannerVision Processing Server from two groups of capture sources. The first group includes sources from where the ScannerVision Processing Server captures or pulls documents while the second group falls under the auspices of the ScannerVision Networking Server. Sources in this group upload or push documents to the ScannerVision Networking Server which in turn hands them over to the ScannerVision Processing Server for processing.

Group 1	Group 2
ScannerVision Processing Server "pulls" from source	**Capture source "pushes" to ScannerVision Networking Server**
FTP server	ScannerVision Desktop Clients
Email server	ScannerVision MFP Clients (Ricoh, Kyocera, Samsung etc.)
Local or shared network folder (called Hot Folder capture)	Secure FTP server used by embedded Kyocera Mita client

Group 1 capture sources are configured on a template level which is to say the template contains information of where documents are to be captured from. All documents arriving in the capture source are processed by the particular template unless more than one template is configured to capture documents from the same source.

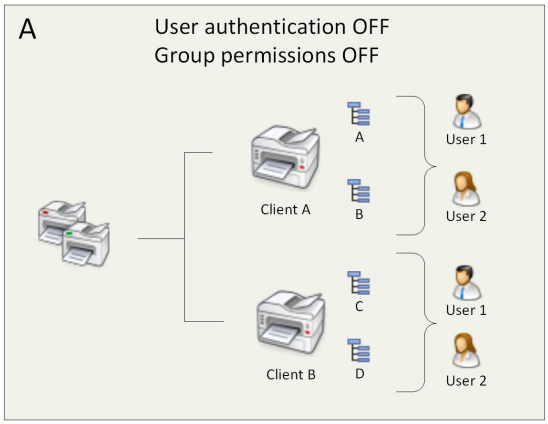
Group 2 capture sources are configured outside the context of a template. These capture sources are interactive and allows the user to select the template that will process the document. The selected template could be configured to configure the capture source for optimal capturing such as controlling the resolution at which the document is scanned.

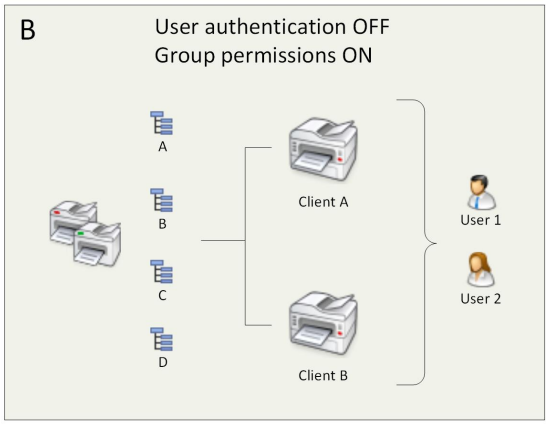
The WHO

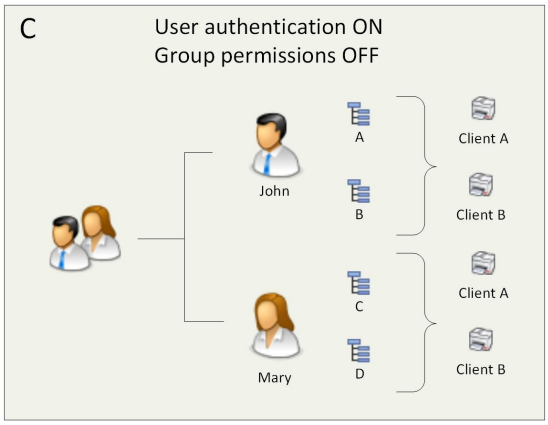
You can control which users are allowed to process documents through ScannerVision in various ways, depending on where the documents come from. With Group 1 sources you have indirect control through the security features offered by the particular source. For example with Hot Folder capture only users with write access to the shared folder can upload documents. Similarly when capturing from an FTP server only users with credentials/permission to connect to the ftp server can upload documents.

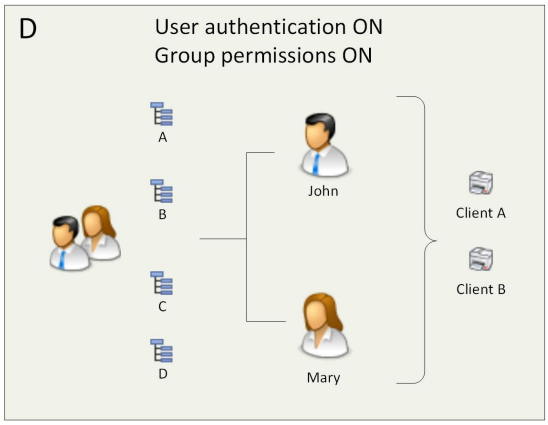
With Group 2 capture sources youhave more fine grained control over which users are allowed to process documents ranging from free-for-all to authenticated user level control. In the free-for-all scenario you don't have to configure users, only clients while in the authenticated scenario you have to configure users who will have to authenticate themselves before they will be able to upload documents. In both these scenarios the user is presented with a list of templates to choose from before a document can be uploaded. The list of templates can be controlled on an individual client/ user level or on a client/user group level.

Below is a graphic representation of the various options:









A Any user can upload documents and the templates they can choose from are specific to the client.

B Any user can upload documents and the templates they can choose from are the same on all clients belonging to the group.

C Only authenticated users can upload documents and the templates they can choose from are specific to the client.

D Only authenticated users can upload documents and the templates they can choose from are the same on all clients belonging to the group.

The WHAT

Templates determine what happens to a document once it reaches the ScannerVision Processing Server. A large part of this help manual is devoted to the functionality available to you when processing a document. Some of the things you can do include:

- Splitting the document into multiple parts
- Route split document to different destinations (including other templates) based on certain conditions

- Read 1D and 2D barcodes
- OCR raster documents such as Tiff images or fax documents in order to convert them to searchable PDF documents or even Microsoft Word or Excel documents
- Image processing such as de-speckling, hole punch mark removal, rotation etc.

How you configure your templates will be determined by your business needs. ScannerVision offers you with a lot of functionality that can transform the way your business operates.

4.1 Document Processors

ScannerVision has a the ability to process multiple documents at the same time. Each document is processed by a "Document Processor". The more Document Processors you have licensed the more documents can be processed in parallel. Each Document Processor is a fully fledged processing engine that is responsible for the entire workflow of every template. This means that each Document Processor will process every active template and it will perform every function specified by the template. It is not a case of a particular Document Processor instance processing only certain templates or certain tasks such as OCR.

Every capture source (Hot folder, Ftp, Email and Client) is limited to 3 minutes processing time per capture source. After 3 minutes of say hot folder capture the Document Processor will move on to Ftp capture, then Email and so forth - whichever capture sources are enabled. After all capture sources have had 3 minutes to execute the template run is complete and the Document Processor moves on the next template. If a particular document takes 3 minutes or longer it is processed to completion without interruption. If a document is processed in less than 3 minutes the Document Processor will continue to process as many documents as it can within the 3 minute window. The 3 minute execution time limit ensures that all templates are executed and that the system does not become bogged down with a particular capture source in a particular template.

Note

Multiple Document Processors will not improve the processing speed of a particular document but rather the multiple document are processed simultaneously.

Performance

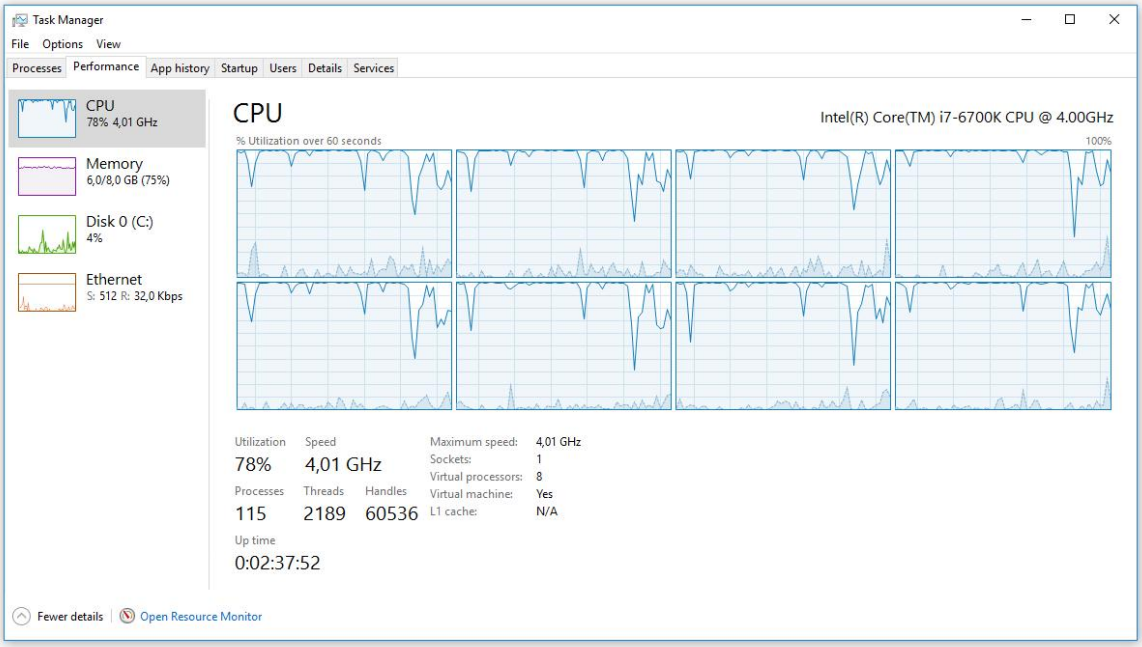
The number of Document Processors you choose to license on your system depends on the type of processing that is performed and how you store the documents. Maximum document throughput will typically be constrained by two primary factors:

1. CPU

2. IO (Disk and network Input/Output)

CPU

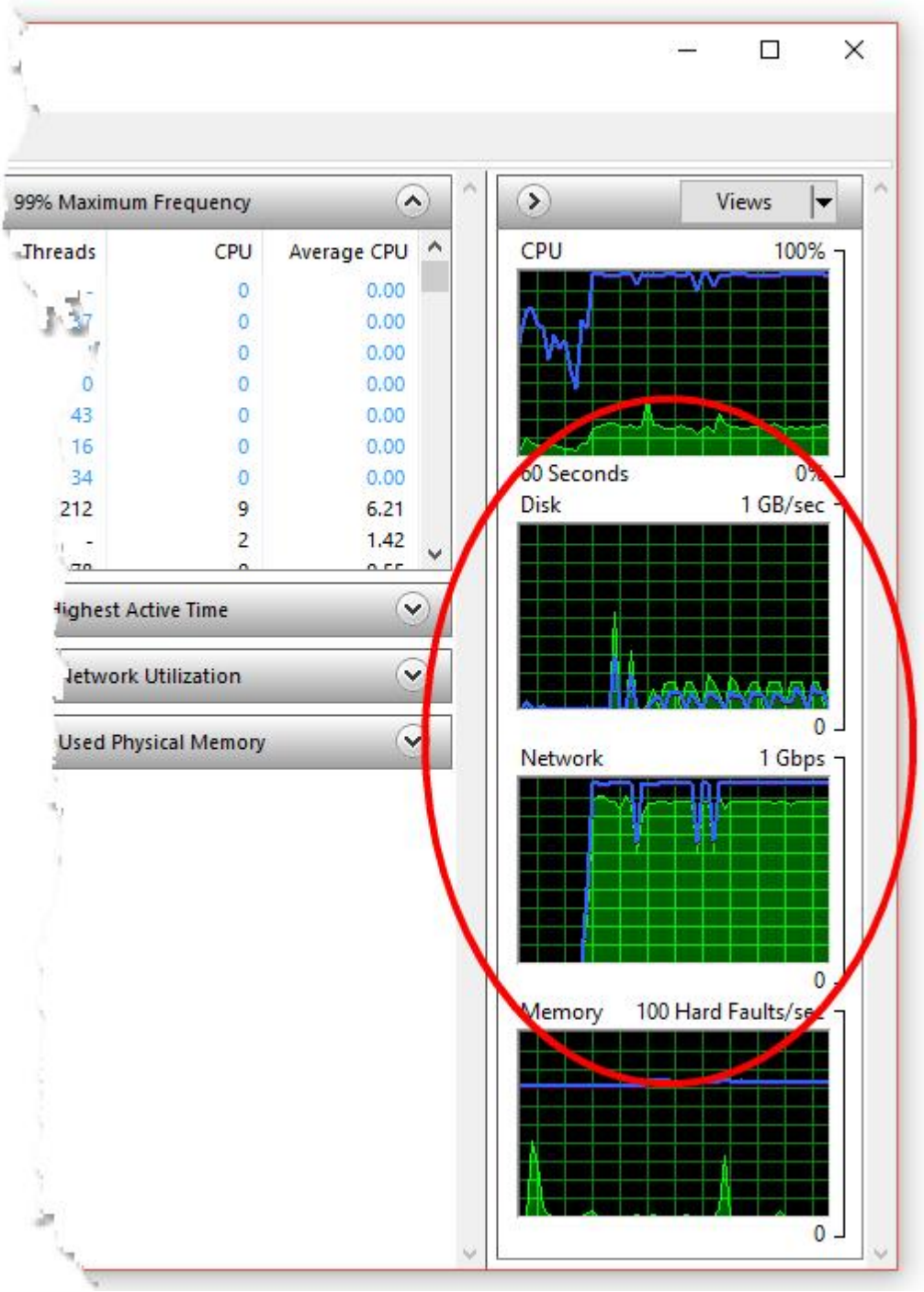
When you have a CPU bound system it means that the logical processors on your machine are working as hard as they can to get the work done. If for example your are doing Automatic Forms Recognition, OCRing, barcode reading or image processing your system will likely become CPU bound as these functions require a lot of processing power. In the screen shot below you can see that the system is performing close to its maximum capacity. This is a virtual machine with 8 logical processors and 8 Document Processors running. Adding more Document Processors will most probably not improve throughput much, if at all, and it may in fact harm throughput on average since the processors' task switching overhead may negate the benefit of the additional Document Processors.



Input/Output

When a system is IO bound it means that the hard disk(s) and/or network interface is running at full capacity. If you are not doing document conversion or any of the CPU heavy functions mentioned before, such as when "Bypass image processing" is enabled, there will be very little CPU dependent activity and a lot of disk and/or network usage and hence your system will likely be IO bound. The CPU may be running at 40-50% capacity while the hard drives are running at 100%. In this scenario

you would gain no benefit by adding additional Document Processors since they will be waiting most of the time for the IO devices - whether it is to get files loaded from or saved to the hard drives or to get it uploaded to the DMS.



It is therefore important to measure the performance of your system under real world conditions. If you require high document throughput it is important to ensure that the factors mentioned above are

in balance. Also bear in mind that if you are running other services on the machine such as a DMS or database engine, they will compete with ScannerVision for the same system resources and all will suffer. It is therefore recommended to run ScannerVision on a dedicated server.

4.2 Running the Servers

The ScannerVision Processing Server and ScannerVision Networking Server are independent 32bit Windows console applications which are installed as Windows services by the ScannerVision installer. This allows them to be run in "headless" environments where a user does not have to be logged in to the server in order for them to run.

You can also run them as console applications.

Note

Both the ScannerVision Processing Server and ScannerVision Networking Server require elevated (administrator) privileges to run as console applications.

To run the servers as console applications do the following:

- Create a shortcut to the respective exe in the ScannerVision installation directory. The ScannerVision Processing Server application is called "svnts-pr.exe" and the ScannerVision Networking Server "svnts-cn.exe".
- Right click on the shortcut and select "Properties" from the pop up dialog.
- Select the "Shortcut" tab.
- In the "Target" edit box add the text **cmdline** at the end of the existing text with a space in between.
- Select the "Compatibility" tab and ensure that the "Run this program as an administrator" check box is selected.
- Press the "OK" button.
- Double click the newly created shortcut

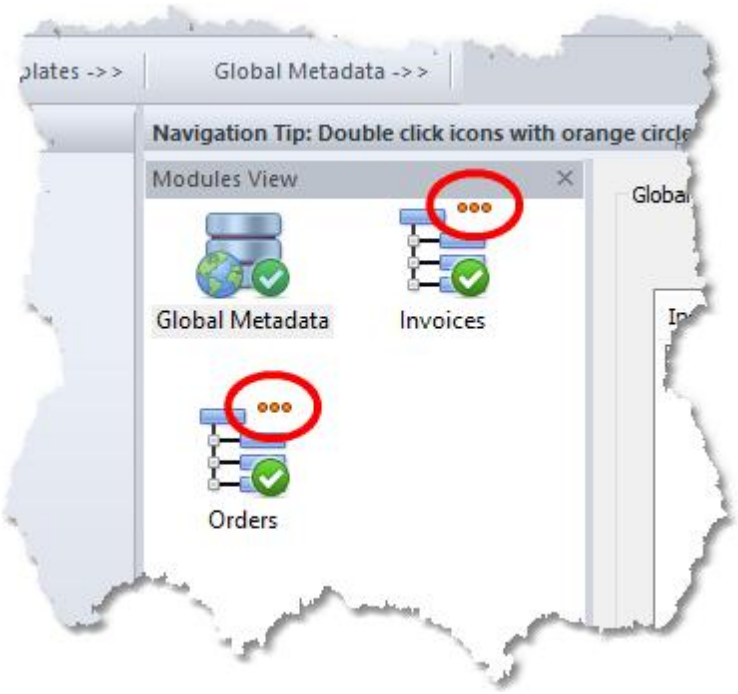
4.3 User Interface Visual Cues

There several visual cues that are used throughout the ScannerVision user interface to help you to navigate and to visually represent the state of settings, options and functions.

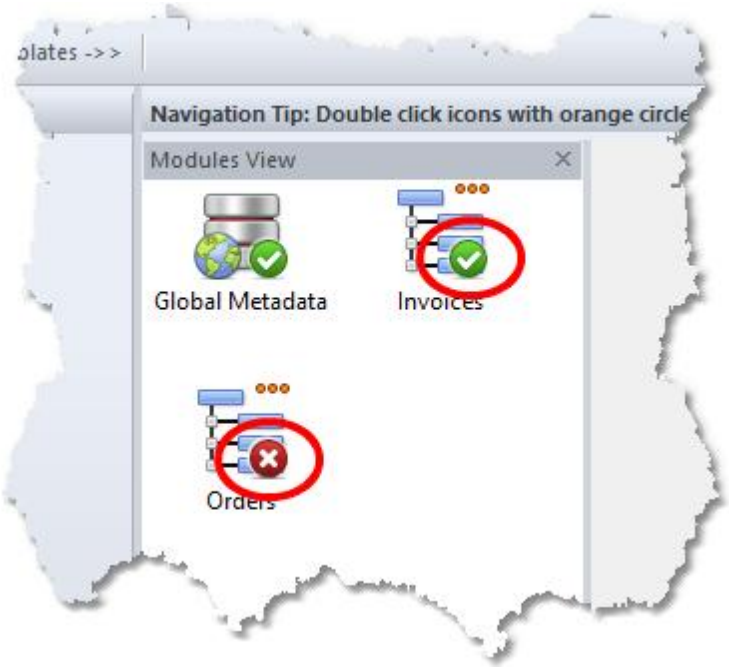
The first queue is the 3 orange dots shown here:

Whenever you see these dots in the top right hand corner of an icon it indicates that there are more

settings "behind" the icon. In other words, you should double click the icon to reveal more settings. In the screen shot of the Templates' modules view shown below, you can see that the Invoices and Orders templates both have the dots and the Global Metadata icon doesn't.

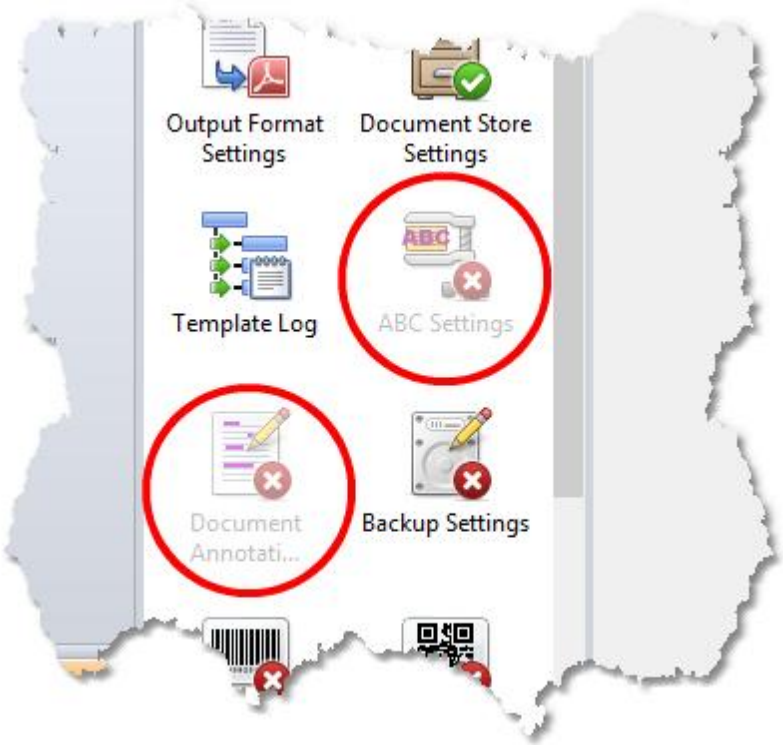


The second queue is a green check mark or a red cross in the lower right hand corner of an icon as shown below:



You can see that the Invoices template above has the green check mark overlay whereas the Orders template has the red cross overlaid. A green check mark indicates that the particular function/option/ settings is enabled while the red cross indicates that it is disabled.

The third queue that is used is grayed out icons. In the screen shot below you can see that the ABC Settings and Document Annotations icons are grayed out. This means that these modules are not licensed.



5 License Activation

ScannerVision offers an Evaluation Mode under which all functionality is available for you to try and play with. In this mode vector documents (Searchable PDF, PDF/A, Microsoft Word & Excel etc.) are truncated to 2 pages maximum while raster formats such as Tiff and PDF are watermarked.

In Production Mode you would need to have a valid license file which removes the above limitations.

5.1 Evaluation Mode

ScannerVision's evaluation mode allows you to evaluate all ScannerVision's functions, modules and connectors but with two restrictions. A "NOT FOR RESALE" watermark is added to every page of raster output documents such as Tiff and PDF while vector output formats such as Searchable PDF, PDF/A, Microsoft Word and Excel are limited to two output pages. A page is inserted at the end of the document that states that it has been created with an evaluation version of ScannerVision and that some pages have been removed from the document.

The number of Document Processor instances that will run is equal to the number of logical processors as reported by the operating system.

An example of a watermarked page is shown below:

General Services (Pty) Ltd.

SERVICE INVOICE

Name General Services
Address 125 High Street
Industrial Park
Capital City, 1000

☒ Apply discount
☒ First time customer

SOLD TO: Customer Name
100 Main Street
Capital City
1000

SERVICED AT:

Sales Tax Rates: 6,50% On Parts 9,50% On Labor

INVOICE #	MAKE OF EQUIPMENT	MODEL #	SERIAL #	INVOICE DATE	SERVICE DATE
123654789	Compressor	C002B	SN002B-10100	January 5, 2014	January 5, 2014

PARTS USED

QTY	DESCRIPTION	PRICE	AMOUNT
1	Product	18,76	18,76
8	Product	4,32	34,56

SERVICE PERSON	DATE	HOURS	RATE	AMOUNT
Name	07-13-09	12,0	$40,00	$480,00
Name	07-14-09	4,0	65,00	$260,00
			TOTAL	**$740,00**

PARTS	$53,32
LABOR	740,00
TAX	73,77
TOTAL	**$867,09**

GUARANTEED 30 DAYS AGAINST LABOR AND MATERIAL DEFECTS

COMMENTS:

T. Moore
SERVICE SUPERVISED BY

T. M
SIGNED

5.2 Production Mode

You have to activate your ScannerVision installation before you can create fully functional templates or run the ScannerVision Processing Server without limitation.

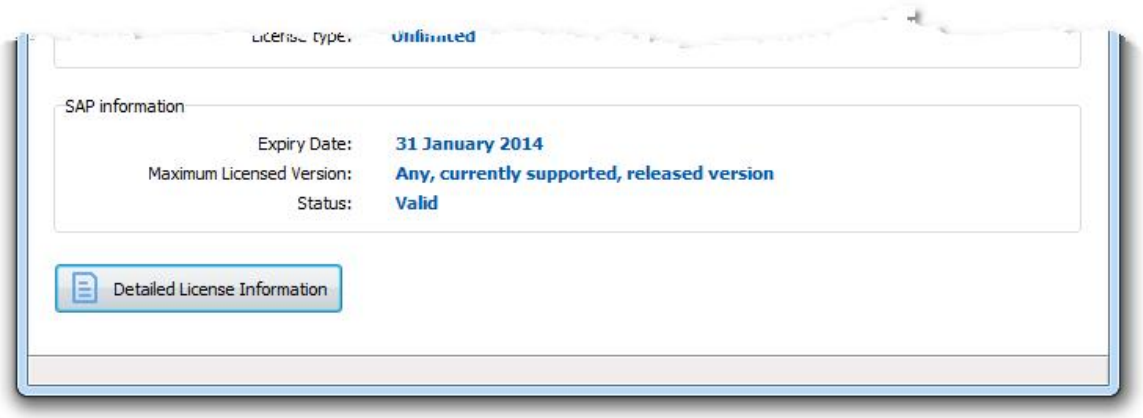
ScannerVision includes a rich set of features as standard but there are several features and connectors that can be purchased separately. A ScannerVision license can be tailored to your specific needs. Your license is tied to the serial number you received with your purchase. If you have not received a serial number or have lost it please contact your reseller. This serial number will not change with future changes to your license.

ScannerVision supports 2 different licensing models namely a full license with no limit on the number of documents that can pass through the system and a Pay Per Scan (PPS) model where you are charged for each document that passes through the system on a prepaid or a post-paid basis. The specifics of these models are discussed later.

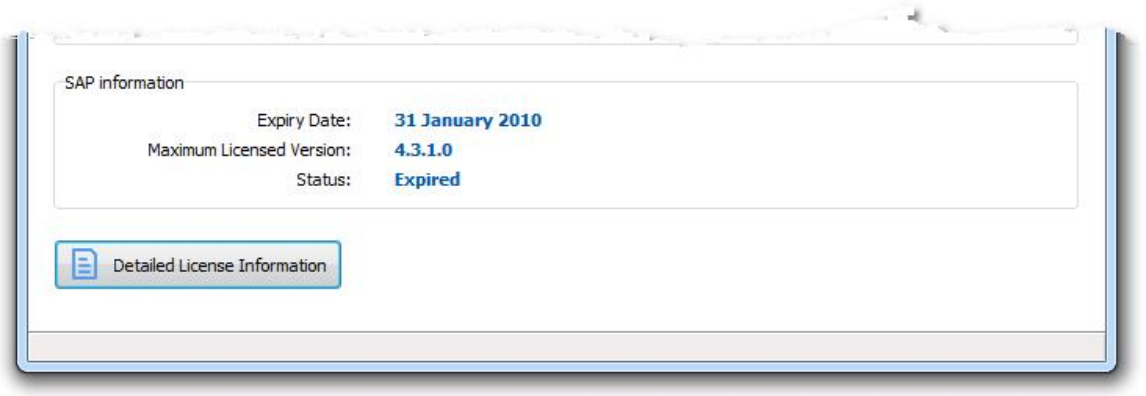
Maintenance Program

With a ScannerVision Maintenance Program subscription you are entitled to support as well as to any major version of ScannerVision that is released during the your maintenance period. If your Maintenance Program subscription expires, you are entitled to run the latest release version of ScannerVision at the time that your subscription expired.

If your Maintenance Program subscription is up to date (not expired) the License Manager will show the following:



If your Maintenance Program has expired the License Manager will show the following:



The "Maximum Licensed Version" indicates the version you are **entitled** to run and may not be the same as the version that you are actually running. If the version you have installed is later than the one shown you will not be able to run it and if it is earlier you are entitled to upgrade your installation to the version indicated.

* Pay Per Scan is only available in certain countries. Please contact your reseller for details.

5.2.1 License Manager

The ScannerVision License Manager is a standalone application with which you can manage and view the specifics of your license. License activation can be done in either online or offline mode. In the online mode all you need to do is to enter your serial number and click the "Activate" button. The License Manager will contact the ScannerVision license portal and retrieve your specific license file.

In the offline mode you need to create a server information file which you will then have to upload to the ScannerVision licensing portal from a computer with Internet connectivity. A license activation file will then be generated that you must downloaded to your computer. You must copy the license activation file to the ScannerVision server from where you can activate your license. Detailed steps to do this are discussed in the Offline Activation section.

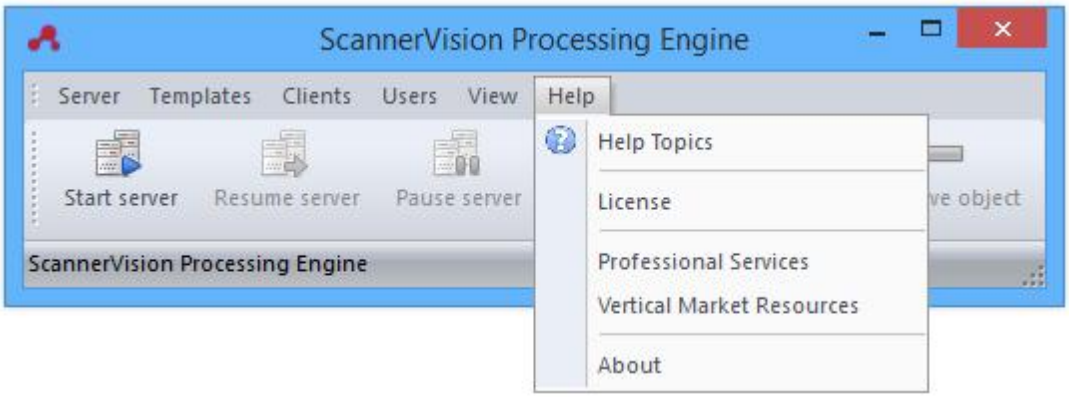
We recommend that you use the online activation as it is much easier and quicker than the offline process. You could temporarily connect your server to the Internet while you activate your license and then disconnect it afterward (please read the important note below if you intend to do this).

IMPORTANT NOTES

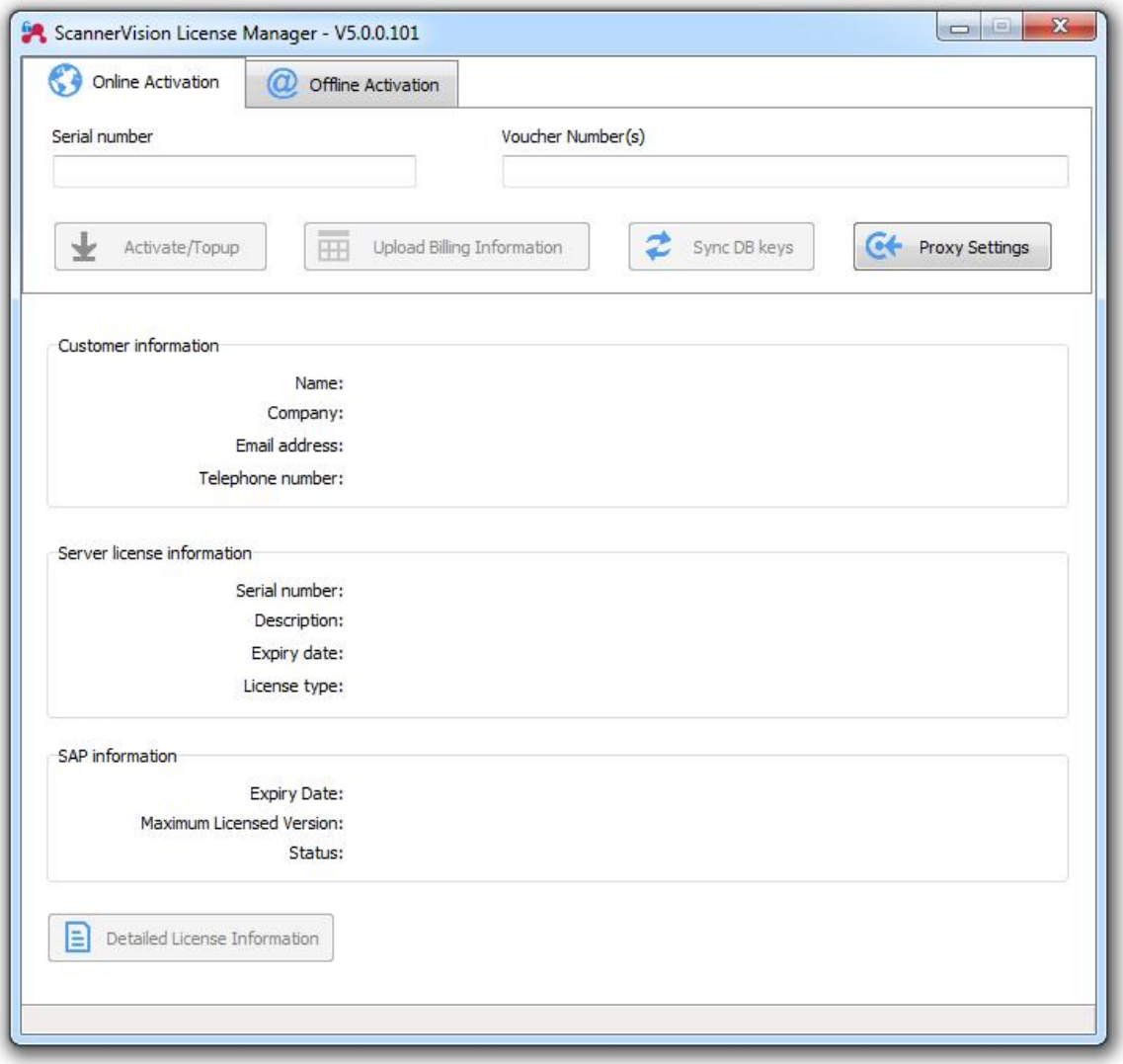
1. ScannerVision requires an active network interface card (NIC) in order to function. An installed NIC that is not connected to the network is not active and does not appear in the list of available NICs on the system so please ensure that at least one NIC is always active. Also ensure that the NIC that was active during the license activation process is also active whenever you want to run ScannerVision. A permanent Internet connection is not required for ScannerVision to function but one would be required for online license activation.

2. **Ensure that the date and time of your** ScannerVision server is correct (we recommend that you synchronize the system time with a time server). If the UTC time of your server is out with more than 15 minutes you will not be able to activate your license.

To launch the License Manager click on the Help->License... menu:



The License Manager window will appear:



Select the activation method you want to use by clicking on the Online Activation or Offline Activation tab at the top of the screen.

5.2.1.1 Online Activation

To do an online activation of your license please follow these steps:

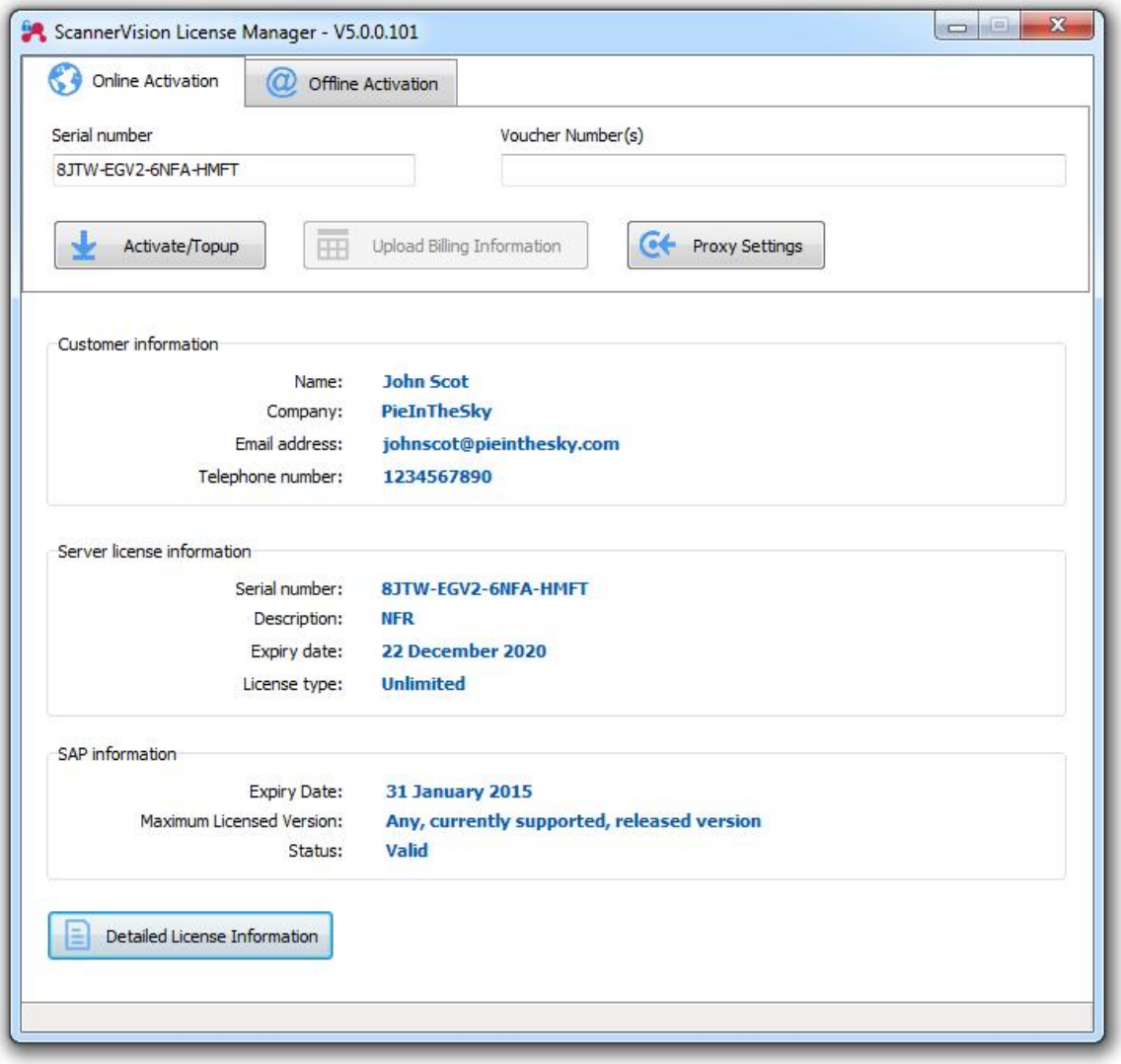
1. Enter your serial number.
2. Follow the optional steps below.

3. Click the "Activate/Topup" button.

Optional steps:

1. Add any per page billing voucher numbers you may have separated by commas.
2. If you connect to the Internet through a proxy server you must provide the details of the proxy server by clicking on the "Proxy Settings" button and completing the information.

If the license activation was successful your screen will update with the particulars of your license:



To view the full details of your license please click the "Detailed License Information" button in the lower left hand corner of the screen. Please verify that all the information in the license is correct.

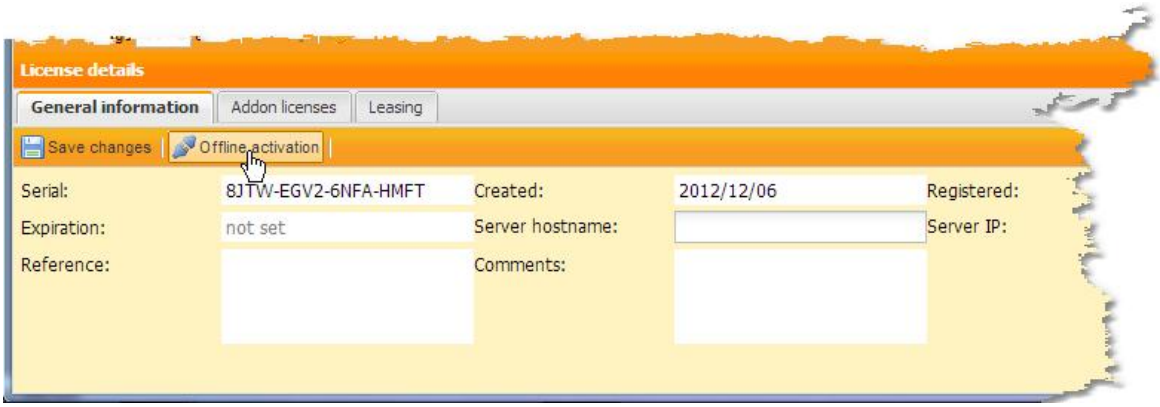
5.2.1.2 Offline Activation

In order to do an offline license activation you would need access to a computer/laptop that has access to the Internet. You also need to be registered on the ScannerVision licensing portal.

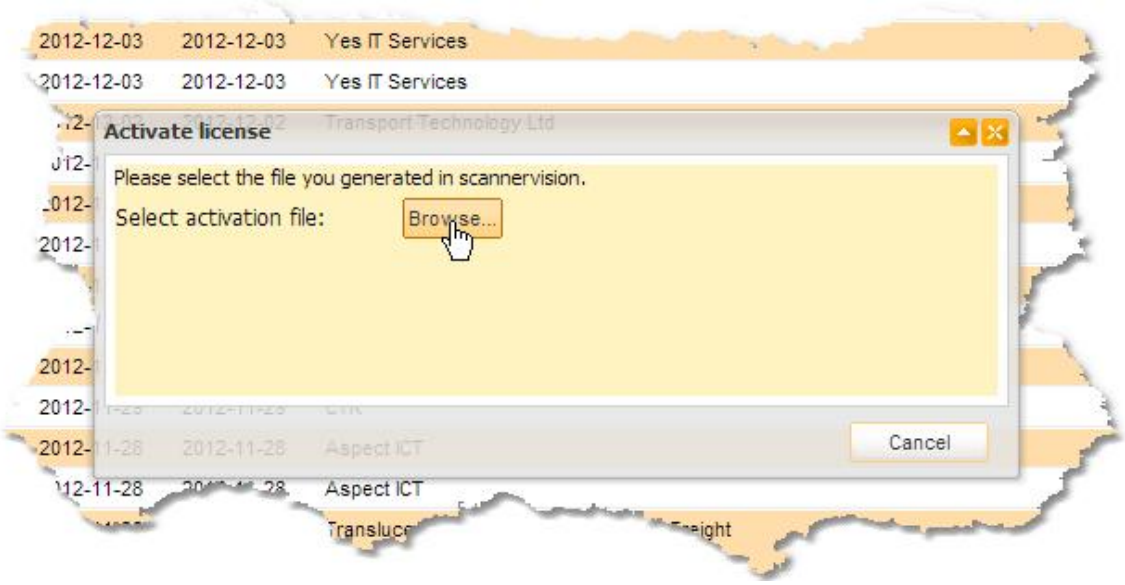
To do an offline activation of your license please follow these steps:

1. Add any per page billing voucher numbers you may have separated by commas (optional).
2. Click the "Generate" button. A "Browse For Folder" dialog will appear.
3. Select the folder where you want the server information file to be saved and click "Ok" on the "Browse For Folder" dialog. A confirmation dialog will appear with the path and file name (ServerInfo.dat) of the server information file.
4. Copy the ServerInfo.dat file to the computer/laptop that has Internet connectivity.
5. Go to the http://my.scannervision.com licensing portal and log in to the portal.
6. Click on the "License" button on the toolbar:

7. Find your serial number (license key) in the list and select it.
8. Click the "Offline activation" button:



9. Click the "Browse" button on the popup window and navigate to the "ServerInfo.dat" file created above.



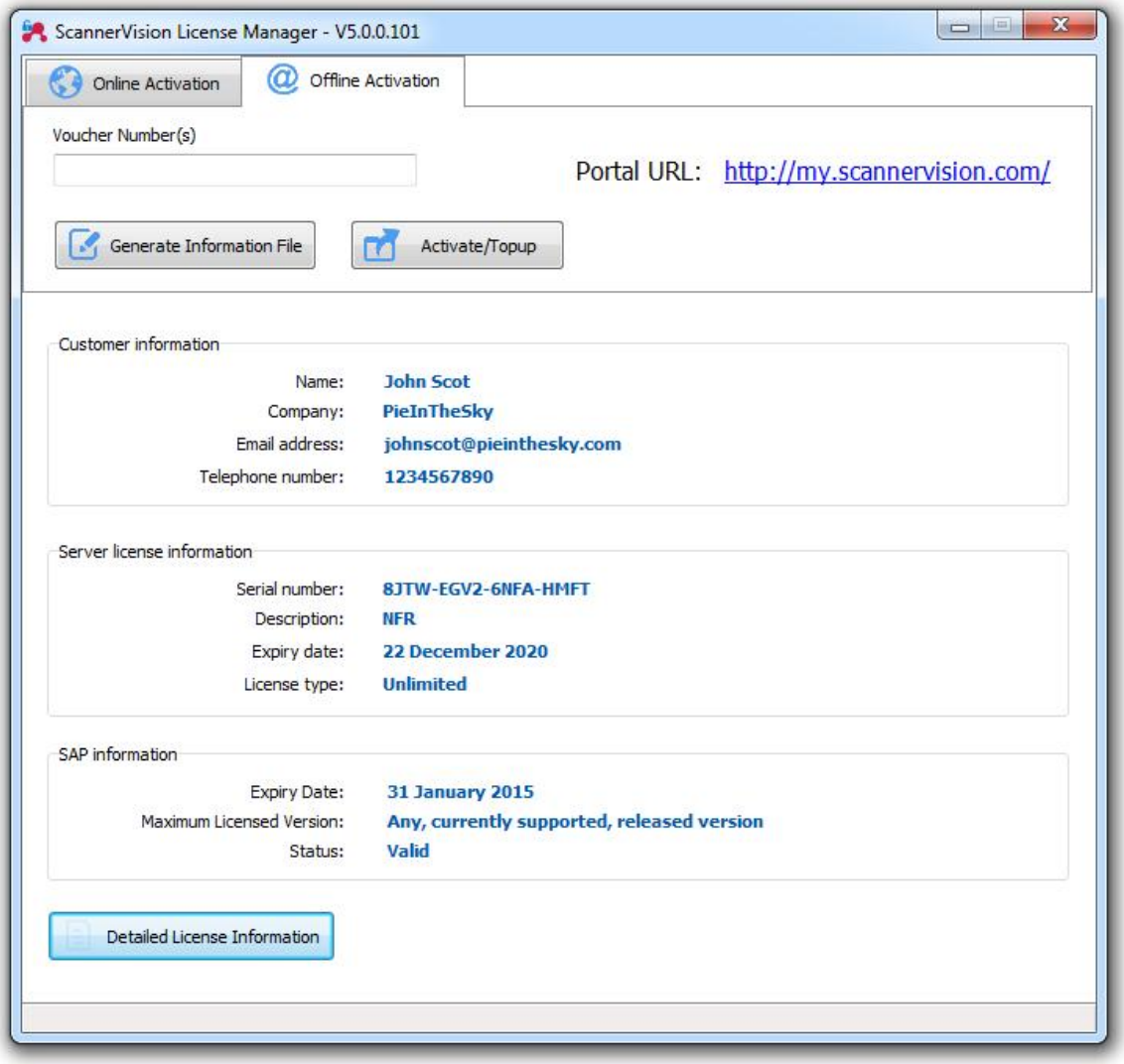
10. A popup will appear confirming the activation of your license:

11. Click the "Ok" button. A license activation file called "activationfile.laf" will be downloaded to your computer and sent to your email address.

12. Copy the "activationfile.laf" to the ScannerVision server.

13. Click the "Activate/Topup" button on the License Manager application and navigate to where you have copied the "activationfile.laf" above and click "Open".

If the license activation was successful your screen will update with the particulars of your license:



To view the full details of your license please click the "Detailed License Information" button in the lower left hand corner of the screen. Please verify that all the information in the license is correct.

5.2.2 Pay Per Scan (PPS)

NOTE

Pay Per Scan is only available in certain countries. Please contact your reseller for details.

PPS allows you to pay only for the pages that passes through ScannerVision on a prepaid or post-paid basis. With the prepaid option you buy vouchers that you load into ScannerVision through the License Manager. The voucher represents a number of credits which are loaded into the database. Once credits have been loaded your ScannerVision is ready to process documents. With the post-paid option you pay for the documents that have passed through the system at the end of the month. You

can also load vouchers for the post-paid option in which case the number of credits will be added to your balance.

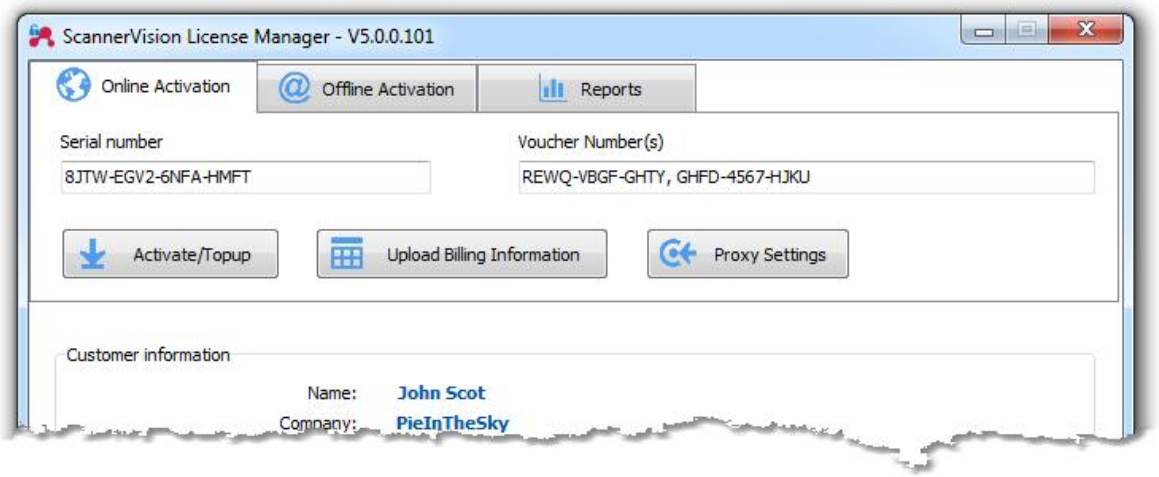
Depending on certain allowances in your license a portion of your unused credits can be carried over to the following month. See the Document cost calculation section for more details.

NOTE

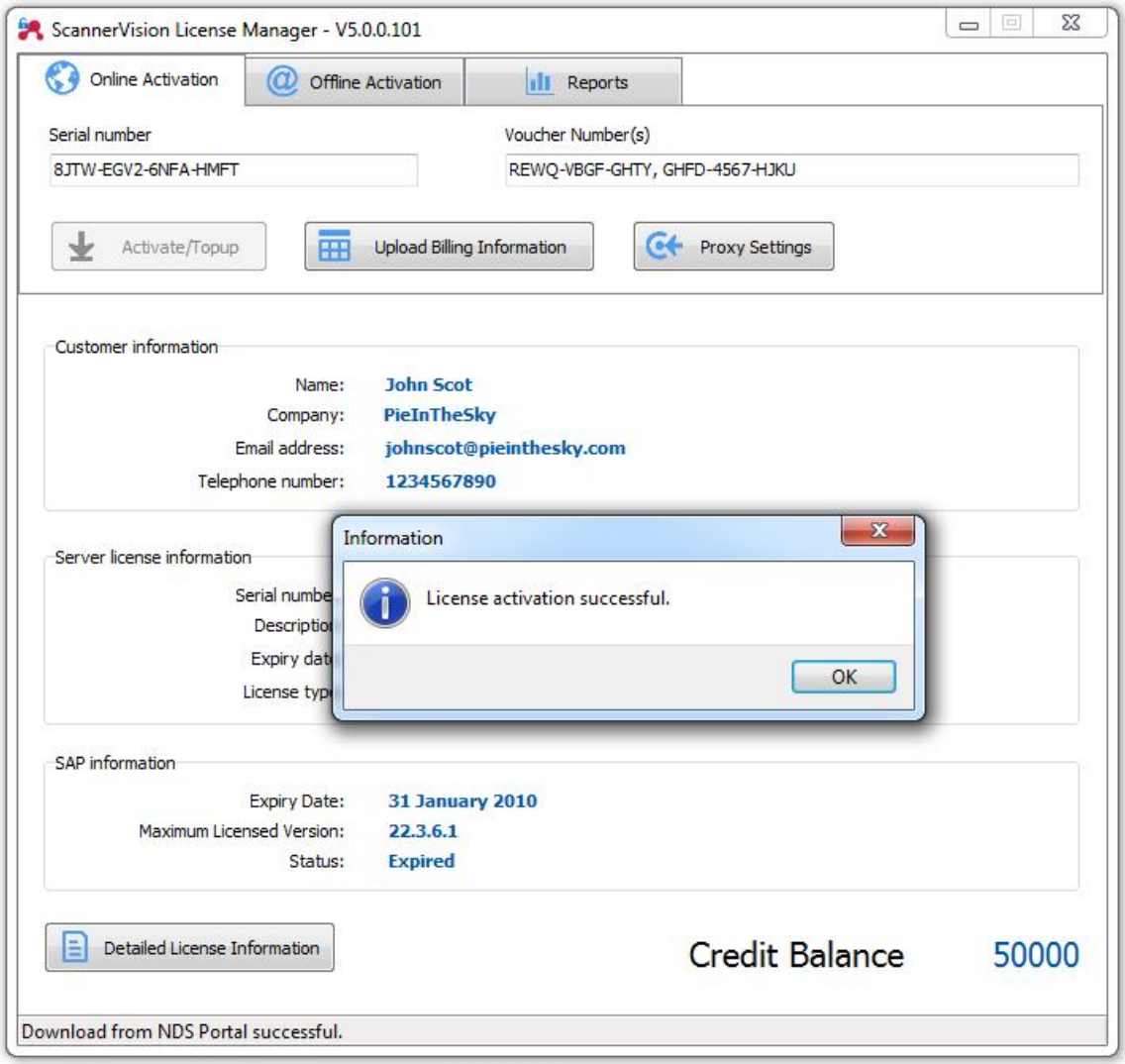
PPS credits do not relate to currency on a one to one basis. In other words, 1 credit does not necessarily equal 1 Euro/Dollar/Pound/Rand etc.

Post-paid PPS requires a permanent Internet connection and is therefore not available as an offline option.

When activating your license you can enter any voucher numbers you have in the "Voucher Number(s)" edit box before you click the "Activate" button as shown below. You can redeem vouchers after you have activated your license by entering them into the "Voucher Number(s)" edit box and clicking the "Activate" button.



Once the license has been activated you will see the following:



Notice the addition of the "Credit Balance" value in the lower right hand corner of the screen as well as the "Reports" tab that has appeared at the top of the screen. These items will appear only when you have a PPS license.

Credit Balance

The credit balance value is your current credit balance at this moment in time. It is a real time value which means it is updated immediately when a document has passed through the system. So while the processing engine is running and processing documents, you will see this value changing continuously.

In the prepaid mode you have to have a large enough positive credit balance before processing of a document will be allowed. In post-paid mode you can have a positive or negative credit balance.

Reports

From the reports tab you can run various reports of the transactions that have gone through ScannerVision. For more information refer to the Reports section.

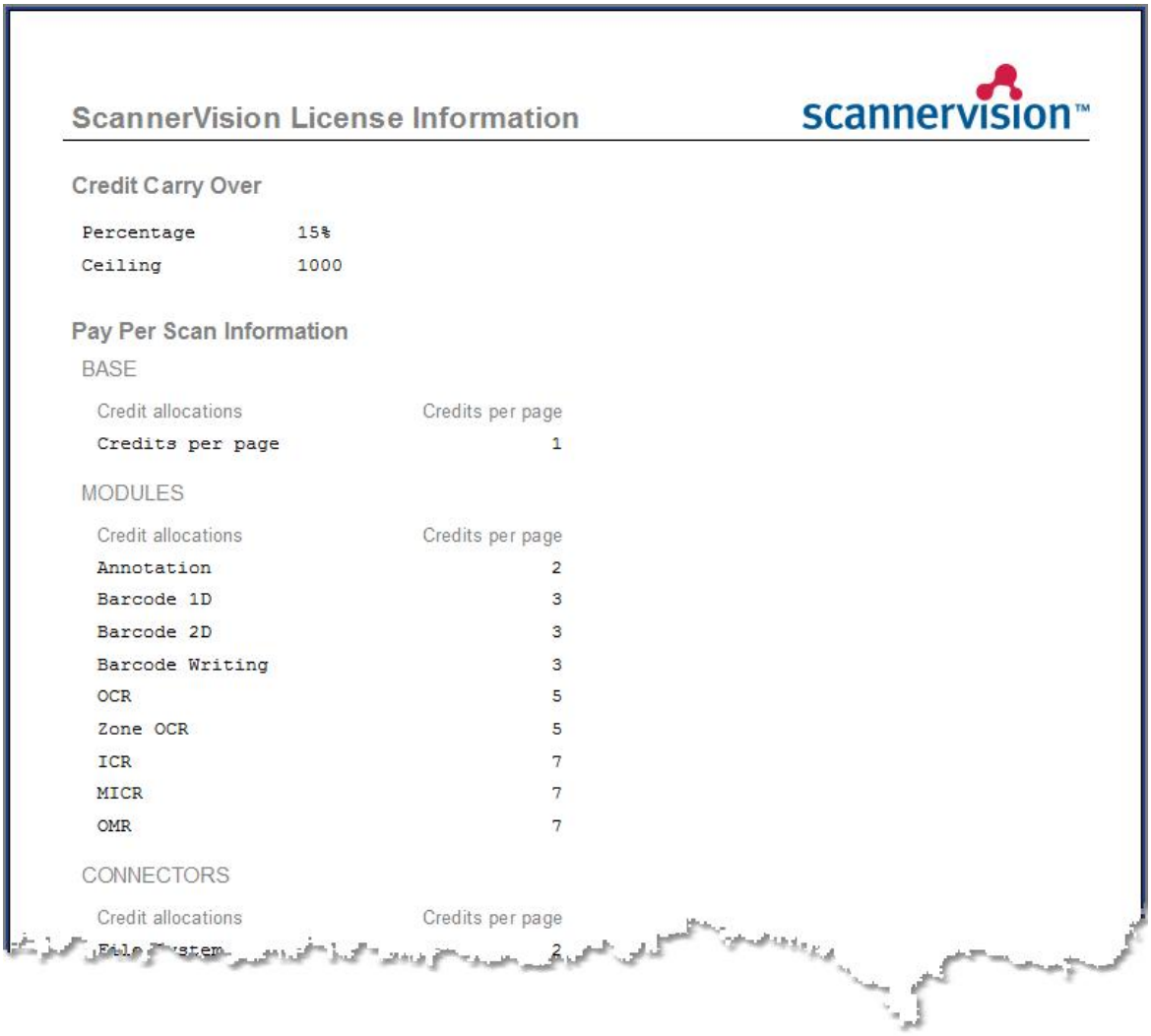
5.2.2.1 Document Cost Calculation

After you have successfully activated your PPS license you should look at the specifics of your license by clicking the "Details License Information" button. Navigate to the page of the license report that contains the PPS information.

NOTE

PPS credits do not relate to currency on a one to one basis. In other words, 1 credit does not necessarily equal 1 Euro/Dollar/Pound/Rand etc.

The license information report will look something like this:



ScannerVision License Information

scannervision™

Credit Carry Over

Percentage	15%
Ceiling	1000

Pay Per Scan Information

BASE

Credit allocations	Credits per page
Credits per page	1

MODULES

Credit allocations	Credits per page
Annotation	2
Barcode 1D	3
Barcode 2D	3
Barcode Writing	3
OCR	5
Zone OCR	5
ICR	7
MICR	7
OMR	7

CONNECTORS

Credit allocations	Credits per page
File System	2

Credit Carry Over

If at the end of the month you have not used all your credits, a portion (or all) of your credits can be carried over to the following month. How much is carried over is calculated as follows:

- If only the ceiling value is specified (Percentage = 0%) all unused credits are carried over up to the ceiling level. So if your credit balance at the end of the month is 1500 credits and the ceiling value is 1000, then 1000 credits are carried over and you loose 500 credits. If your credit balance is 700 at the end of the month all 700 credits are carried over.
- If only the percentage value is specified (Ceiling = 0) then the specified percentage of unused credits are carried over with no upper limit. So if your credit balance at the end of the month is 15000 and the percentage value is 15% then 2250 credits are carried over (15000 x 15% = 2250).
- If both values are specified the percentage of unused credits are carried over up to the ceiling limit. So if your credit balance at the end of the month is 15000 and the percentage value is 15% and the ceiling value is 1000 then 1000 credits are carried over (15000 x

15% = 2250 which is greater than the ceiling of 1000).

- If both the percentage and ceiling values are 0 no credits are carried over.

Pay Per Scan Information

The cost (in credits) of processing a document through ScannerVision depends on the functions that are used when processing the document like OCR, barcode reading etc. and the number of pages in the document. The cost of using a function is specified next to the respective function on a cost per page basis. In the image above for example the cost of using OCR is 5 credits per page and Barcode 1D is 3 credits per page.

The "Credits per page" value is a rate that is applied for every page regardless of the functions used. So if you don't use any of the functions specified above and process a 15 page document the cost of that document would be 15 credits (1 credit per page x 15 pages = 15 credits).

Cost Calculations Example

Let's say you have enabled Barcode 1D, Zone OCR and ICR in your template and you are processing a 20 page document, the cost of the document would be as follows:

3 credits for barcode 1D x 20 pages	= 60 credits
5 credits for Zone OCR x 20 pages	= 100 credits
7 credits for ICR x 20 pages	= 140 credits
1 credit per page x 20 pages	= 20 credits
Total	**= 320 credits**

NOTE

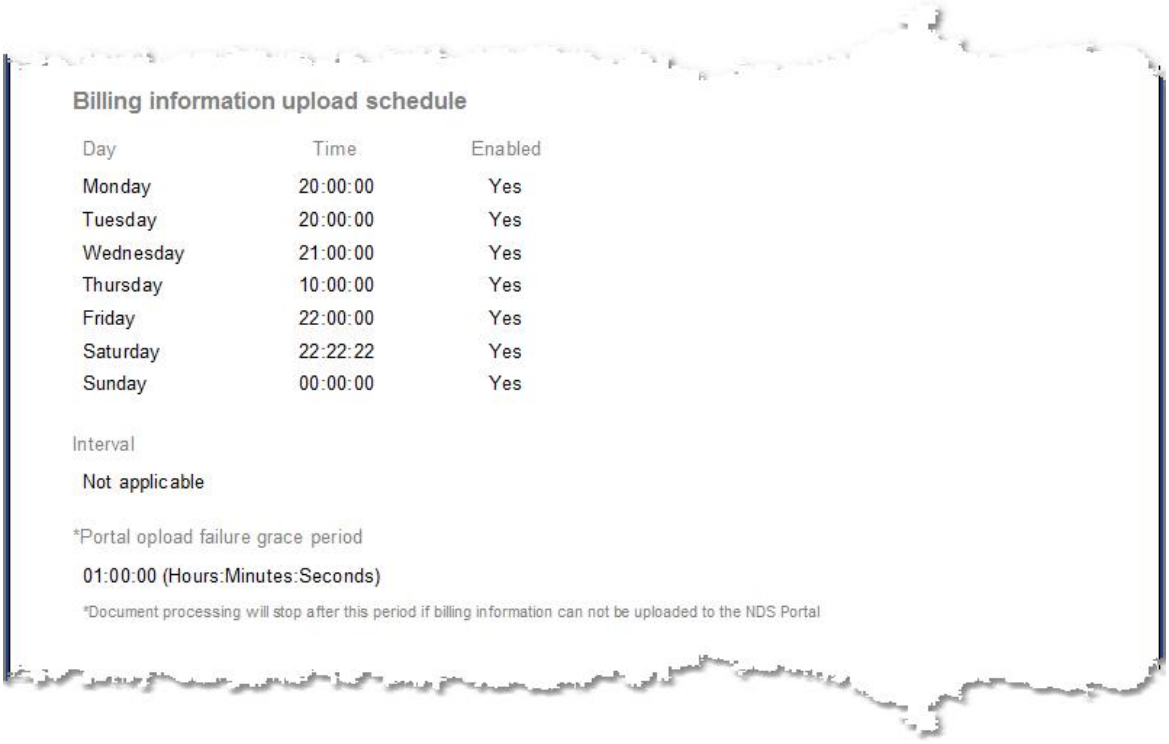
All enabled functions' cost accrue to the total document cost regardless of whether the function produced any outcome. For example, if Barcode 1D is enabled and your documents does not contain any barcodes the cost of Barcode 1D is still applied to every page.

5.2.2.2 Uploading of Billing Information

When you have a post-paid PPS license your transaction data is uploaded to the ScannerVision license portal on a regular basis. You therefore need to ensure that your server is permanently connected to the Internet. If ScannerVision is not able to contact the license portal document, processing will stop as described below.

Uploading of billing information is a function of the processing service and therefore will only occur while the processing service is running.

To view details of how often ScannerVision will upload billing information to the license portal click on the "Detailed License Information" button on the License Manager and scroll the page where you see something like this:



Billing information upload schedule

Day	Time	Enabled
Monday	20:00:00	Yes
Tuesday	20:00:00	Yes
Wednesday	21:00:00	Yes
Thursday	10:00:00	Yes
Friday	22:00:00	Yes
Saturday	22:22:22	Yes
Sunday	00:00:00	Yes

Interval

Not applicable

*Portal opload failure grace period

01:00:00 (Hours:Minutes:Seconds)

*Document processing will stop after this period if billing information can not be uploaded to the NDS Portal

Billing information upload schedule

The upload schedule indicates how often billing information is uploaded to the ScannerVision licensing portal. If your license does not contain explicit scheduling information, billing data will be uploaded daily at 12 o'clock at night.

The schedule can be specified in terms of explicit week days and time of day or as an interval. If any of the days of the week are enabled, i.e. the value under the "Enabled" heading is "Yes", then uploads will happen on those days at the specified local time as indicated under the "Time" heading. If no days are enabled, uploads will happen on a scheduled basis.

In the above image you will notice that all the days are disabled and "Interval" now has a value which indicates that uploads will happen every 6 hours.

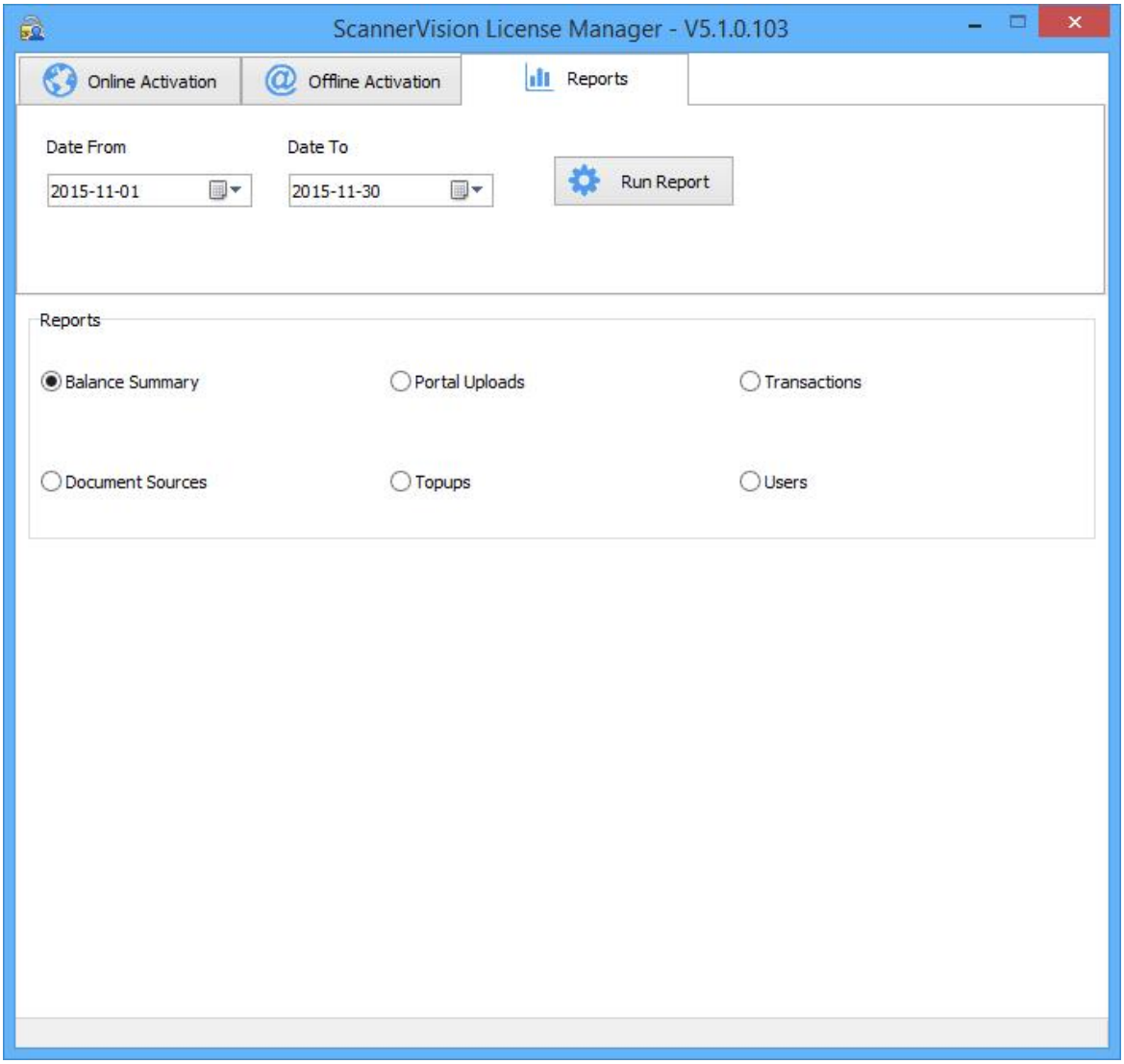
Failure to upload

If for some reason ScannerVision is not able to upload billing information to the portal, it will enter a grace period which is specified in the license file. As per the image above the license allows for a grace period of 1 hour but it can be any value as dictated by the license.

During the grace period ScannerVision will continue to process documents while it retries to upload billing information to the portal every 5 minutes. If the grace period lapses without any successful upload processing will stop. If you restart the processing ScannerVision will first try to upload the billing information and only if it was successful will it allow document processing to resume.

5.2.2.3 Reports

There are 6 reports that you can run as shown below:



To run a report

When you open the License Manager it defaults the dates over which report are run to the beginning and end of the current month. You can change the date range with the "Date From" and "Date To" date pickers. Select the report you want to run and then press the "Run Report" button.

Reports

Below is a description of the information each report offers for the reporting period:

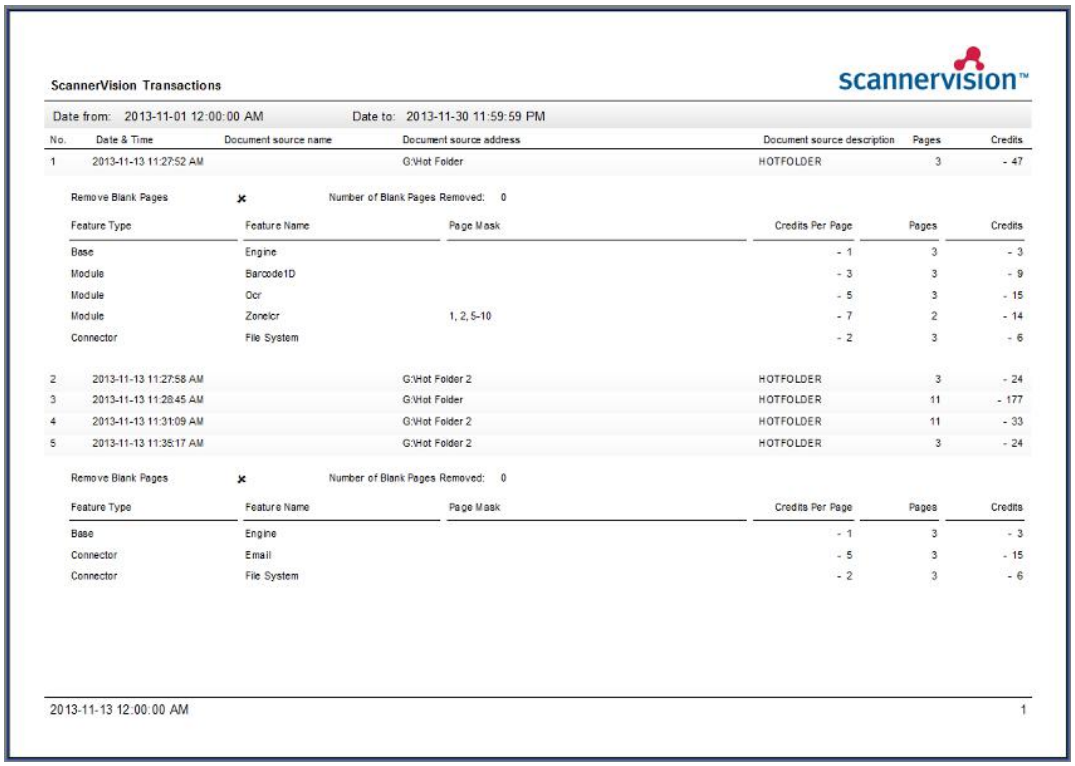
Balance Summary	A summary of all credit topups, transactions and portal uploads.
Portal Uploads	A list of all portal uploads and the number of credits that have accrued since the last upload.

Document Sources A list of document sources that produced documents for processing. In this report you will see all the hot folders, ftp servers and clients etc. from where documents entered the system. The report does not show which sources are configured but rather the ones from which documents arrived for processing.

Topups A list of all vouchers that were loaded into ScannerVision showing when the voucher was loaded, the voucher number and the number of credits.

Users A list of users that produced documents.

Transactions Below is a screen shot of a transaction report:



ScannerVision Transactions

scannervision™

Date from: 2013-11-01 12:00:00 AM Date to: 2013-11-30 11:59:59 PM

No.	Date & Time	Document source name	Document source address	Document source description	Pages	Credits
1	2013-11-13 11:27:52 AM		G:\Hot Folder	HOTFOLDER	3	- 47

Remove Blank Pages ✘ Number of Blank Pages Removed: 0

Feature Type	Feature Name	Page Mask	Credits Per Page	Pages	Credits
Base	Engine		- 1	3	- 3
Module	Barcode1D		- 3	3	- 9
Module	Ocr		- 5	3	- 15
Module	ZoneIcr	1, 2, 5-10	- 7	2	- 14
Connector	File System		- 2	3	- 6

2	2013-11-13 11:27:58 AM		G:\Hot Folder 2	HOTFOLDER	3	- 24
3	2013-11-13 11:28:45 AM		G:\Hot Folder	HOTFOLDER	11	- 177
4	2013-11-13 11:31:09 AM		G:\Hot Folder 2	HOTFOLDER	11	- 33
5	2013-11-13 11:35:17 AM		G:\Hot Folder 2	HOTFOLDER	3	- 24

Remove Blank Pages ✘ Number of Blank Pages Removed: 0

Feature Type	Feature Name	Page Mask	Credits Per Page	Pages	Credits
Base	Engine		- 1	3	- 3
Connector	Email		- 5	3	- 15
Connector	File System		- 2	3	- 6

2013-11-13 12:00:00 AM 1

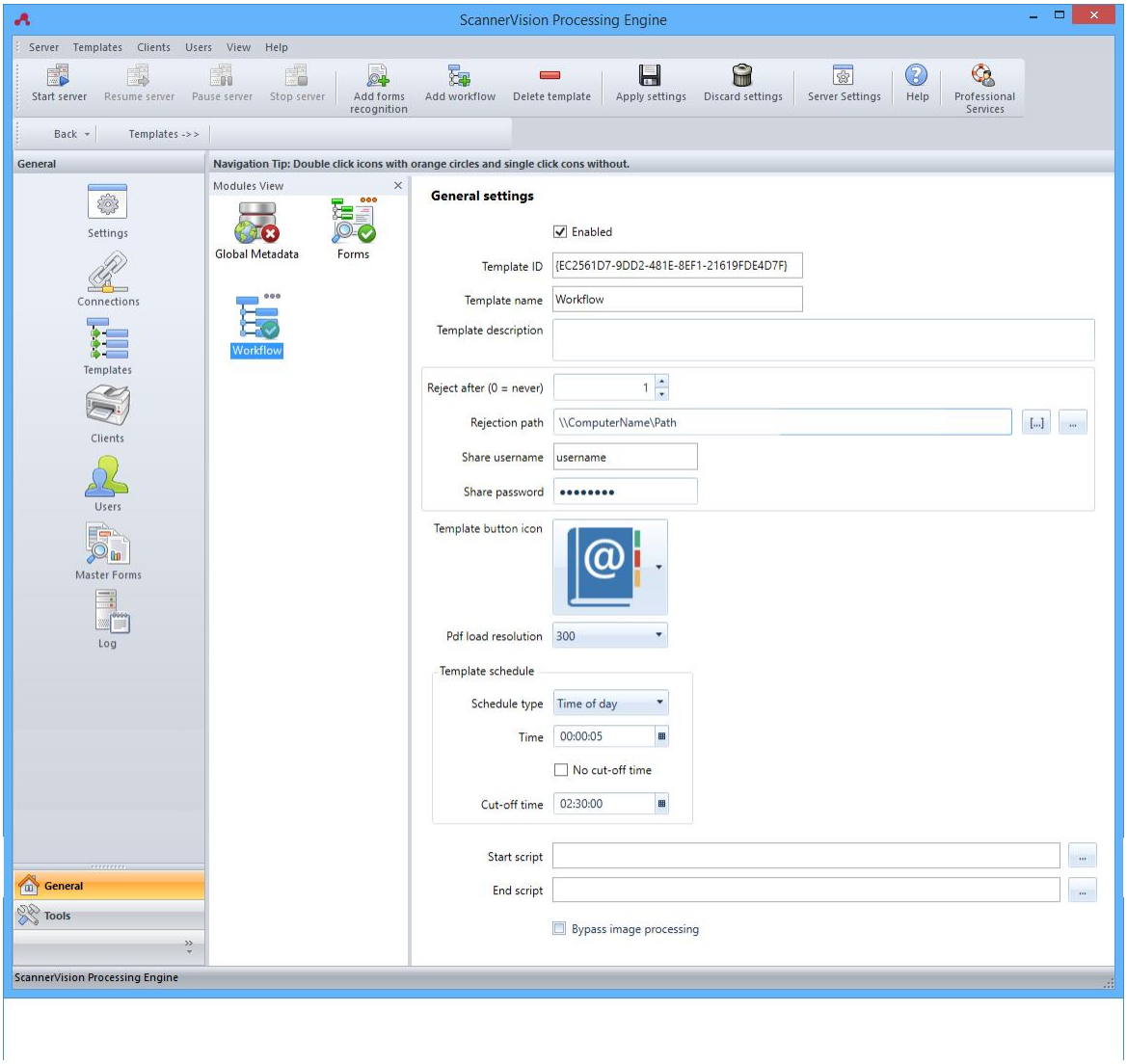
With the transaction report you can see details of each document that was processed by ScannerVision. Details that are shown include the source of the document, the number of pages, the cost of the document and specifics of which functions were used during the processing of the document. The latter is not shown by default as you can see in document numbers 2, 3 & 4 in the image above. Documents 1 & 5 on the other hand show the details of the functions used. To reveal the details of a transaction click on the line in which it appears.

6 Setup

Every aspect of ScannerVision's operation is configured using the ScannerVision Processing Engine User Interface application shown below. In this section we will cover:

- ScannerVision Processing Engine User Interface
- Settings
- Configuring Connections
- Configuring Clients
- Configuring Users
- Log

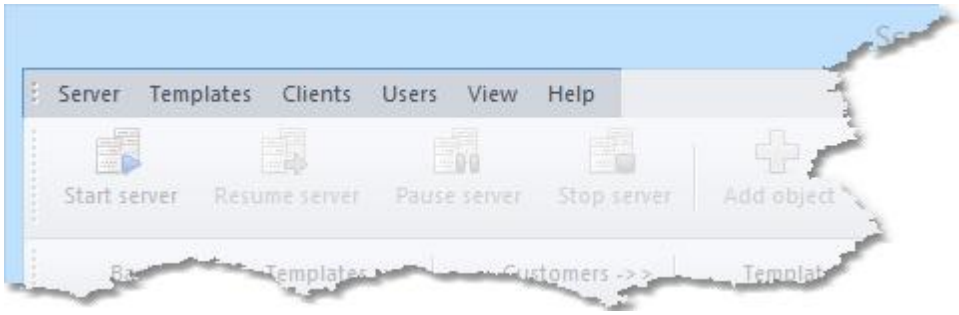
Templates are covered in detail in the Creating Templates section.



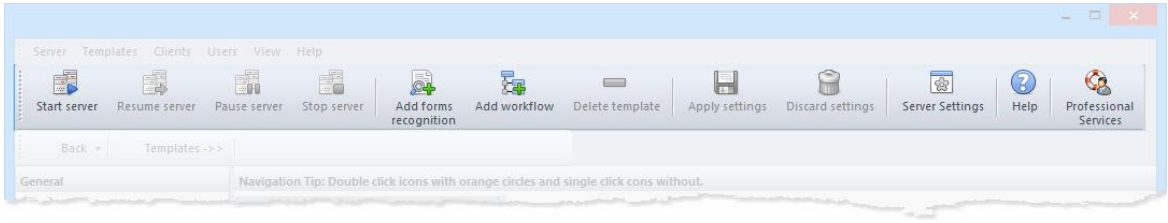
6.1 Processing Engine User Interface

The Processing Engine User Interface can be divided into the following areas:

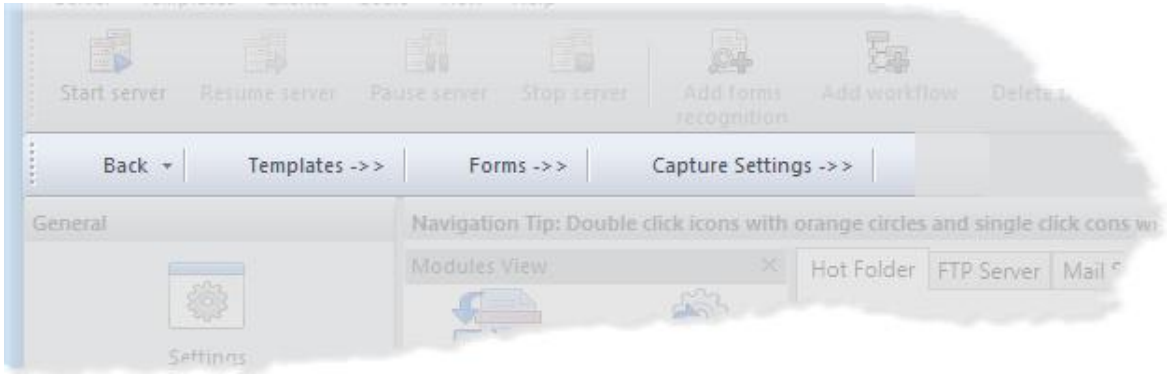
- Menu Bar



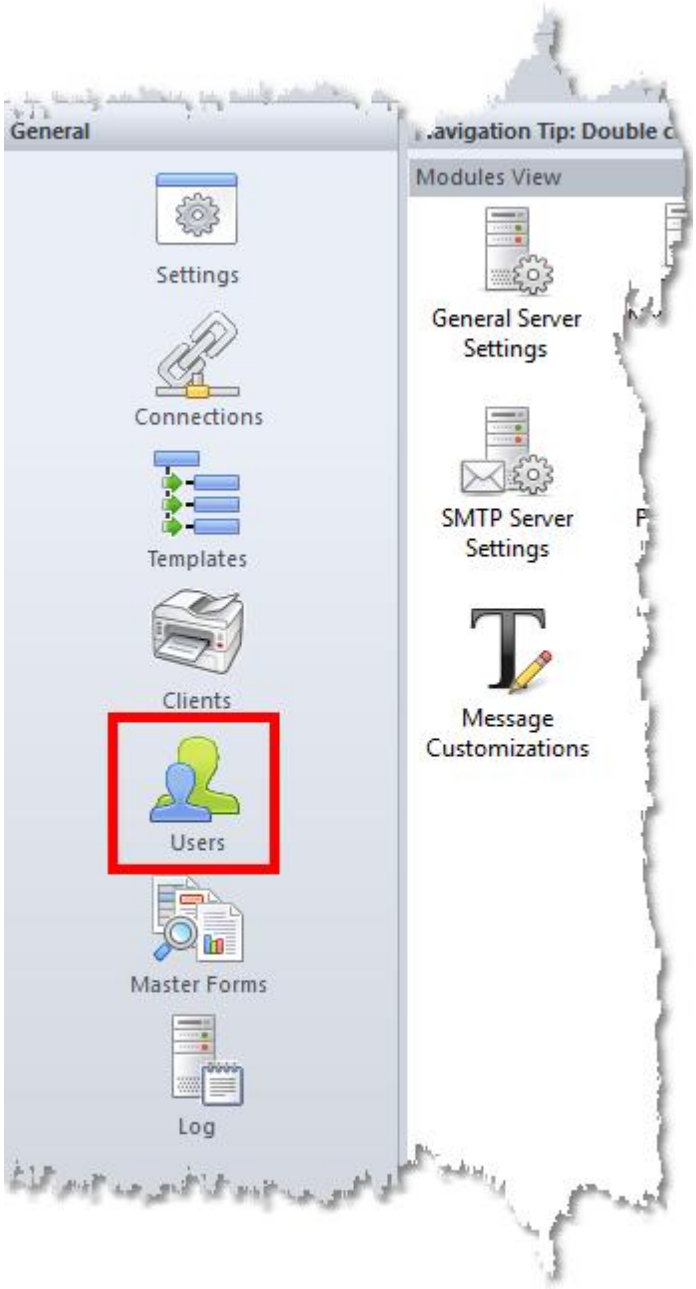
- Command Toolbar



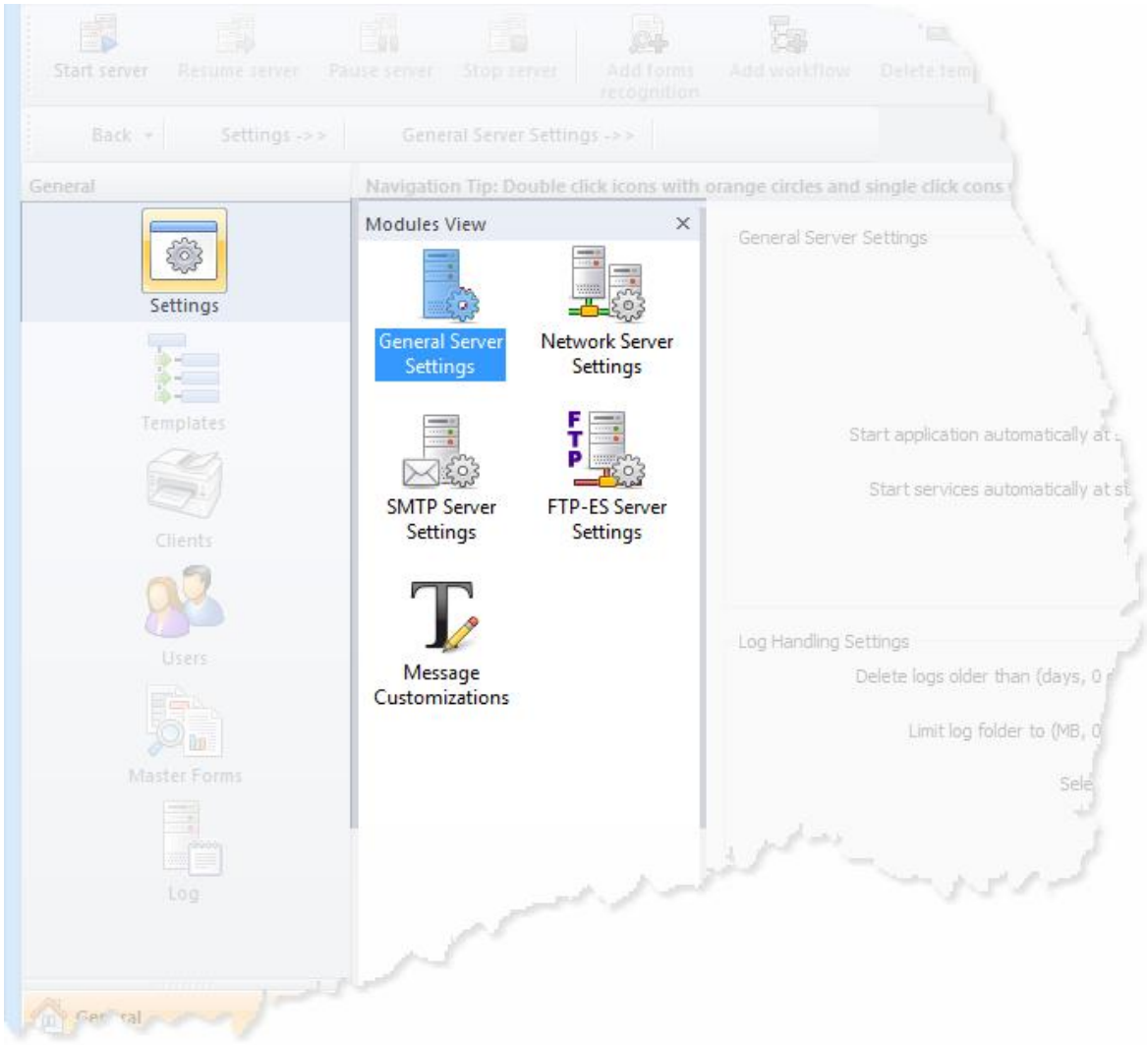
- Navigation Toolbar



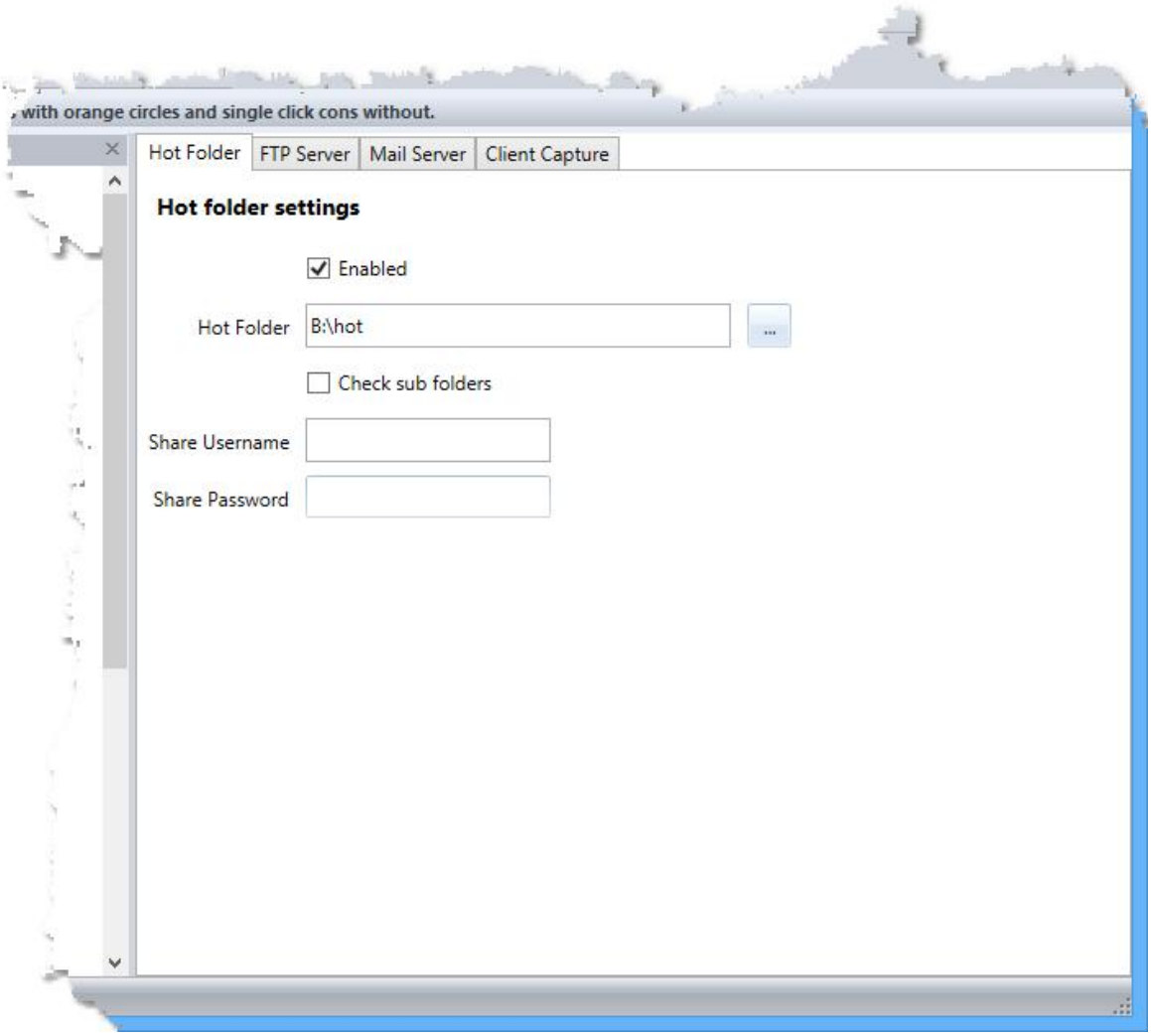
- Navigation Pane



- Module View

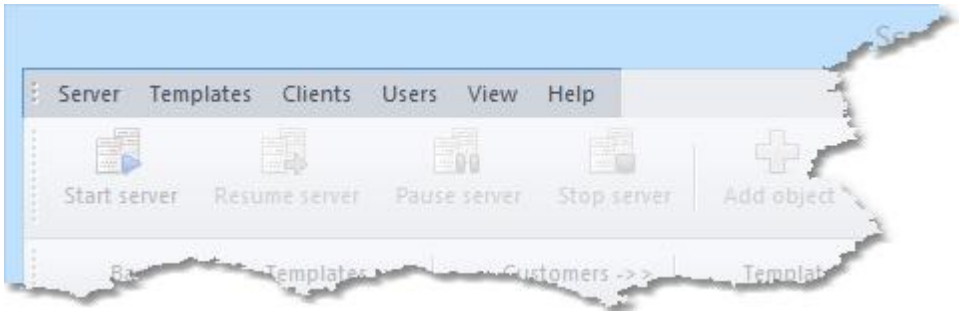


- Configuration Pane



6.1.1 Menu Bar

The menu bar contains links to configuration screens, commands and options which in certain cases have equivalent buttons on the Toolbar and Navigation Pane. Certain options and commands are available in the menu bar only.



Server

Start	Starts the ScannerVision Processing Server and ScannerVision Networking Server services.
Resume	Resumes the processing service after it has been paused.
Pause	Pauses processing service only. The ScannerVision Networking Server continues to run.
Stop	Stops both the ScannerVision Processing Server and ScannerVision Networking Server services.
Settings	Navigates to the general settings screen.
Exit	Exits the application.

Templates

Show	Navigates to the templates' settings screen.
Add	Add's a new template.
Delete	Deletes an existing template.
Duplicate	Duplicates an existing template with a new template name and ID.

Clients

Show	Navigates to the clients' settings screen.
Add Group	Adds a new client group.
Remove Group	Deletes an existing client group.
Add	Add's a new client.
Delete	Delete an existing client.
Duplicate	Duplicates an existing client with a new client

	name and ID.

Users

Show	Navigates to the users' settings screen.
Add Group	Adds a new user group.
Remove Group	Deletes an existing user group.
Add	Add's a new user.
Delete	Deletes an existing user.
Duplicate	Duplicates an existing client with a new user name and ID.
Import	Launches the "Import Users" tool.
CSV Import	Launches the "CSV User Import" tool.

View

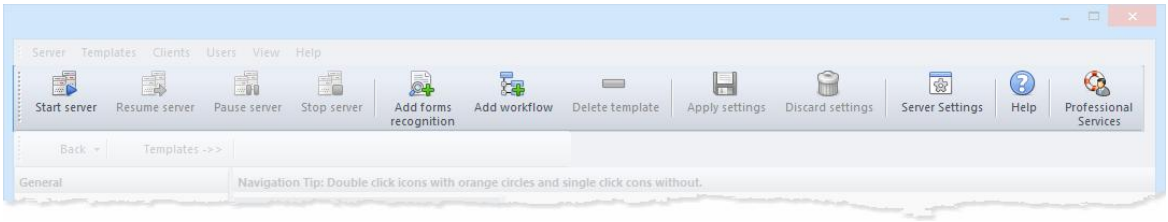
Templates	Switches the templates modules bar from an icon list to a text list.
Navigation Pane	Shows/hides the navigation pane.
Tool Tip Bar	Shows/hides the tool tip bar.
Modules View	Shows/hides the modules bar.
Statuc Bar	Shows/hides the status bar.

Help

Help Topics	Opens this help file.
License	Launch the License Manager.
Professional Services	Shows the Professional Services information dialog.
Vertical Market Resources	Opens the default browser and navigates to the ScannerVision Vertical Market web site.
About ScannerVision	Shows information about your ScannerVision installation.

6.1.2 Command Toolbar

The command toolbar hosts buttons that activates the most often used commands.

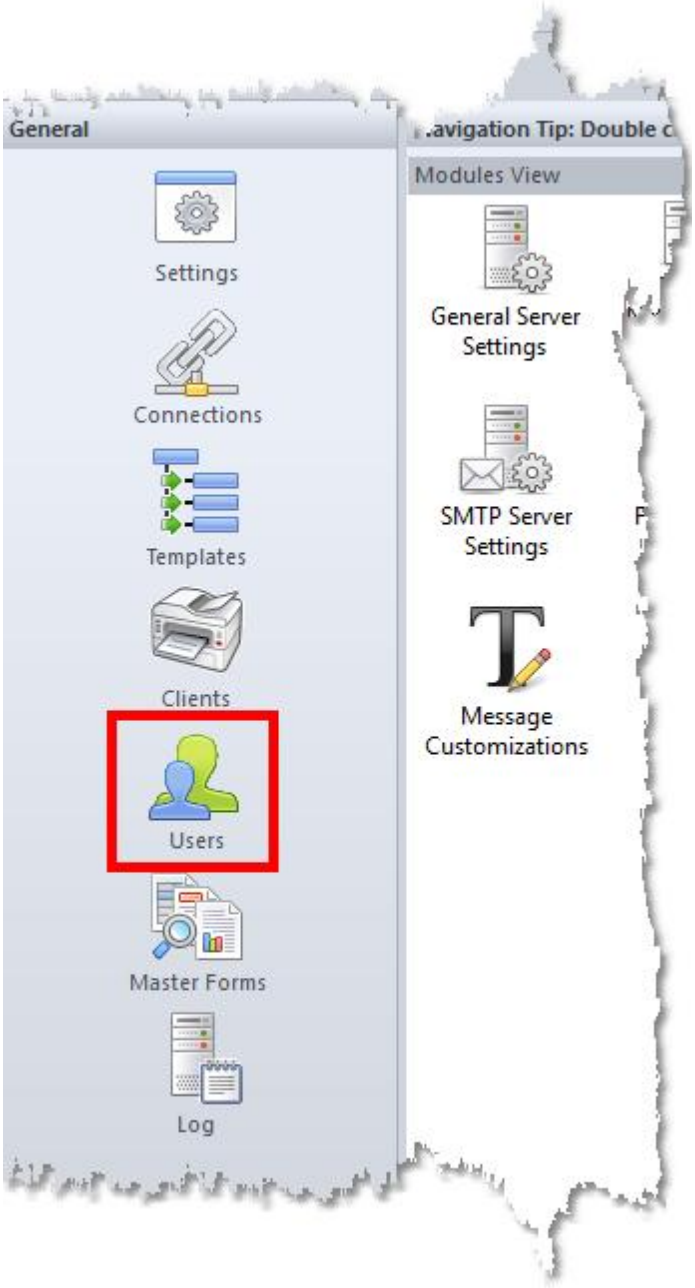


	Start the ScannerVision Processing Server and ScannerVision Networking Server services.
	Resumes the ScannerVision Processing Server service.
	Stop the ScannerVision Processing Server service only.
	Stop the ScannerVision Processing Server and ScannerVision Networking Server services.
	Adds a Forms Recognition Template.
	Adds a Workflow Template.
	This button is context sensitive so depending on in which section of the application you are its function will change. For example if you are in the Templates section the selected template will be deleted and if you are in the clients section the selected client will be deleted.
	Saves your configuration changes. Be sure that you save your changes before you leave the section of the application you are in.
	Discard any configuration changes since your last save.
	Navigates to the Settings pane.
	Opens this help file.

	Shows the Professional Services dialog.

6.1.3 Navigation Pane

The Navigation Pane is used to navigate to the Settings, Templates, Clients, Users and Log screens.

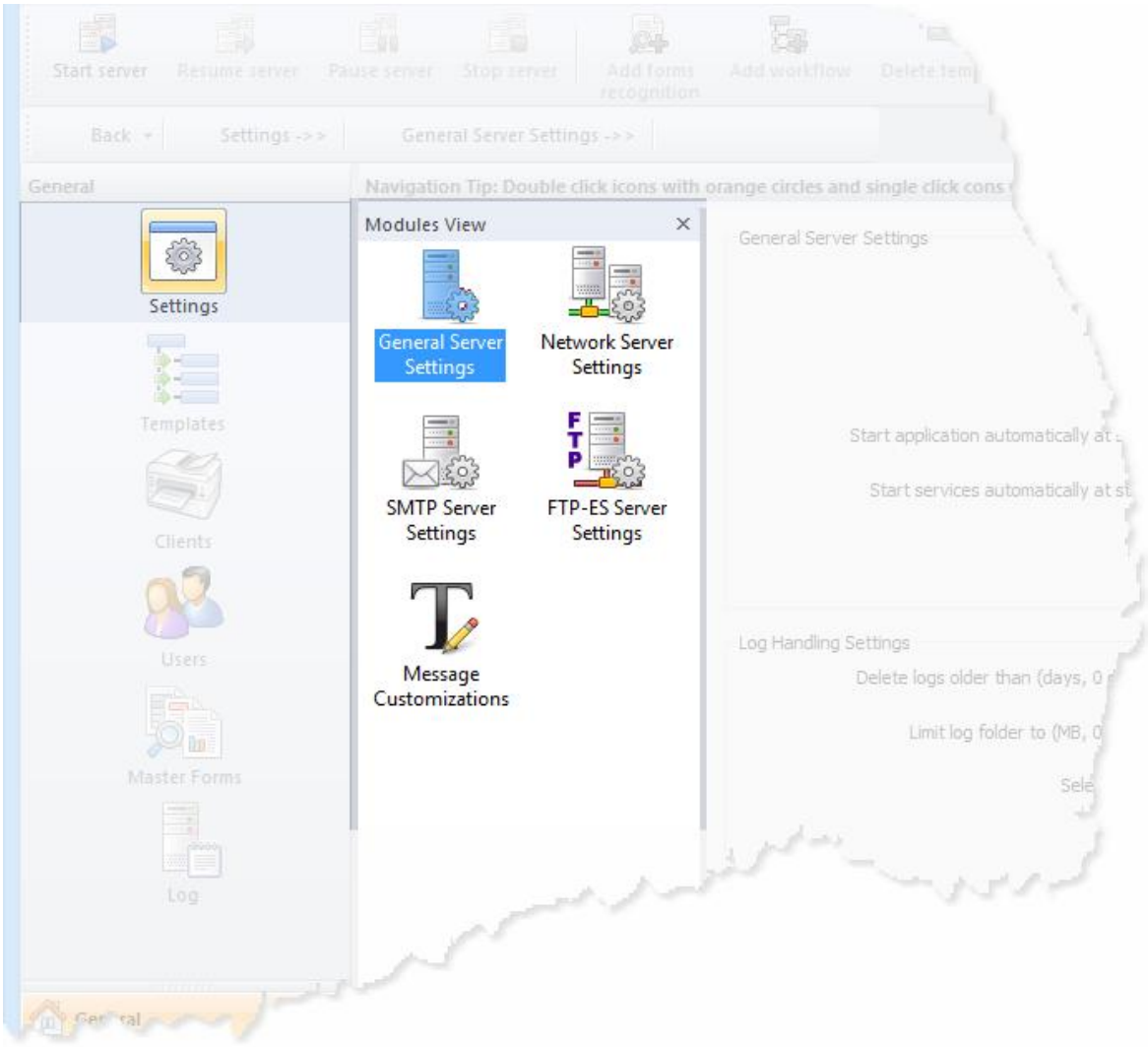


	Navigates to the Settings screen.
	Navigates to the Connections screen.
	Navigates to the Templates screen.
	Navigates to the Client and Client Groups screen.
	Navigates to the Users and User Groups screen.
	Navigates to the Automatic Forms Recognition screen.
	Navigates to the Log window.

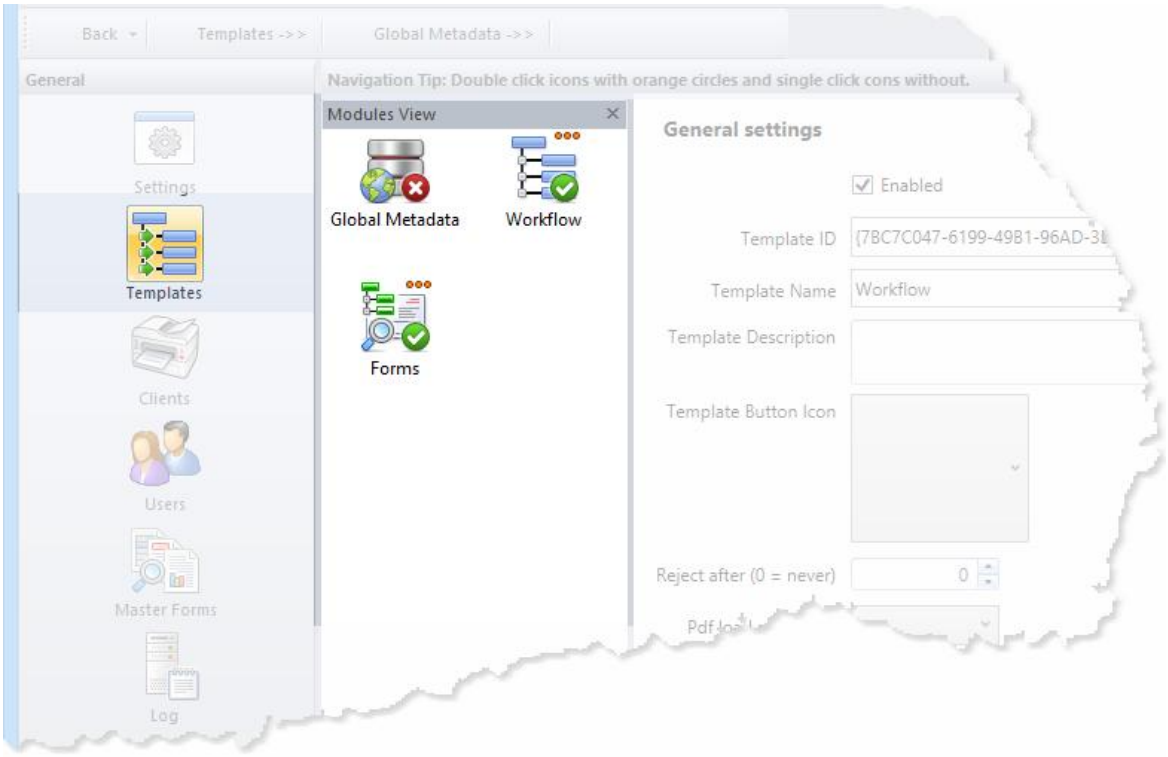
6.1.4 Module View

The Modules View changes depending on the option selected from the Navigation Pane.

In the example below the Settings option has been selected which displays the general server settings screen.

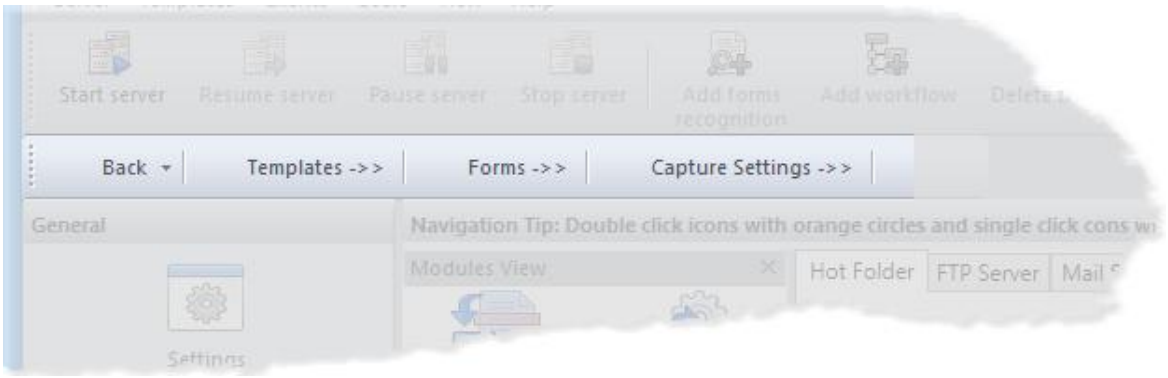


Clicking on the Templates option in the Navigation pane will display the existing templates in the Modules View window.



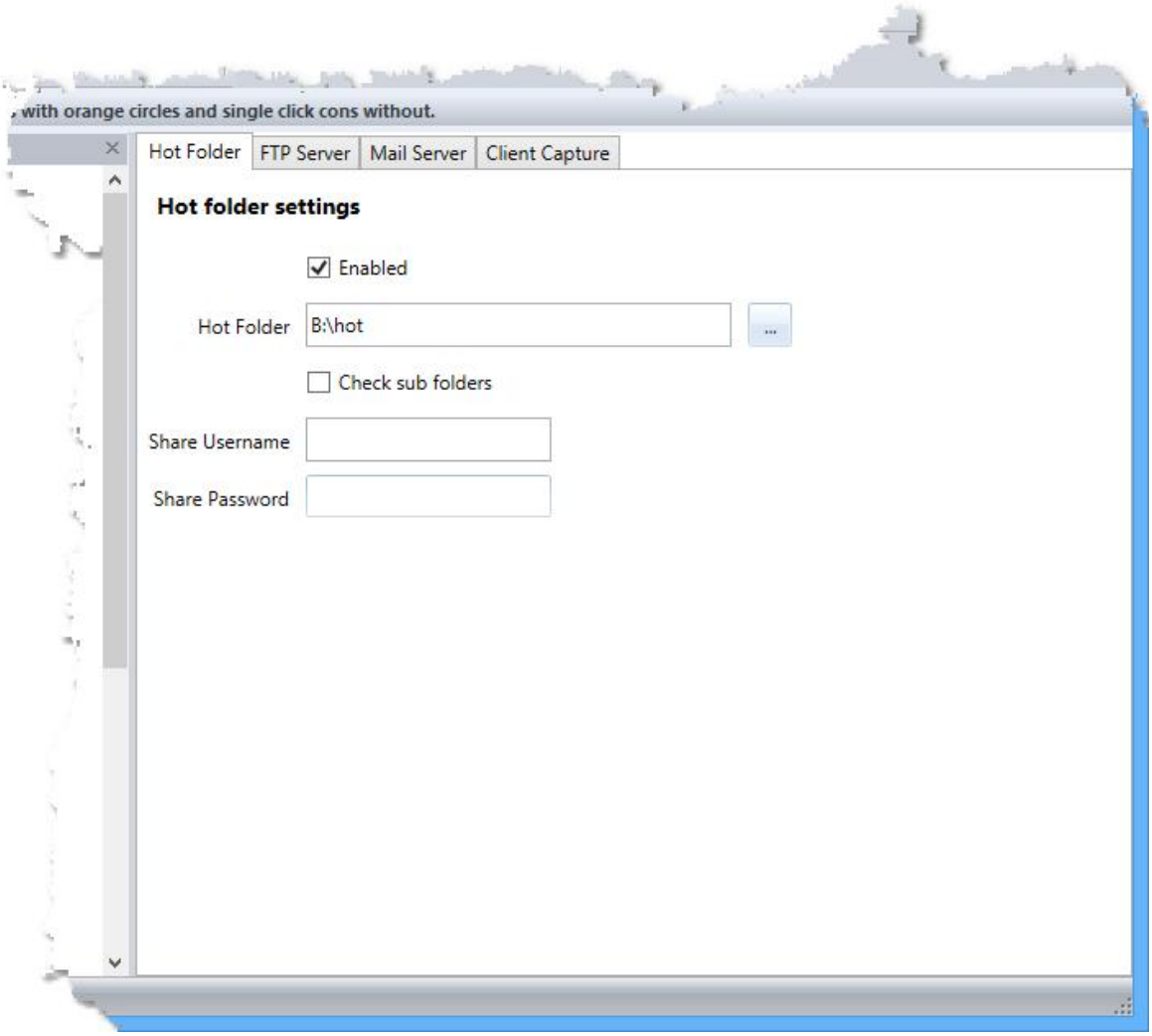
6.1.5 Navigation Toolbar

The navigation toolbar allows you to navigate up the settings hierarchy to any level you want.



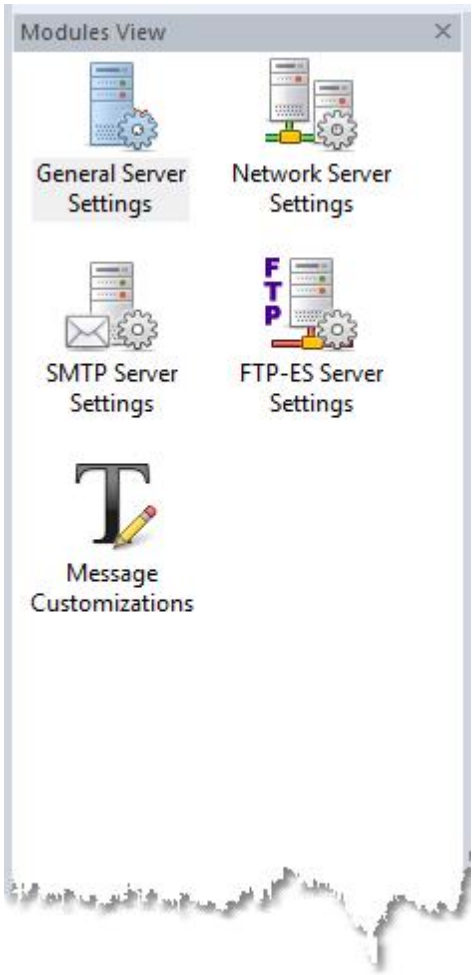
6.1.6 Configuration Pane

The Configuration Pane shows the available settings for the particular options you have selected in the modules bar.



6.2 Settings

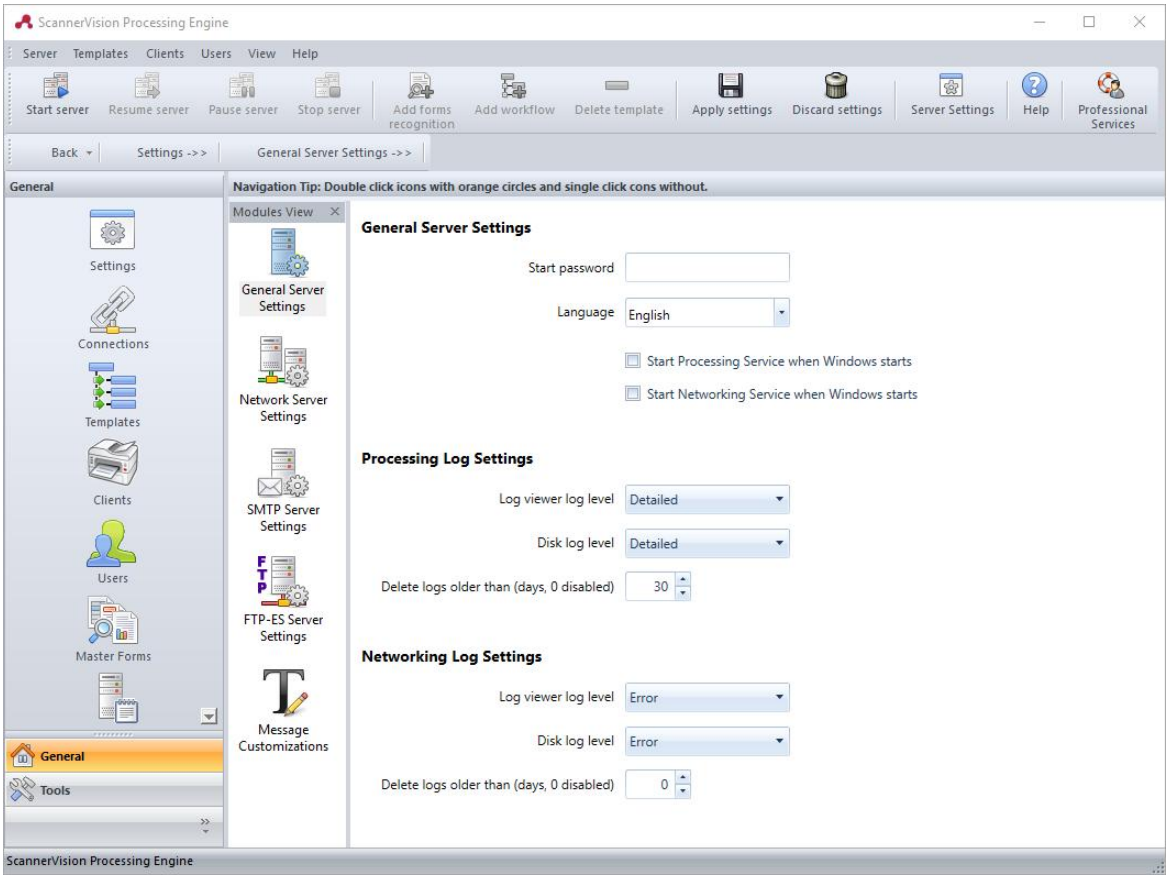
ScannerVision settings apply to ScannerVision in general and are not specific to any template, user or client.



Settings are grouped as follows:

- General Server Settings
- Network Server Settings
- SMTP Server Settings
- FTP-ES Server Settings
- Message Customizations

6.2.1 General Server Settings



Start password

If you specify a password the Processing Engine User Interface will prompt the user to enter the password when it is launched in order to use the application.

Note

This password applies only to the Processing Engine User Interface application and not the ScannerVision Processing Server and ScannerVision Networking Server.

Language

Select the language you want the application and log messages to be presented in. You have to restart the application after you have made a language change.

Start Processing Service when Windows starts

When selected the ScannerVision Processing Server will start automatically when Windows starts.

Start Networking Service when Windows starts

When selected the ScannerVision Networking Server will start automatically when Windows starts.

Processing/Networking Log Settings

Processing (this includes all logging other than Networking) and Networking logs can be configured independently. In addition, you have independent control over the log level of the log viewer and logging to disk. Options include:

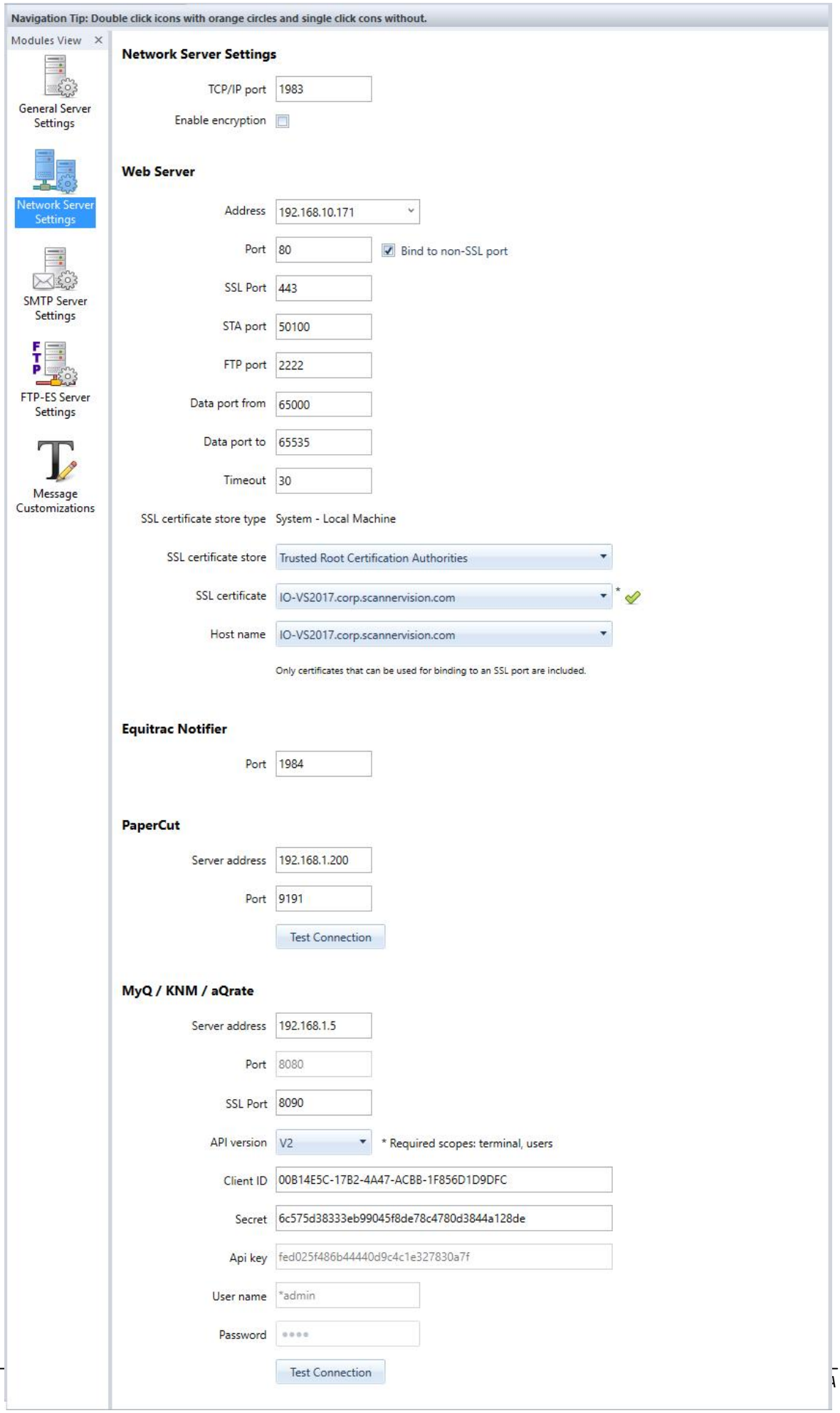
Errors	Log only errors.
Warning	Log errors and warnings.
Info	Log errors, warnings and informational messages.
Detailed	Log everything.
Off	Logging is disabled.

When logging to disk is enabled, you can have log files deleted automatically if they are older than the number of days you have specified by specifying a number larger than 0 (zero) in the "Delete logs older than..." control. If this value is set to 0 (zero), no log files will be deleted.

6.2.2 Network Server Settings

The ScannerVision Networking Server communicates with clients using a proprietary TCP/IP protocol that has been optimized for speed and that add very little overhead to the data being transferred. The ScannerVision Networking Server binds to port 1983 by default but you can change this by specifying a different port.

Communication can also be encrypted between clients and the server. When encryption is enabled on the server, communication to all clients is encrypted. If encryption is disabled on the server, clients can still be configured to encrypt communication in which case the communication between the particular client and server will be encrypted.


Navigation Tip: Double click icons with orange circles and single click cons without.
Modules View
General Server Settings
Network Server Settings
SMTP Server Settings
FTP-ES Server Settings
Message Customizations
Network Server Settings
TCP/IP port 1983
Enable encryption
Web Server
Address 192.168.10.171
Port 80
Bind to non-SSL port
SSL Port 443
STA port 50100
FTP port 2222
Data port from 65000
Data port to 65535
Timeout 30
SSL certificate store type System - Local Machine
SSL certificate store Trusted Root Certification Authorities
SSL certificate IO-VS2017.corp.scannervision.com
Host name IO-VS2017.corp.scannervision.com
Only certificates that can be used for binding to an SSL port are included.
Equitrac Notifier
Port 1984
PaperCut
Server address 192.168.1.200
Port 9191
Test Connection
MyQ / KNM / aQrate
Server address 192.168.1.5
Port 8080
SSL Port 8090
API version V2
* Required scopes: terminal, users
Client ID 00B14E5C-17B2-4A47-ACBB-1F856D1D9DFC
Secret 6c575d38333eb99045f8de78c4780d3844a128de
Api key fed025f486b44440d9c4c1e327830a7f
User name *admin
Password
Test Connection

TCP/IP port

The port number on which the ScannerVision Networking Server binds.

Enable encryption

Select this option if you want encryption between all clients and the server to be encrypted.

Note

The ScannerVision network protocol implements a hand shaking mechanism during which the server and client transmits a few bytes of data in an unencrypted form. This is always the case regardless of whether or not encryption is enabled on either end. The data that is transmitted contains no sensitive information such as user names or password.

The data that is sent by the client (which always initiates the communication) is:

Client protocol version number:	1 byte. A number indicating which version of the protocol is implemented by the client. The ScannerVision Networking Server supports multiple versions.
Client identifier:	1 byte. A number identifying the client as a Desktop client, Ricoh MFP client etc.
Force encryption flag:	1 byte. A flag that tells the server to encrypt all data after the handshaking has completed - regardless of whether the "Enable encryption" option on the server has been disabled.
Client host name:	Unicode character array of varying length - depending on the length of the client's network host name.

The data that is sent by the server in response to the client is:

Force encryption flag:	1 byte. A flag that tells the client to encrypt all data after the handshaking has completed.
Authentication type:	1 byte. A flag that tells the client what type of authentication challenge to present to the user.
Minimum supported protocol version:	1 byte. A number indicating the minimum protocol version that the server will accept.
Maximum supported protocol version:	1 byte. A number indicating the maximum protocol version that the server will accept.

Web Server Address

The IP address to which web based MFP devices will connect. Select an address that is visible on you

LAN. Do not select an address that is local to the server.

Web Server Port

The TCP/IP port on which the ScannerVision Web Server binds.

Bind to non-SSL port

Enable/disables non-secure communication between MFP clients and the ScannerVision Networking Server.

Web Server SSL Port

The TCP/IP port on which the ScannerVision Web Server binds for secure connections. This is used by the Connections web page and when web clients use the "https" protocol.

STA Port

The TCP/IP port on which the ScannerVision Kyocera Web Server binds. Kyocera MFP establishes a connection to the ScannerVision Kyocera Web Server before they connect to the ScannerVision Web Server.

Web Server FTP

This ftp server is used by MFP's to upload scanned documents to the ScannerVision Networking Server.

FTP Port

The TCP port on which the FTP server will bind. The default port for FTP is 21 but it could be any **available** port up 65536. The FTP server will bind to the specified port for all active IP addresses on the system. So if you have 2 network cards in the server the FTP server will be reachable on both addresses.

Data port from and to

The FTP clients on the Samsung MFPs run in passive mode which is to say that the clients initiate the connection to both the command and data channels of the ftp server. This means that both the command and data ports on which the FTP server binds need to be opened in the server's firewall software. If you want to limit the range of ports that you have to open in your firewall you can do so by limiting the range of data ports the FTP server is allowed to bind on by specifying the upper and lower limit in the respective edit boxes.

Timeout

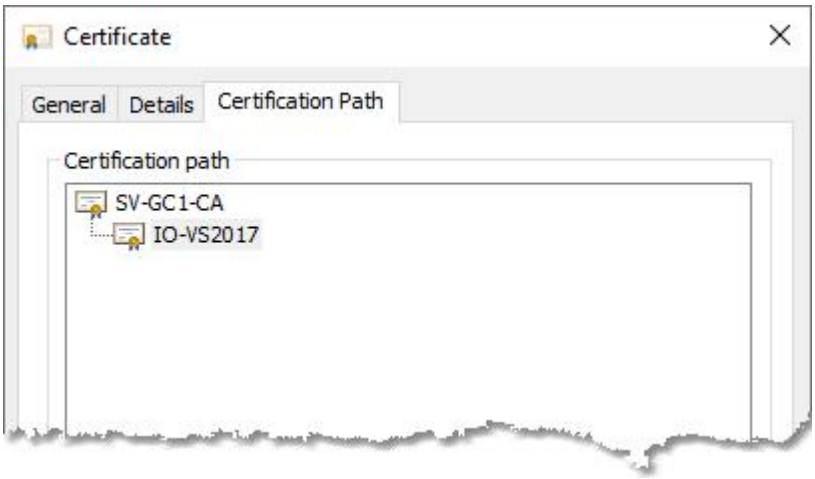
The timeout used by MFPs to connect to the ScannerVision Networking Server.

Secure network communication

In order to secure network traffic to and from the ScannerVision Networking Server you have to enable SSL port binding and select a port to bind to. The default SSL port is 443. In order for the binding to succeed you have to provide an SSL certificate that was created for this purpose. ScannerVision can generate what is known as a "self signed certificate" which is one where the issuer and subject are the same entity. Depending on how strict the certificate validation is that the client enforces that is connecting to the ScannerVision Networking Server, this self-signed certificate may be all you need. Not all clients (e.g. some MFP brands) do not allow self-signed certificates but insist that the certificate be signed by a trusted Certificate Authority (CA). ScannerVision gives you the options to select such a certificate as described below.

When you configure ScannerVision for the first time, it will generate a self-signed certificate and install it in the "Local Computer->Trusted Root Certification Authorities" store and automatically select it.

It is essential to ensure that clients are able to connect to a DNS server to resolve the IP addresses of all the links in the Certification Path chain included in the certificate:



SSL certificate store type

This is an informational label that indicates in which storage type ScannerVision is expecting to find the SSL certificate.

SSL certificate store

Select the store in the which the certificate is installed that you want to use.

SSL certificate

Select the certificate you want to use for SSL communication.

Certificates that can be used for SSL communication have to meet specific criteria. Only the certificate that do are shown in the list.

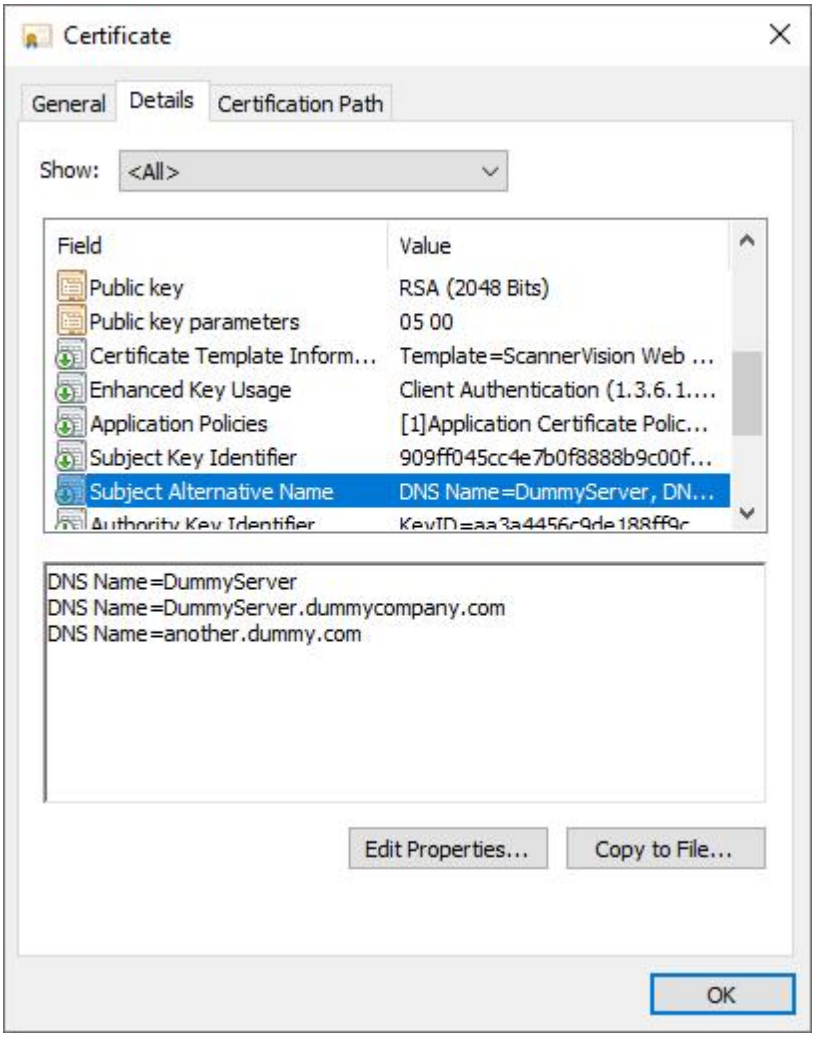
These criteria are:

1. It must not be a CA certificate, in other words it "Basic Constraints->Subject Type" must not be "CA" but "End Entity".
2. Key Usage must include:
   a. Digital Signature
   b. Key Encipherment
3. Extended Key Usage must include:
   a. Server Authentication
   b. Client Authentication

Automatic validation is performed on the certificate when it is selected. The result of which is indicated with a green check as indicated above if the validation was successful or with a red exclamation mark otherwise. To find out what the problem is, hover your mouse over the image. The tooltip will contain details of the validation failure.

Host name

When a client such as an MFP web client connects to the ScannerVision Networking Server over an SSL connection it will validate the certificate it received from the server as part of the SSL handshake protocol. Part of the validation process is to ensure that the host name of the server to which it connected matches one of the alternative name entries included in the certificate - see below. To ensure that this is the case, ScannerVision retrieves the list of alternative names from the selected certificate for you to choose from. Your selection will then be used in the URL that is deployed to MFPs.



It follows therefore that whenever there is a change in the certificate and/or host name details you will have to re-configure the host name here and redeploy all MFP clients that connect to this server. Some MFP client do not support Netbios lookups which mean you would not be able to use an entry with just the host name such as "DNS Name=DummyServer" above. In this case you would have to use "DNS Name=DummyServer.dummyserver" entry - which is the preferred one anyway.

If you encounter connectivity issues from a client make sure of the above and that the CA root certificate is installed on the client machine.

Equitrac Notifier Port

The port to which the Equitrac Notifier application connects.

PaperCut server address

The IP address of the PaperCut server. This is used when PaperCut authentication is selected on any

Client or Client Group.

PaperCut server port

The port of the PaperCut server. This is used when PaperCut authentication is selected on any Client or Client Group.

Test Connection

Press the "Test Connection" button to verify that a connection can be made to the PaperCut server.

MyQ / KNM / aQrate server address

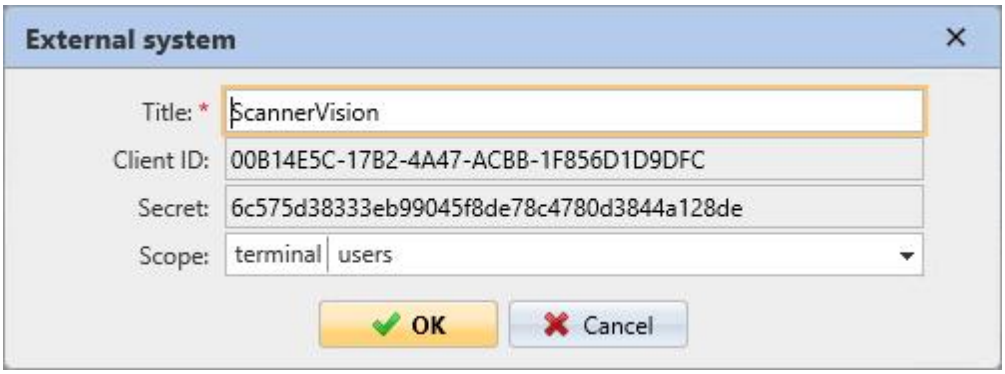
The IP address of the MyQ/KNM/aQrate server. This is used when MyQ/KNM/aQrate authentication is selected on any Client or Client Group.

MyQ / KNM / aQrate server ports

The SSL and non-SSL ports of the MyQ/KNM/aQrate server. These are used when MyQ/KNM/aQrate authentication is selected on any Client or Client Group. When connecting with API V2 SSL is compulsory.

MyQ / KNM / aQrate API version

The version of the web service API that is supported by your MyQ/KNM/aQrate installation. Currently only MyQ V7.4.1.30 or later supports API V2.



* Please ensure that you add "terminal" and "users" to the Scope selection.

MyQ / KNM / aQrate Client ID (API V2 only)

The Client ID that is generated by the MyQ/KNM/aQrate Extern System window shown above.

MyQ / KNM / aQrate Secret (API V2 only)

The Secret that is generated by the MyQ/KNM/aQrate Extern System window shown above.

MyQ / KNM / aQrate Api key (API V1 only)

The API key that is generated on the MyQ/KNM/aQrate server. Refer to the MyQ/KNM/aQrate documentation for further details.

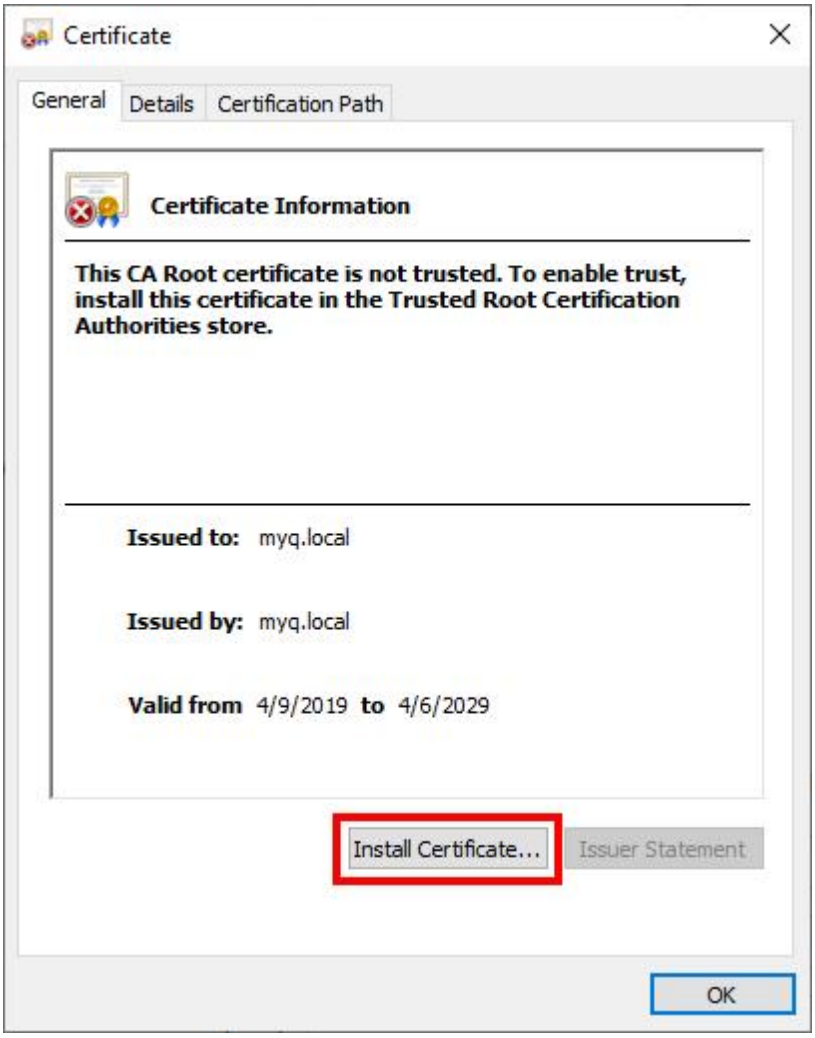
MyQ / KNM / aQrate user name & password

The user name and password the is required to connect to the MyQ/KNM/aQrate server.

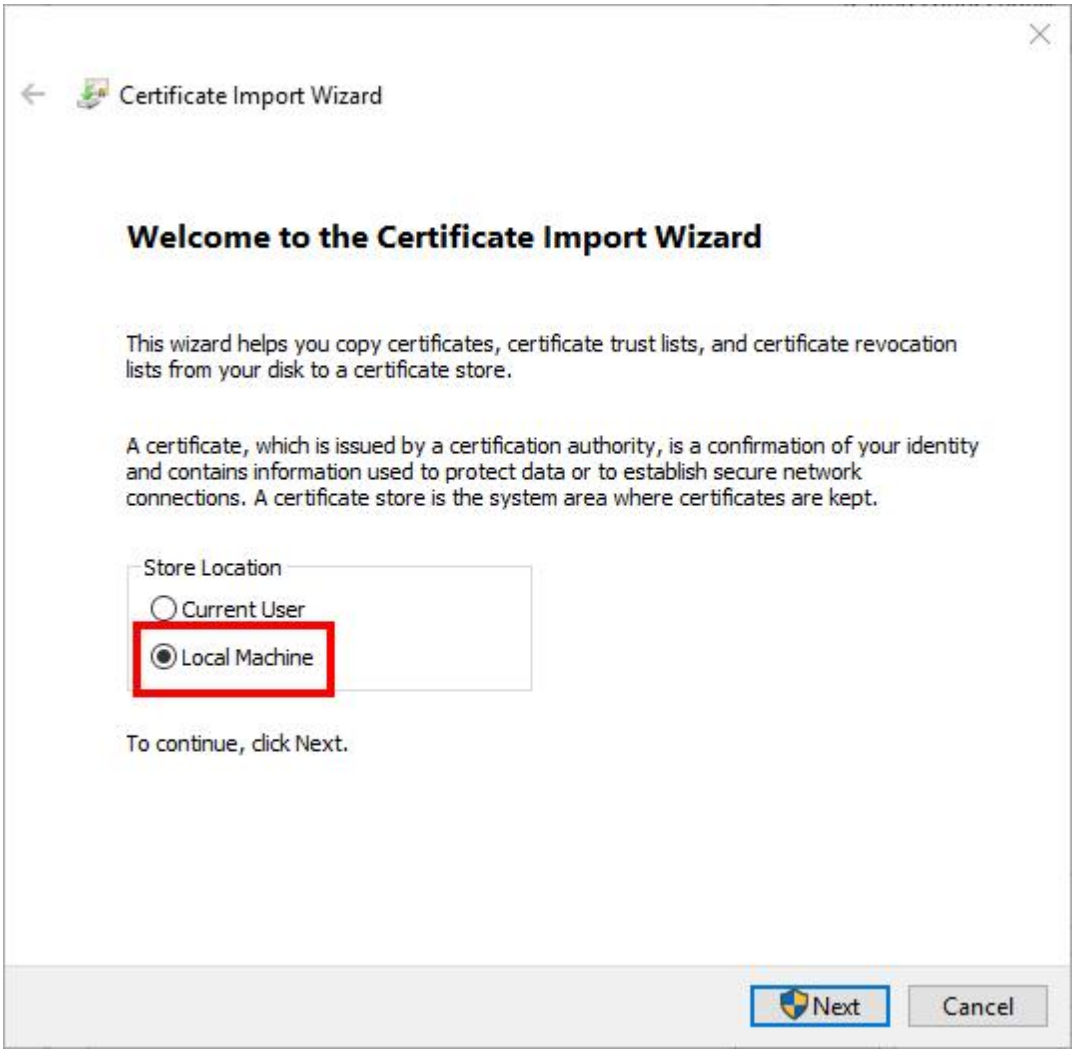
MyQ / KNM / aQrate Test Connection

Press the "Test Connection" button to verify that a connection can be made to the MyQ/KNM/aQrate server.

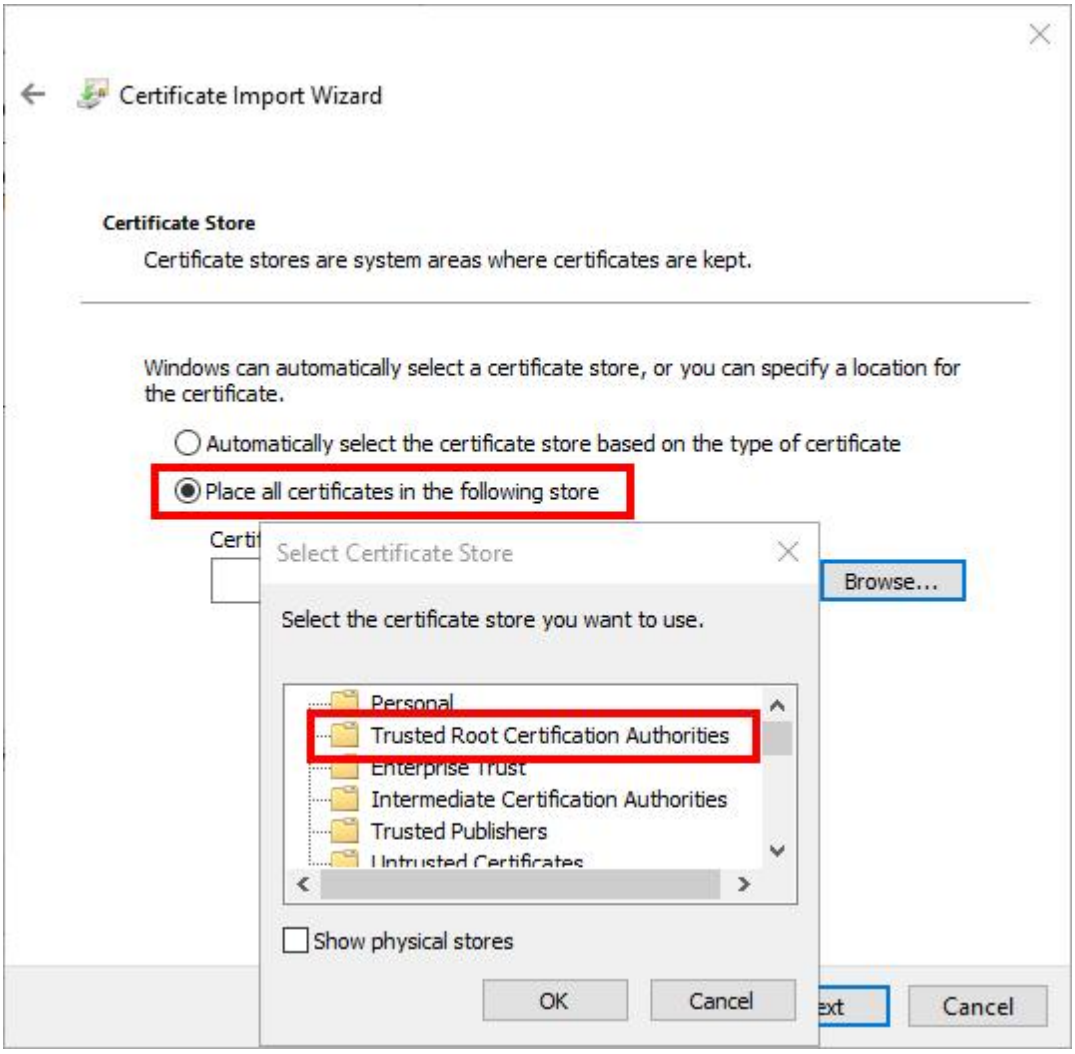
If this is the first time you are connecting to a MyQ server with API V2 selected you will be prompted to install an SSL certificate:



Click the "Install Certificate..." button. You will be presented with the following screen:



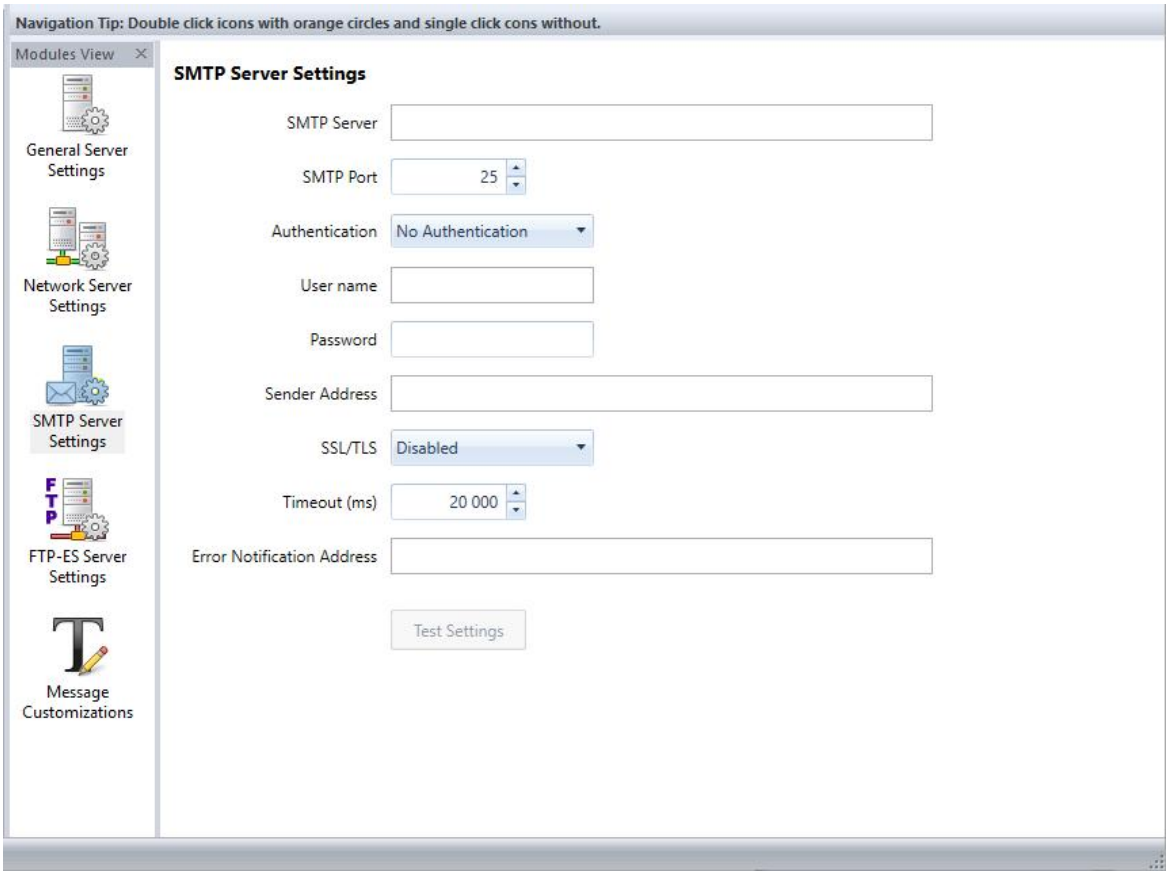
Select "Local Machine" and click "Next". You will be presented with the following screen:



Select the "Place all certificates in the following store" option and press the "Browse" button. On the popup that appears select the "Trusted Root Certification Authorities" option and press "OK". Continue with the remaining options until the certificate is installed.

6.2.3 SMTP Server Settings

ScannerVision can send email notifications out to administrators under certain scenarios.



SMTP Server

The IP address or host name of the email server.

SMTP Port

Port number of the email server - port 25 is the default.

Authentication

The authentication method required by the email server.

User name & Password

Authentication credentials to be used by ScannerVision when sending emails.

Sender Address

The return address to be put in the email.

SSL/TLS

Select the encryption option for communication between ScannerVision and the SMTP server.

Timeout (ms)

The time allowed for sending the email message.

Error Notification Address

The email address(es) to which notifications emails will be sent. Separate email addresses with a semi-colon " ; ".

The email address to which a test email is sent when the "Test Settings" button is clicked.

Test Settings

You can test that your SMTP server settings are correct by clicking the "Test Settings" button upon which a test email is sent to the provided email address.

When the "Use same settings as configured in main SMTP Server Settings" is checked the "Sender Address" is used as the "From" address and the "Error Notification Address" is used as the "To" address of the test email.

Note

When using Google's Gmail as your SMTP server it is recommended that you enable 2-step verification on the account with which you will be connecting and to set up an "App password".

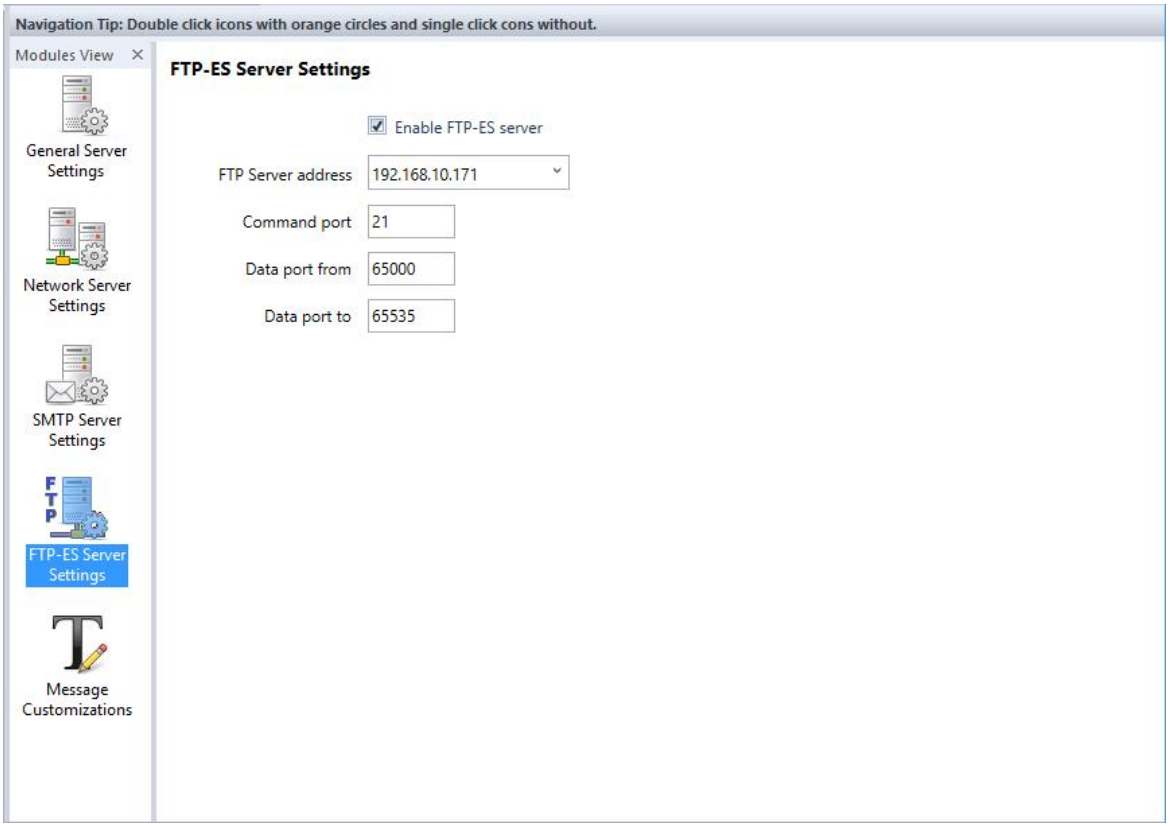
6.2.4 FTP-ES Server Settings

The FTP-ES Settings section is where you configure ScannerVision's built-in secure ftp server which is used by Kyocera MFP clients to upload scanned documents for processing. This should not be confused with a template's FTP capture section. The latter is an FTP server from where ScannerVision **downloads** documents for processing when the template runs.

The embedded Kyocera MFP client requires specific configuration settings and the ScannerVision FTP server has been built with this in mind. It is therefore important that the machines on which the Kyocera MFP client is installed are configured accordingly. Please see the Kyocera MFP Configuration section for details.

The "Enable FTP-ES server" check box is selected by default. To disable the ftp server de-select this check box and restart the server. NOTE if the ftp is not running Kyocera MFP clients will not be able to upload documents to ScannerVision.

The FTP server is part of theScannerVision Networking Server so whenever changes are made to the settings below you have to restart the ScannerVision Networking Server for the changes to take effect.



Enable FTP server

This check box enables the FTP server when the ScannerVision Networking Server is started.

FTP Server address

The address that will be returned to the MFP client. This address has no bearing on the address(es) to which the FTP server binds. It is only used to let the MFP client know to which address to connect to. This would normally be the same address as the ScannerVision server's but it could be different. When an MFP client is on a different network segment than the server and is not able to connect to the server directly, it would need to connect to a different address which will be forwarded to the server by a router. This true for the MFP client's own connection to the ScannerVision Networking Server as well the device's FTP client's connection to the FTP server on the ScannerVision server that receives scanned documents.

Command port

The TCP port on which the FTP server will bind. The default port for FTP is 21 but it could be any **available** port up 65536. The FTP server will bind to the specified port for all active IP addresses on the system. So if you have 2 network cards in the server the FTP server will be reachable on both addresses.

Data port from and to

The FTP clients on the Konica Minolta MFPs run in passive mode which is to say that the clients initiate the connection to both the command and data channels of the ftp server. This means that both the command and data ports on which the FTP server binds need to be opened in the server's firewall software. If you want to limit the range of ports that you have to open in your firewall you can do so by limiting the range of data ports the FTP server is allowed to bind on by specifying the upper and lower limit in the respective edit boxes.

6.2.4.1 Kyocera MFP Configuration

The embedded Kyocera client requires specific configuration settings in order to communicate with the ScannerVision FTP server. Please ensure that you have configured your Kyocera MFP using the COMMAND CENTER as follows:

Advanced -> Secure Protocols

Secure Protocol Settings

SSL - **On**

Clientside Settings

Effective Encryption - **Select at least one**

Certificate Verification - **Off**

Scanner -> FTP

FTP Settings

FTP - **On**

FTP Port Number - **This is retrieved from the** ScannerVision server

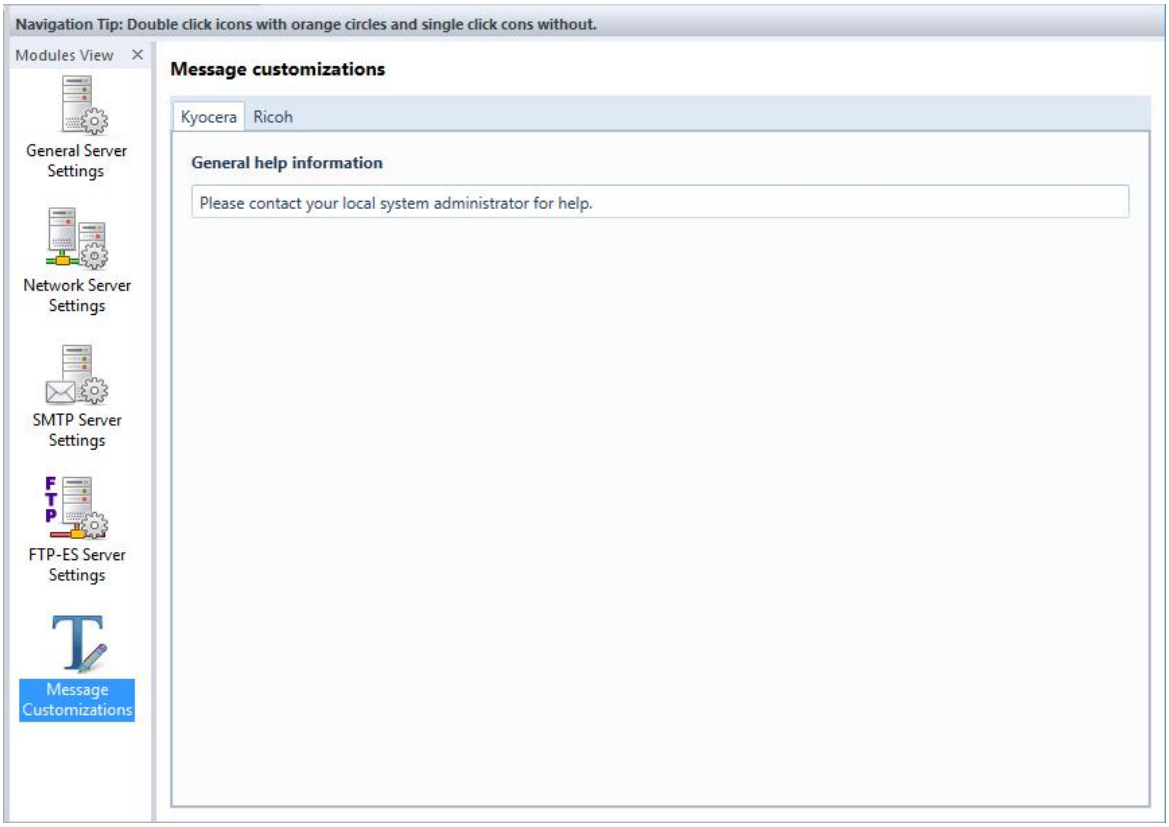
FTP Encrypted TX - **On**

6.2.5 Message Customizations

ScannerVision offers you the ability to customize some of the informational messages displayed to users on the clients such as when a help button is pressed. You can use this facility to provide information to the user that is relevant to your environment.

The messages that you are able to customize is specific to each client and are grouped in tabs with headings that identify the client. Each message has a heading that tells you where in the client the message appears. In the screen shot below you can see that the Kyocera client allows you to customize the message that the user sees when the help button is pressed.

Whenever you have made changes to any of these messages, the clients will retrieve them from the server the next time they connect.



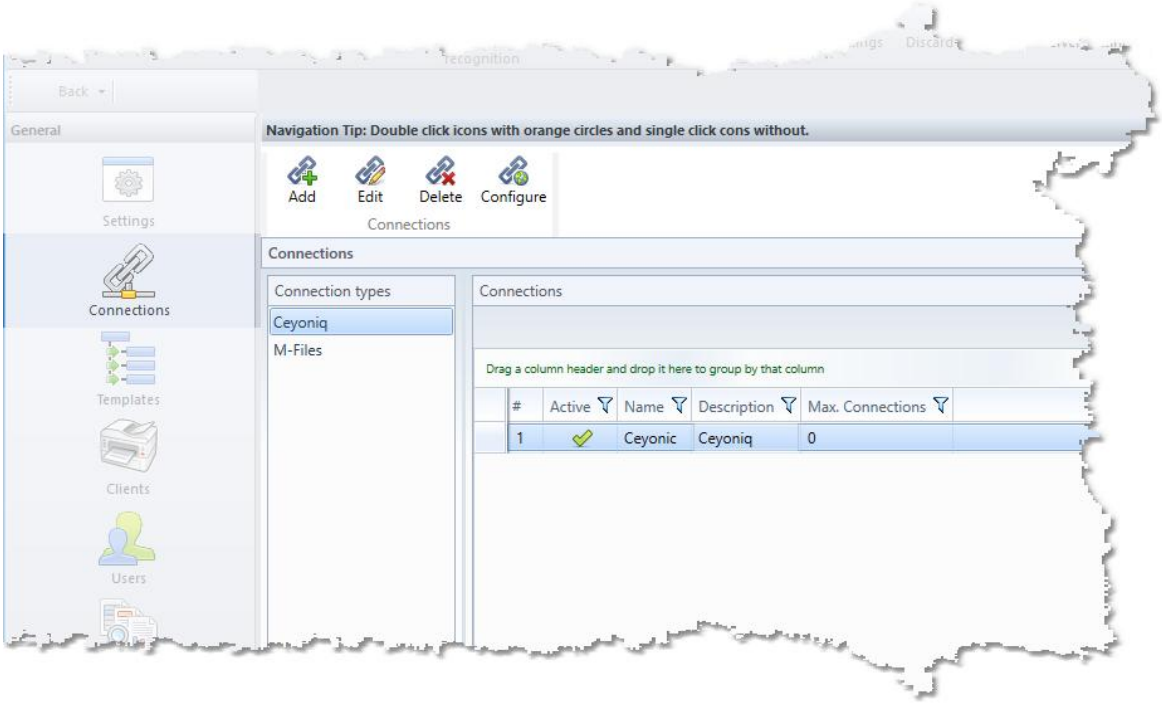
6.3 Configuring Connections

Connectors and Picklists generally require a server address and login credentials or a connection string in the case of a database. In the past you had to configure these parameters for every connector instance you added to a template or picklist you created - even if you were connecting to the same server and using the same credentials. Connections allow you to configure this once only and then to reference the connection in every connector and/or picklist, in any template, that needs a connection to the particular server. If you need to change the address of the server you want to connect to you can now do it in one place - on the connection level. All the connectors and/or picklists in all templates that reference the connection will automatically use the updated connection. In

addition, you can control which users and or groups are allowed to connect to the particular server through permissions. We cover permissions in the Permissions section.

Connections also facilitate the Single Sign On (SSO) capability of connectors and picklists. When a connector or picklist executes it is done with the identity of the user who is logged in at the MFP or Desktop client. When an uploaded document is processed by the ScannerVision Processing Server every connector defined in the template will process the document as the user who uploaded it.

The Connections screen is shown below. The entries in the "Connection types" list is determined by the connectors you have installed.



6.3.1 Authentication Tokens

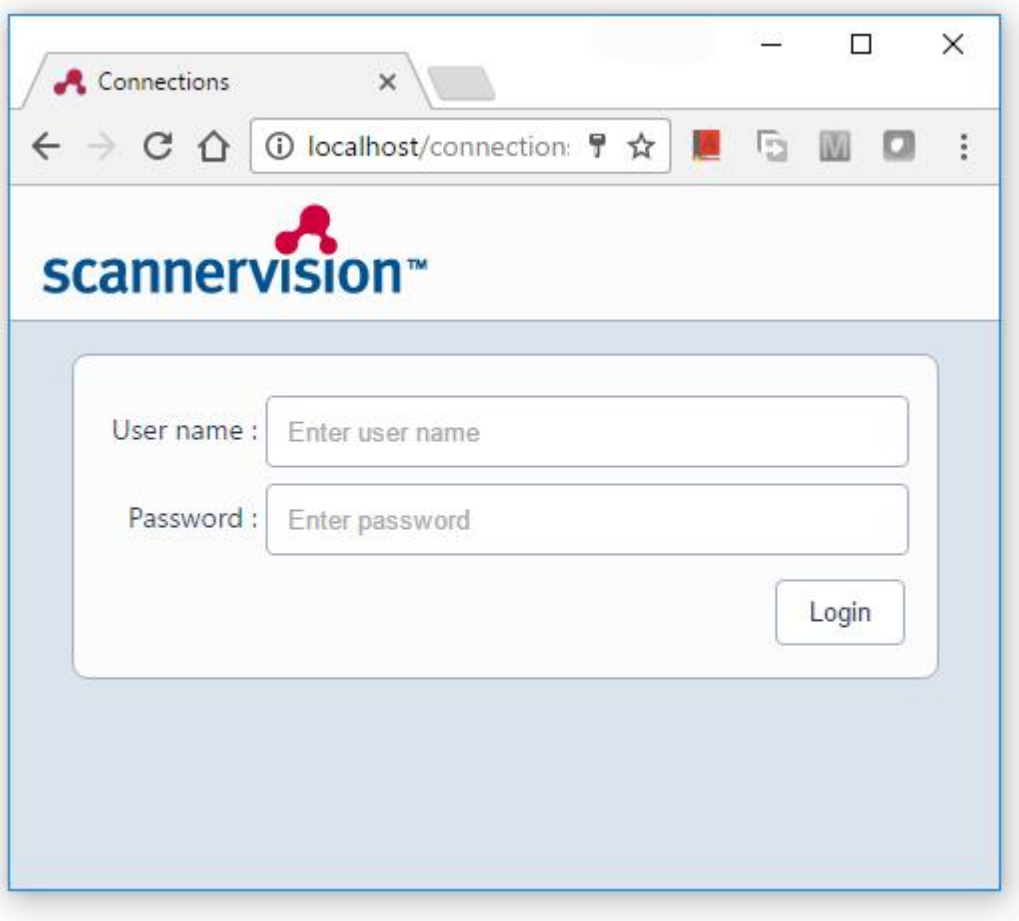
The authentication mechanisms employed by document management systems (DMS) vary widely. From a simple username/password combination to sophisticated OAuth authentication tokens. The ScannerVision's Connections system links the authentication information of a DMS to a ScannerVision user. After a connection has been configured and user permissions assigned, users have to provide ScannerVision with the authentication information needed to connect to the particular DMS. This is done through a web site that is hosted by the ScannerVision Networking Server.

The steps a user has to follow to provide ScannerVision with his/her authentication information are as follows:

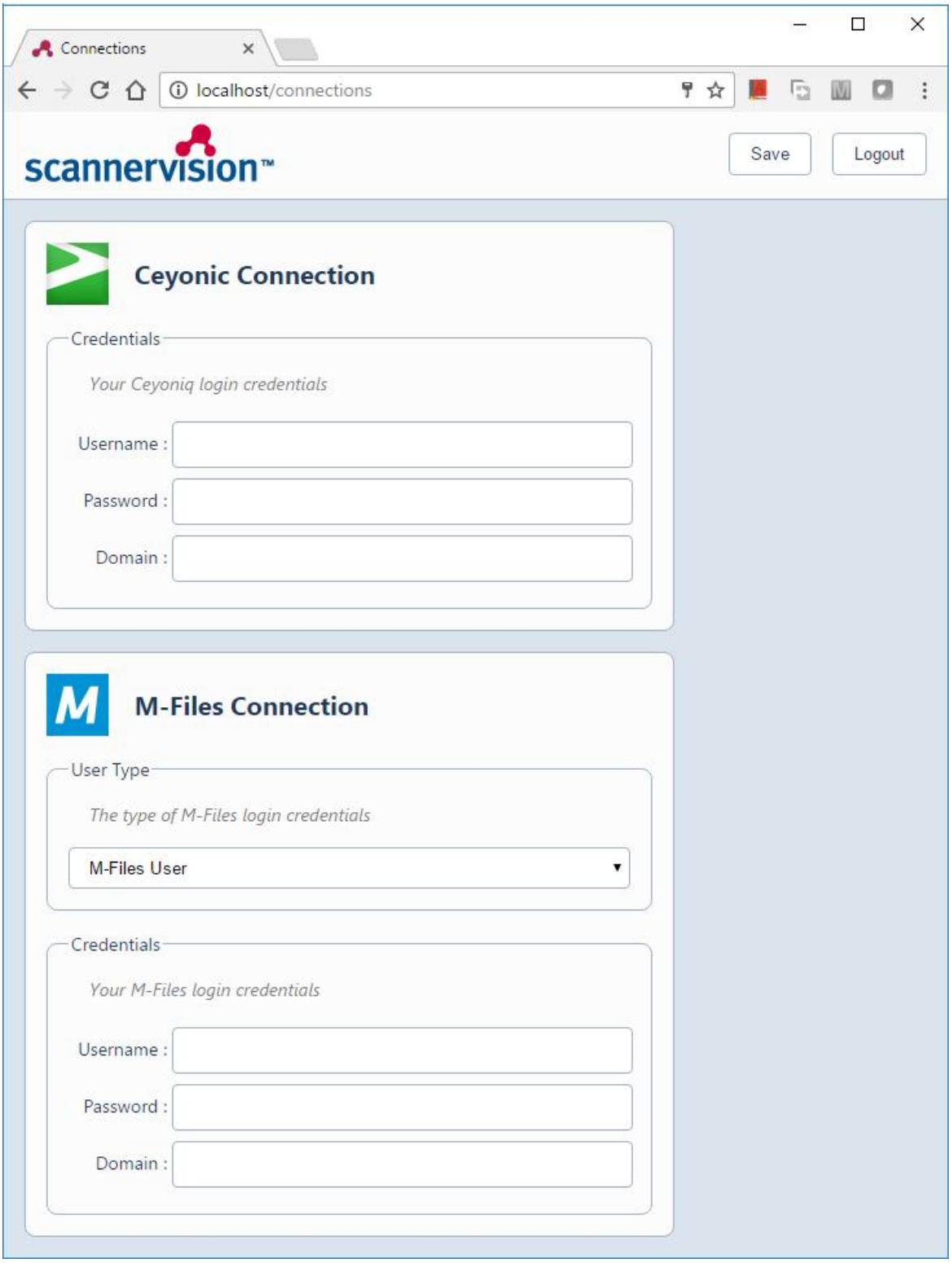
1. Open an Internet browser (Internet Explorer, Microsoft Edge, Google Chrome etc.).
2. Enter the following URL: https://[SERVER]:[PORT]/Connections where [SERVER] is the IP address or host name of the ScannerVision Networking Server and [PORT] is the SSL port you have configured in the Network Server Settings.
3. Log in to the web site that appears using your **ScannerVision** user name and password.

4. Complete the information that appears on the page that opens after you have logged in and click the "Save" button in the top right-hand corner.

On the login screen enter your **ScannerVision** username and password and click "Login".

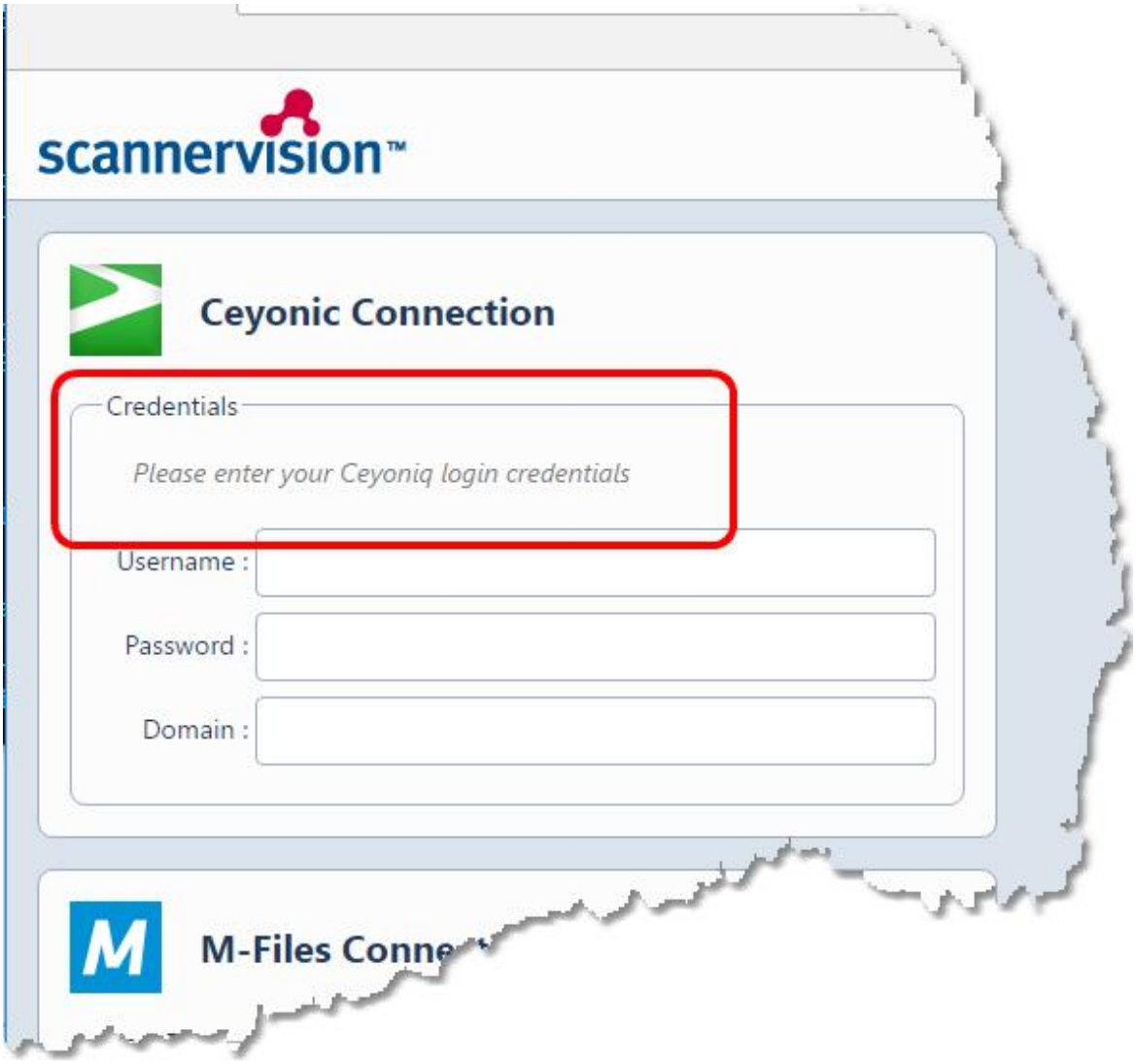


You should then see a web site similar to this:



Note

The number and type of entries you will see on the page is determined by the number and type of connections you have been configured in the Connections section. From the screen shot above it is evident that a Ceyoniq and M-Files connection have been configured. Further, the information users see may be different than the default for a connection type. For example you could configure your page to look like this:



You have the ability to customize certain aspects of the information that is presented on the "Connections" web page. This can be done when you Add or Edit the Connection.

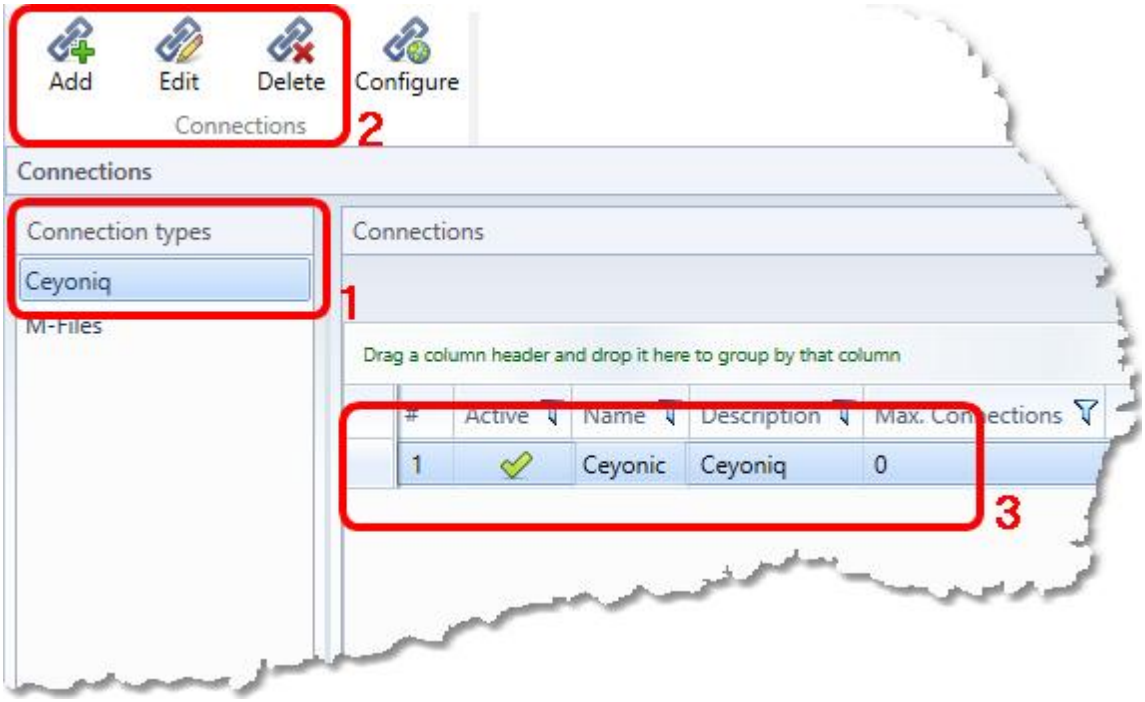
When you click the "Save" button the web page submits the authentication information you have entered to ScannerVision which then stores it securely in a database. The information is linked to your ScannerVision user account. This means that if you are authenticated in ScannerVision you are also authenticated in all the DMSes that you have provided credentials/authentication tokens for.

6.3.2 Adding/Editing Connections

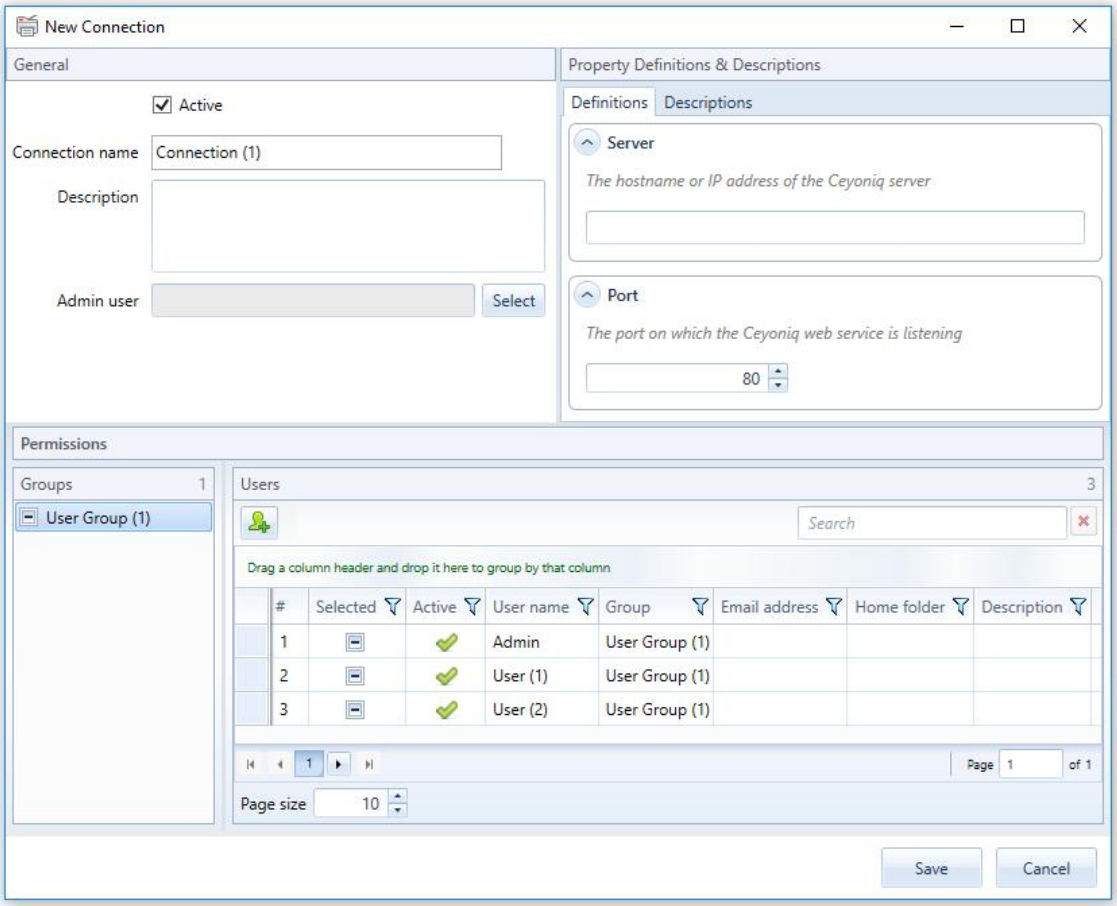
To create a new connection select the Connection type (1) then click the "Add" button (2).

To edit a connection select the Connection type (1) then double click the connection you want to edit (3) or click the "Edit" button (2).

To delete a connection select the Connection type (1) then click the "Delete" button (2).



The Add/Edit connection screen is shown below:



The screen has 3 sections namely "General", "Property Definitions & Descriptions" and "Permissions".

General

Active — Enables/disables the connection. When a connection is disabled no connector or picklist that references the connection will be able to connect to the respective back end system.

Connection name — The name of the connection. This name appears in the "Connections" grid on the "Connections" screen as well as on the "Connections" web site where users configure their credentials. It is a good idea to give your connections meaningful names that users will understand.

Description — This allows you to give a short description of the connection for your own or other administrators' benefit.

Admin user — The ScannerVision user who's connection information will be used when configuring connectors and picklists. This user needs to be given permission to use the connection. If authentication is not enabled on a Client, this user's credentials/authentication token is

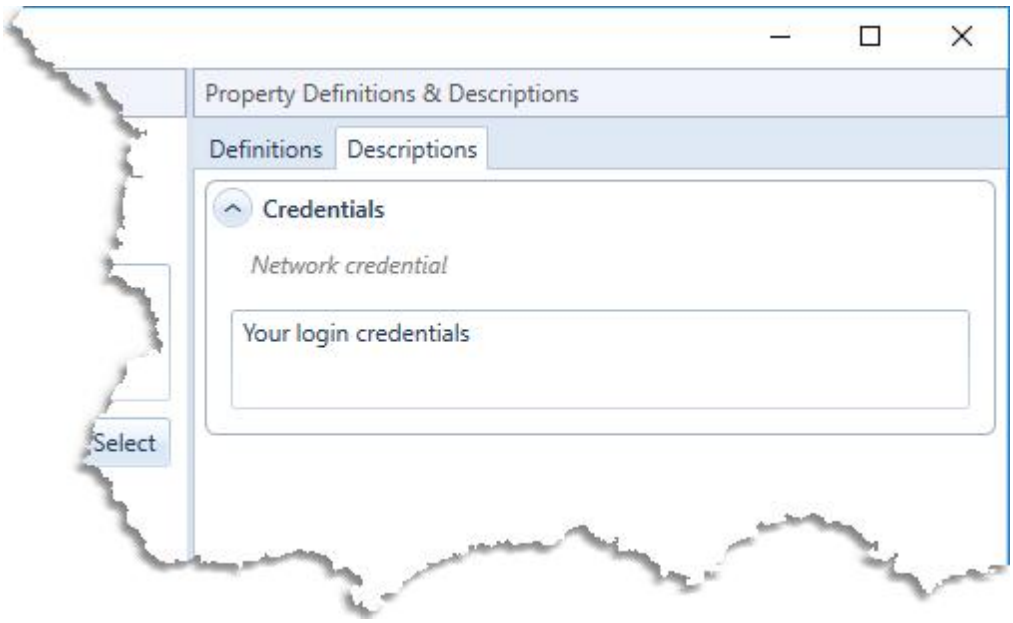
used for the connection when executing connectors and picklists.

Property Definitions & Descriptions

The "Property Definitions & Descriptions" section contains 2 tabs called "Definitions" and "Descriptions". The contents of both these tabs are determined by the particular Connection type.

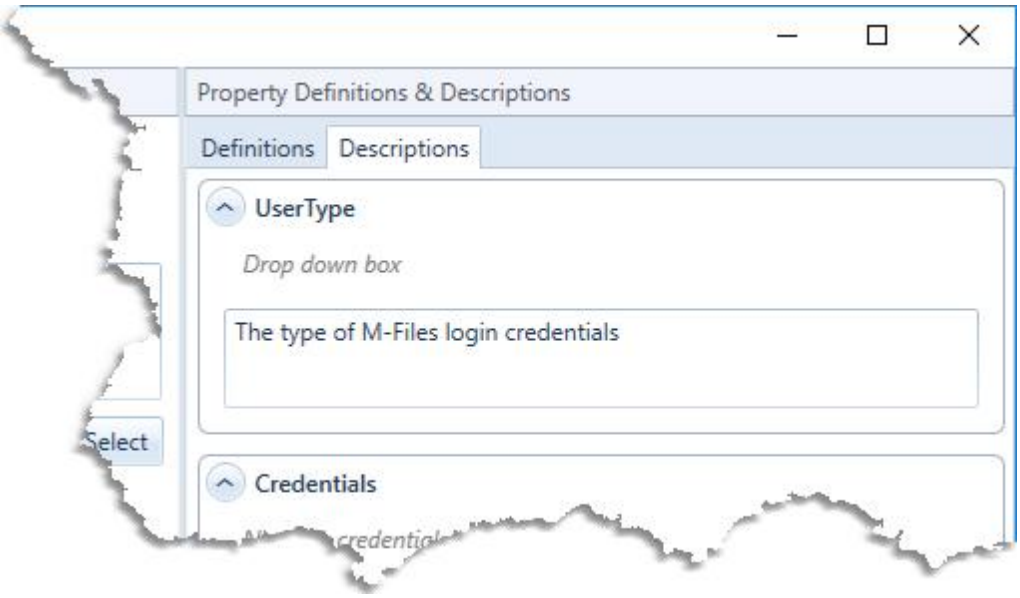
The "Definitions" tab contains controls to take on the information that is required to connect to the particular back end system. This could be a URL, server address or host name, port, database connection string etc.

The "Descriptions" tab lists the elements that will be shown on the "Connections" web site. In the screen shot below is listed and element called "Credentials" and it is of type "Network credential".

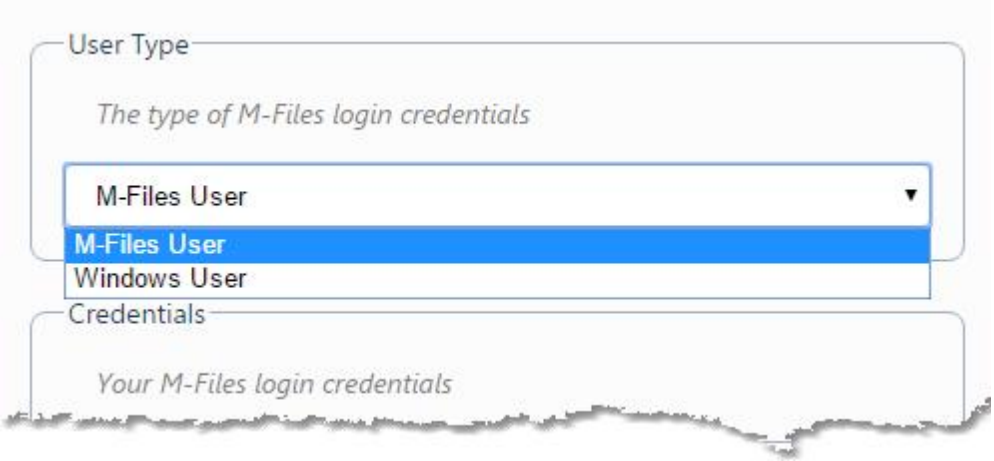


Below is a screen shot of the content that is shown on the "Connections" web page for a network credential type element.

An element of type drop down:



Presents as a drop down box on the web page:



The types that are available and how they are presented are listed in the table below.

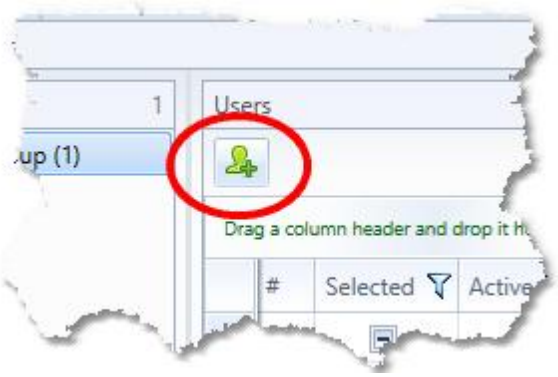
Type	Presentation
Boolean	Check box
Text	Text box
Number	Text box or numeric up/down edit box (browser dependent)
Network credential	Text box for user name, password box, text box for domain
Drop down box	Drop down box

Permissions

Please refer to the Permissions section.

Add User

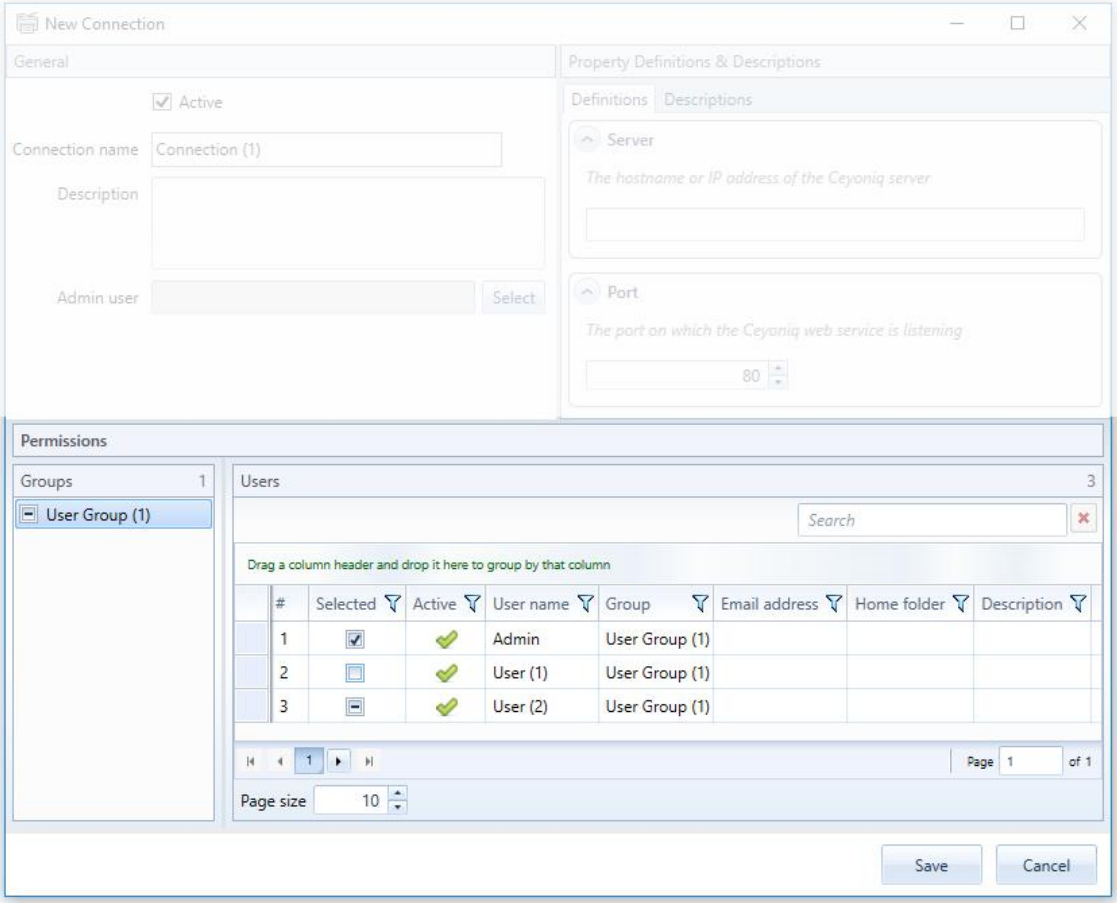
You can add a user by click the button shown below. This is useful if you have not added the admin user yet that you need to assign to the connection.



6.3.3 Permissions

Connection permissions allow you to control which users and groups are allowed to use the connection. If a user or the group to whom the user belongs is not explicitly granted permission to use a connection, the user will not be allowed to connect to the respective back end system.

Permissions are configured on the New/Edit Connection screen:



Permission is granted or denied to users or groups through the check boxes next to the user or group. The check boxes have three states namely "Checked", "Unchecked" and "Undefined". When the check box has a check mark such as the "Admin" user in the screen shot above permission is **granted.** If the check box has no mark such as "User (1)" permission is **denied.** If the check box has a horizontal line such as "User (2)" then permission is neither granted nor denied and the permission set on the group level is applied.

In the screen shot above you will notice that the permission of both 'User Group (1)" as well as "User (2)" is undefined. In this scenario "User (2)" will not be able to connect to the back end system since permission has not been granted to either the user or the group.

User level permission supersedes group level permissions. The "Admin" user above will be able to connect while "User (1)" will not regardless of whether permission has been granted or denied on the group level.

6.4 Configuring Clients

Clients in ScannerVision are devices or applications that capture documents and metadata and then uploads them to the ScannerVision Networking Server. Clients fall into one of two categories namely Desktop Clients and MFP Clients. Desktop clients are Microsoft Windows applications that run on computers with one of the supported Microsoft Windows operating systems installed. MFP clients run on supported multi-functional devices from manufacturers such as Ricoh, Kyocera, Samsung, Konica Minolta and more.

Desktop clients

Desktop clients offer two modes of operation namely Twain scanning and loading of existing documents from disk or the network. The Twain scanning mode of operation allows the user to scan documents from any Twain source that is installed on the system. Most of the often used Twain scan settings can be controlled by the ScannerVision template. This assures consistency across multiple desktop clients and Twain sources. In the second mode of operation the user loads existing documents from the client application. The two methods are not mutually exclusive. It is possible for example to scan a document using a Twain source and then to load one or more existing documents from disk. When the document is uploaded to the ScannerVision Networking Server it becomes a single document.

MFP clients

MFP clients are applications that run on multi-functional devices that offer the ability to install third party applications on the device and for which a ScannerVision client has been developed.

Configuration

There is no distinction between the configuration of a desktop client and an MFP client in ScannerVision apart from selecting the type of client that you want to configure. Clients belong to one or more Client Groups. It is up to you to decide how you want to group the clients you configure. Your decision may be based on the physical location of the devices in your organization, the department in

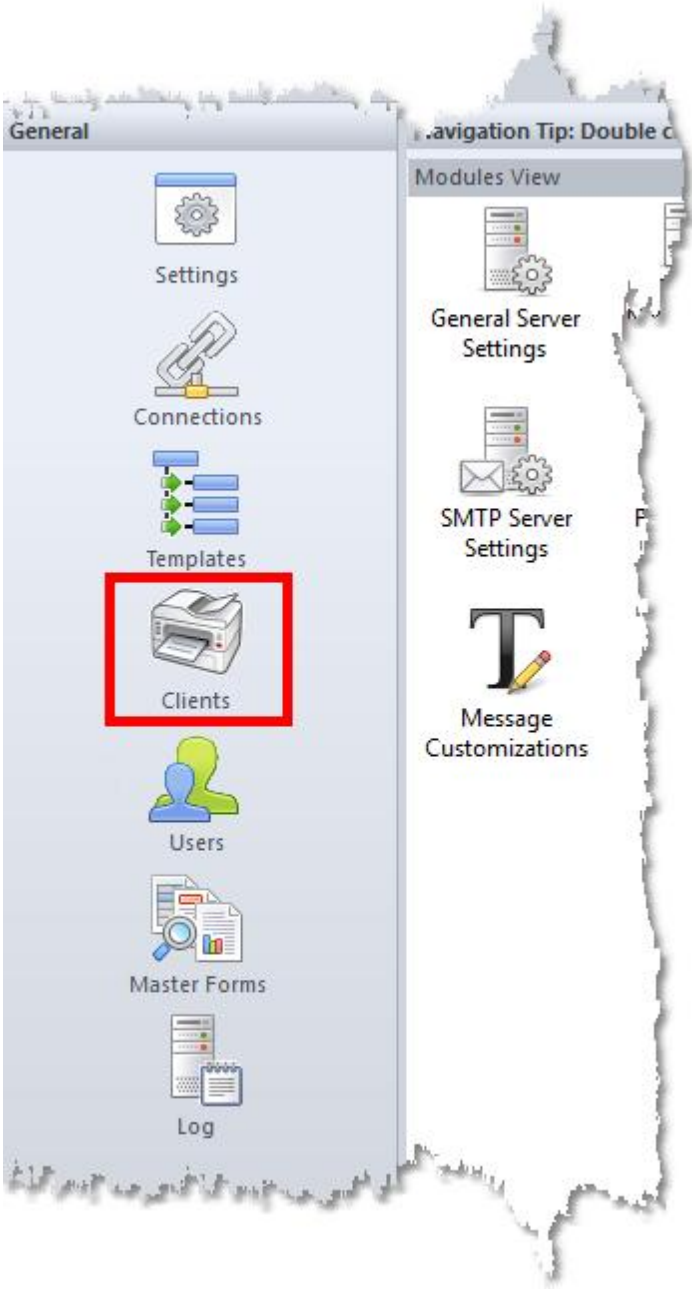
which they will be used, the templates that will be available on them or whatever other grouping criteria you may wish to apply.

When a user wants to submit a document to the ScannerVision server she has to do it in the context of a template. The template defines the entire work flow around a document and as such form the foundation of the document capture process. The templates that are available on an MFP or desktop client could be determined by various configuration options as is discussed in the Overview section.

One of the ways you can control which templates are available on clients belonging to a group, is to assign them on a group level. This makes life a little easier when it comes to deciding which templates should be available on which clients. You do of course also have the option to configure clients individually. When you want to control clients individually you have to manually configure all the clients that belong to the particular group. You can't configure only certain clients in the group individually while others are controlled on the group level.

If a client belongs to multiple groups and group permissions apply on all the groups, the combination of templates defined on the groups will be available on the client. In the scenario where a client belongs to two groups one of which has group permissions applied and one not, then the templates defined on the group with group permission enabled as well as any defined on the client level will be available on the client.

To configure clients press the "Clients" button in the navigation pane shown below:



6.4.1 Client Groups

Clients can belong to one or more groups or none at all. Putting clients in groups means that you can manage them on a group level. There are three aspects which can be controlled on a group level:

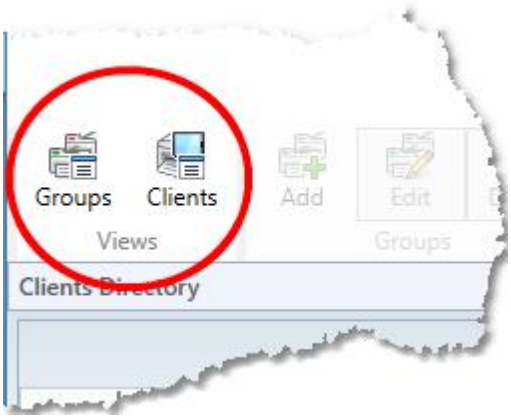
1. Activation/Deactivation

2. Authentication method

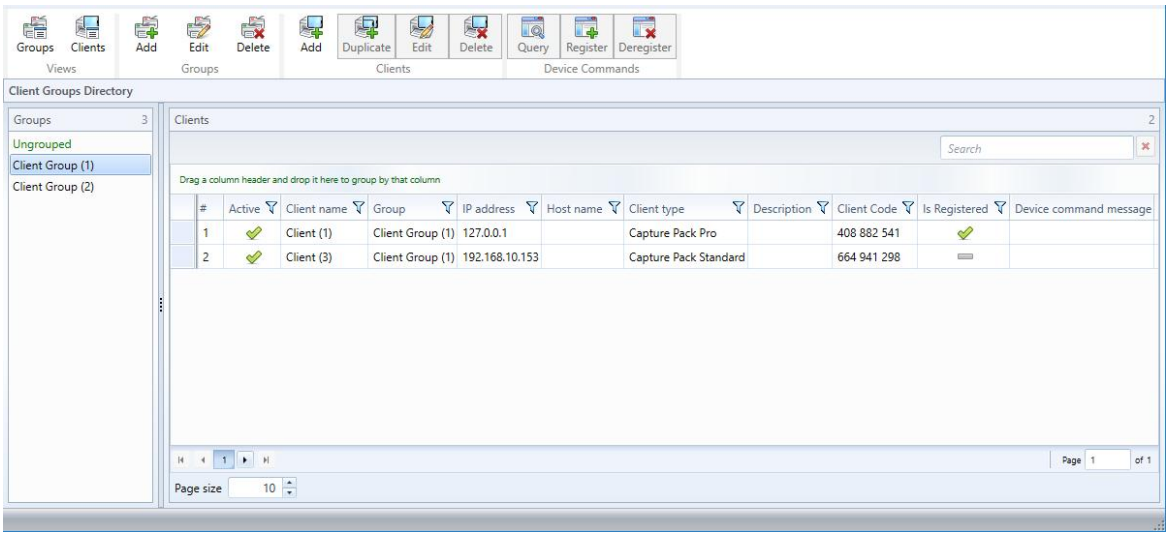
3. Templates

You can choose the sequence in which to configure clients. You could create the groups first and then the clients or you could create the clients first and then the groups and then assign the clients to groups afterward.

The client settings screen can be viewed in 2 modes namely "Groups" and "Clients". The mode can be set using the respective buttons in the "Views" section of the toolbar:

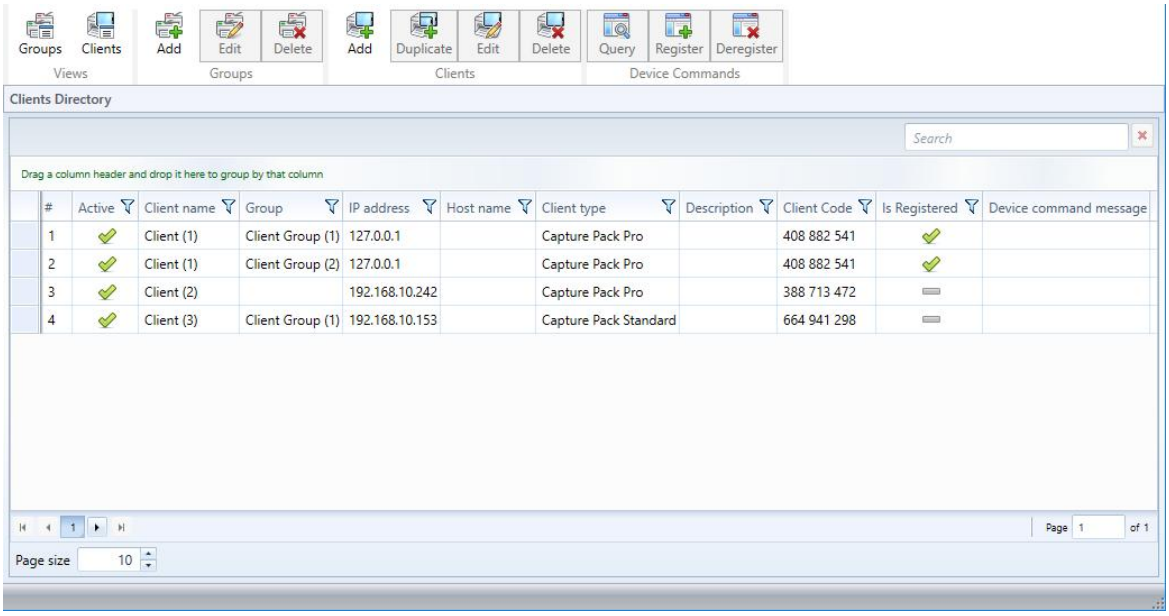


The "Group" view - which is the default - is shown below.



In the "Groups" view only the clients that belong the groups you have selected in the "Groups" list are shown.

The "Clients" view looks like this:

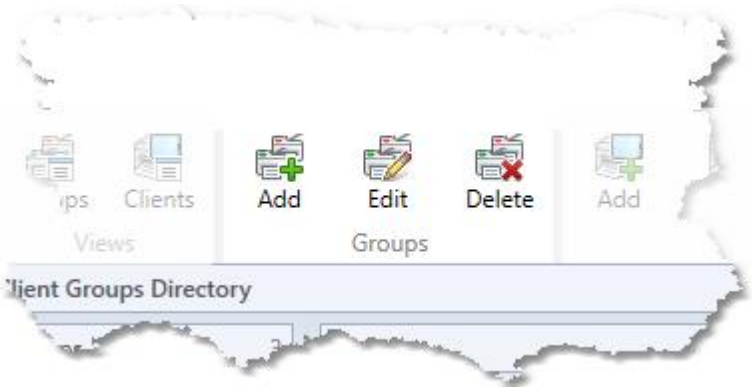


"Clients" view shows all clients. Note that clients that belong to multiple groups are duplicated in the view for each group to which they belong such as "Client (1)" in the screen shot above.

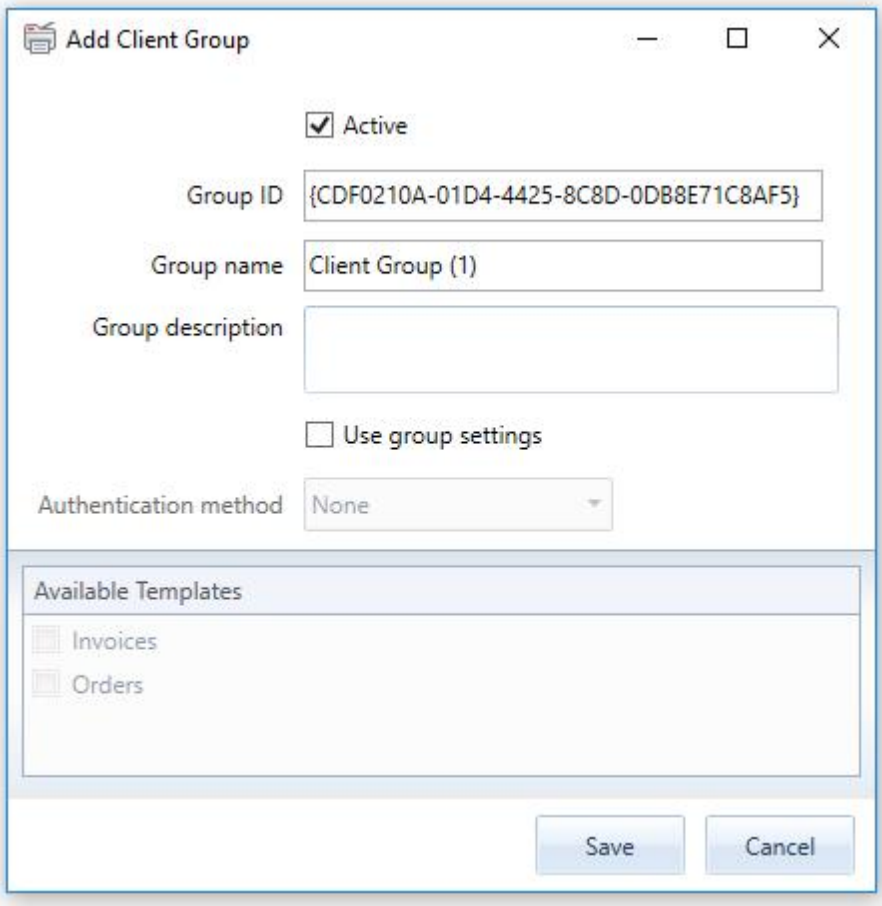
If a client does not belong to a group it will appear in the virtual group called "Ungrouped" as shown in the screen shot below.

The "Ungrouped" group is a visual grouping mechanism and it cannot be edited. If all clients belong to a group the "Ungrouped" group will not appear in the list.

To Add, Edit or Delete a group you can click the respective button in the "Groups" section of the toolbar or in the groups context menu (Context Menu):



When you click the "Add" button the following screen appears:



When a new group is created it is given the name "Client Group" with a number in brackets. This number is normally one higher than the number of existing groups with the same name, unless there is a gap in the numbering in which case the first available number is used.

Active

Enables/disables the group. If a group is disabled no client belonging to the group will be able to connect to the ScannerVision Networking Server.

Group ID

The identifier of the group which is assigned automatically when the group is created. This is a read only field.

Group name

A descriptive name for the group.

Group description

A short description of the group. This could be used to provide further information about the group to other ScannerVision administrators.

Use group settings

When this option is enabled clients' authentication method and templates are configured on the group level. In other words, all clients belonging to the group will share the list of templates you select here as well as the authentication method. If this option is disabled both the authentication method and templates list have to be configured on a per client basis.

Authentication method

The authentication method determines how a client is authenticated when a connection request comes in to the ScannerVision Networking Server. If you select "None" all the templates you select will appear on the clients belonging to the group. If either "ScannerVision", "Equitrac", "PaperCut" or "MyQ" is selected the templates specified under the **User** or **User Group** settings will be used.

The authentication method is only available when the "Use group settings" check box is checked. Available options are:

None	No authentication is required so any person will be able to use any client in the group.
ScannerVision	Only ScannerVision users will be able to use the clients in the group.
Equitrac	Only signed in Equitrac users will be able to use the clients in the group. See Appendix F - Equitrac for more information.
PaperCut	Only signed in PaperCut users will be able to use the clients in the group.

MyQ	Only signed in MyQ users will be able to use the clients in the group.
Kyocera Net Manager	Only signed in Kyocera Net Manager users will be able to use the clients in the group.
aQrate	Only signed in aQrate users will be able to use the clients in the group.

Available Templates

A list of all templates in ScannerVision that have client capture enabled. Select the templates in the list which you want to be available on clients in the group. You can re-arrange the order of the templates which will be the order in which they are displayed on the client. To change the order of a template select the template and right click on it to show the context menu. From the context menu you can move the selected template up or down.

It is advisable to select a template icon for every template that will appear on a client as it makes identification of templates easier.

6.4.1.1 Context Menu

The "Groups" context menu is shown below:

Views

The views section lets you change the view to "Groups" or "Clients". See Client Groups.

Groups

Add/Edit

Adds a new client group or edits the selected group respectively.

Delete

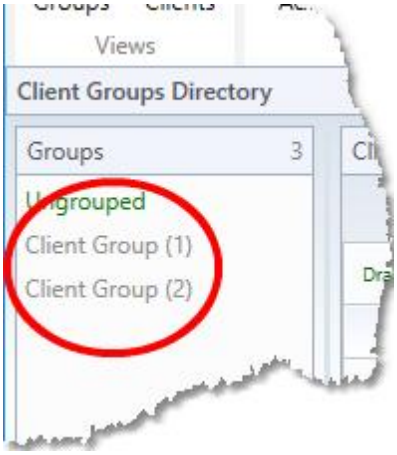
Deletes all the groups. Deleting a group **does not** delete the clients that are assigned to the group also. If all the groups that clients belong to are deleted the clients will become un-grouped.

Clear selection

Clears the selection of the "Groups" list.

Enable/Disable selection

Enables/Disables all selected groups. Disabled groups appear grayed out in the Groups list:



Reload

Reloads all groups from the database.

Clients

Add/Edit/Delete

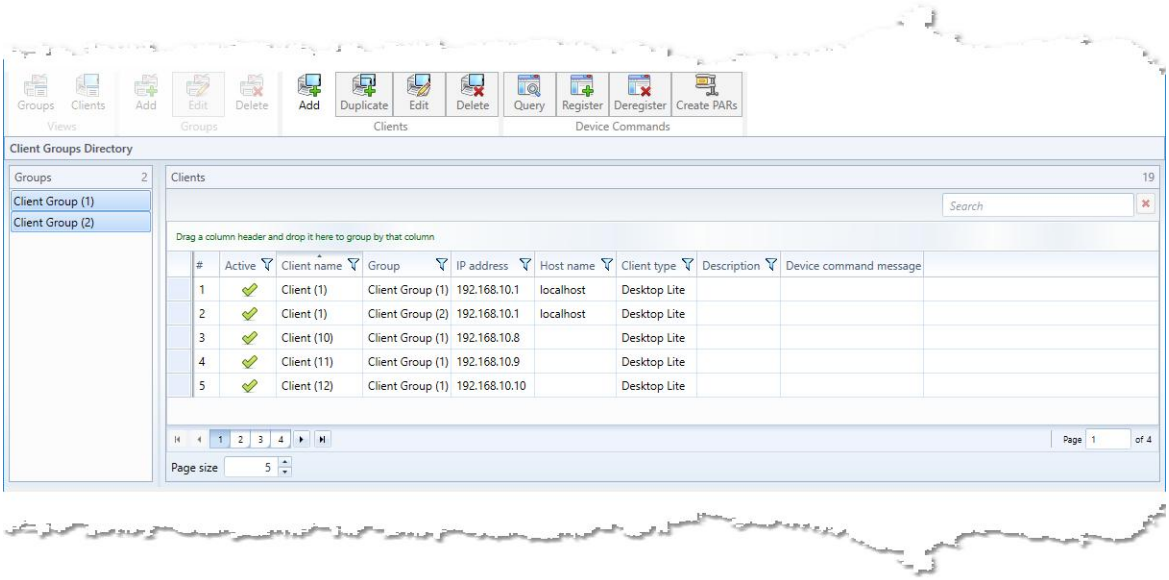
Adds a new client or edits/deletes the selected clients respectively.

Reload

Reloads all clients in the selected group(s) from the database.

6.4.2 Clients

The Client settings screen is shown below. To add a client, select the group(s) to which the client should belong and press the "Add" button on the toolbar or in the context menu (Context Menu).

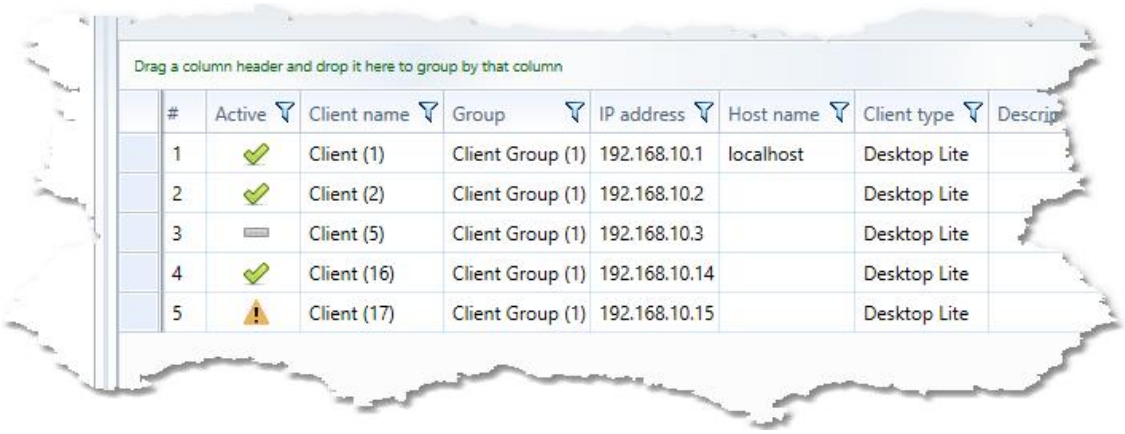


When multiple groups are selected all the clients that belong to the selected groups are shown as is seen above. Clients that appear multiple times do so only because they belong to multiple groups and not because there are multiple instances of the clients in the database. The number of entries in the grid can be controlled by the value in the "Page size" field below the grid to the left. If you have more clients than can fit on one page, additional pages are added that you can navigate to by clicking the respective page number (above the "Page size" field) or by entering a page number in the "Page" field to the right below the grid.

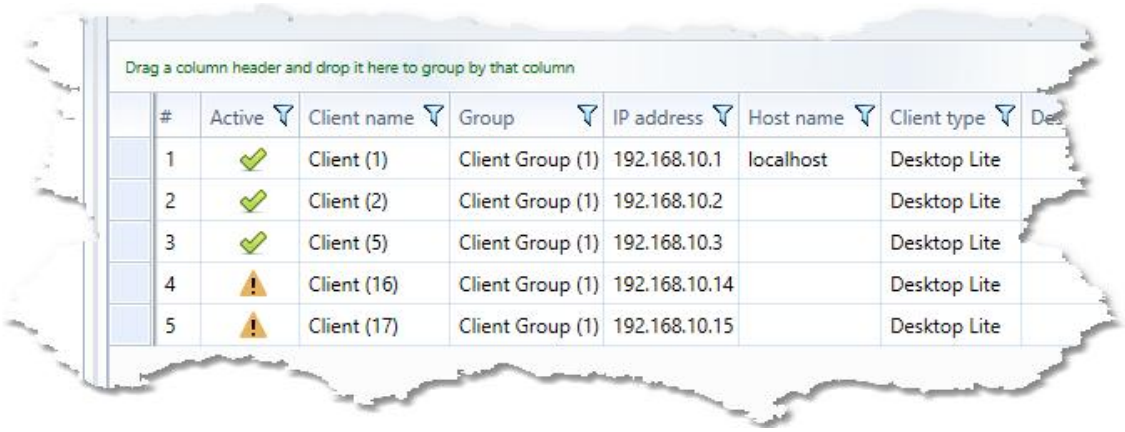
You can also search for a client by name by entering a search term in the "Search" edit box above the grid to the right.

Licensing

Client licensing in ScannerVision consist of 2 parts. Firstly, the Client Type is licensed such as Desktop Client Boost, Fuji Xerox, HP and KYOCERA. Secondly, the number of clients is licensed e.g. 20 x Desktop Client Boost, 10 x Fuji Xerox etc. You are able to configure any type and number of clients but they may not all be licensed.



In the screen shot above you will notice an exclamation mark in line 5. This indicates that "Client (17)" is not licensed. Only Active clients are licensed and in the order they were created. In the example above "Client (5)" is disabled as is indicated by the horizontal bar and "Client (16)" is active and licensed. If we were to enable "Client (5)" **it** will become licensed and "Client (16)" will become unlicensed:

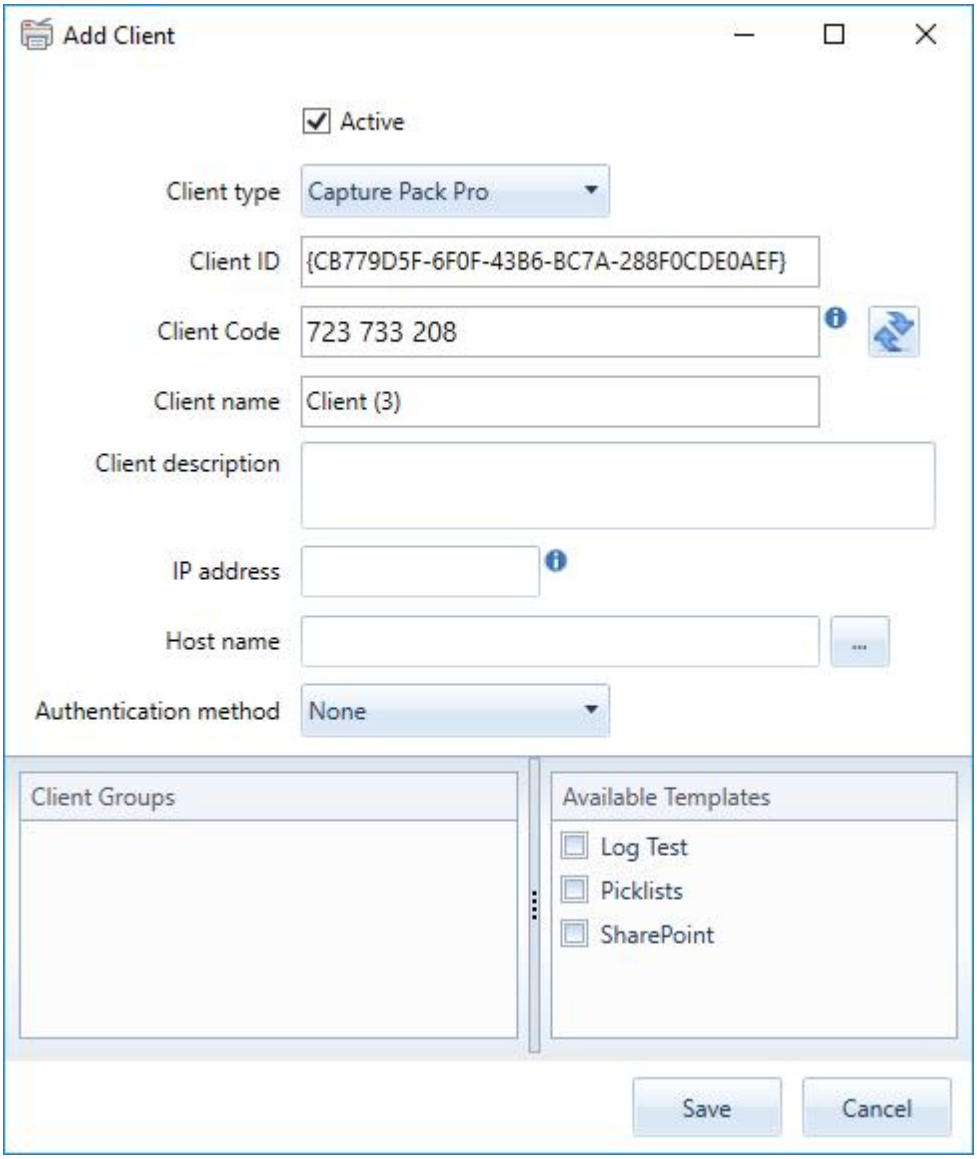


In trial mode all clients (type and number) are implicitly licensed.

6.4.2.1 Adding a Client

When you have multiple groups selected and then add a client, the client will be assigned to the selected groups automatically.

When you click the "Add" button the following screen appears:



Active

Enables/disables the client. If a client is disabled it will not be able to connect to the ScannerVision Networking Server.

Client type

The type of client. Options are:

1. Capture Pack Pro
2. Capture Pack Standard
3. Desktop Automate

4. Desktop Boost

5. Desktop Expert

6. Desktop Lite

7. Canon

8. Epson

9. Fuji Xerox

10.HP

11.KYOCERA

12.MyQ

13.Ricoh

14.Sharp

15.TA (Triumph-Adler)

16.UTAX

Client ID

The identifier of the client which is assigned automatically when the client is created. This is a read only field.

Client Code

*Only applicable to Aivika Capture Pack Pro and Aivika Capture Pack Standard clients.

A client code is a 9 digit number that is automatically generated when a new Aivika Capture Pack Pro or Aivika Capture Pack Standard client is being created. This code is used to register an Aivika Mobile client with the ScannerVision Server application. Since a mobile device is often used outside of a corporate wifi network the device's IP address cannot be used to as an identifier by the server. When an Aivika Mobile client is launched for the first time after installation on a mobile device the Client Code must be entered on the device which then requests to be registered by the server. If another client is not already registered against the client code, it is registered by the server and a registration code is returned to the mobile client which is stored on the device. The registration code is used in subsequent calls to the server to identify the client.

To reset a mobile client's registration, click the  button shown below. This will generate a new client code and clear registration code. When the mobile client that was registered before tries to connect to the server after the client code has been reset the connection will fail and the client will need to be re-registered.

Client name

A descriptive name for the client.

Client description

A short description of the client. This could be used to provide further information about the client to other ScannerVision administrators.

IP address / Host name

The IP address or host name of the MFP or computer on which the client application runs. You have to specify either the IP address or the host name and if you provide both and a client connects to the ScannerVision Networking Server it is first validated against the IP address and if that fails against the host name.

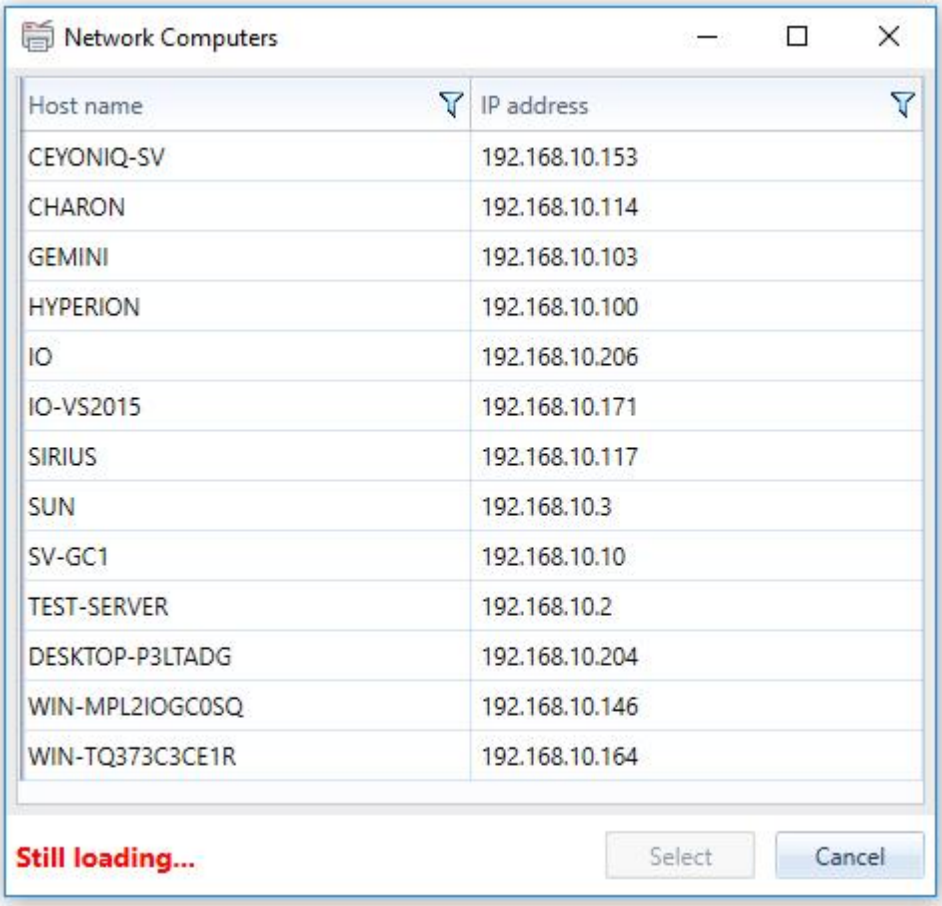
When you press the "..." button to the right of the "Host name" edit box a window appears with a list of the addresses of devices that were discovered on the network. You can select an address from the list.

User name / Password

Some clients require a user name and/or password in order to remotely deploy the client. In the case of Epson clients a non empty password is mandatory.

Note

The discovery of devices on the network could take several minutes to complete. Please be patient when you use this function.



Authentication method

The authentication method determines how a client is authenticated when a connection request comes in to the ScannerVision Networking Server. If you select "None" the templates you select below will appear on the clients belonging to the group. If either "ScannerVision", "Equitrac", "PaperCut" or "MyQ" is selected the templates specified under the **User** or **User Group** settings will be used.

The authentication method is only available when the "User group settings" check box is checked. Available options are:

None	No authentication is required so any person will be able to use a client in the group.
ScannerVision	Only ScannerVision users will be able to use the clients in the group.
Equitrac	Only signed in Equitrac users will be able to use the clients in the group. See Appendix F - Equitrac for more information.
PaperCut	Only signed in PaperCut users will be able to use the clients in the group.

MyQ	Only signed in MyQ users will be able to use the clients in the group.
Kyocera Net Manager	Only signed in Kyocera Net Manager users will be able to use the clients in the group.
aQrate	Only signed in aQrate users will be able to use the clients in the group.

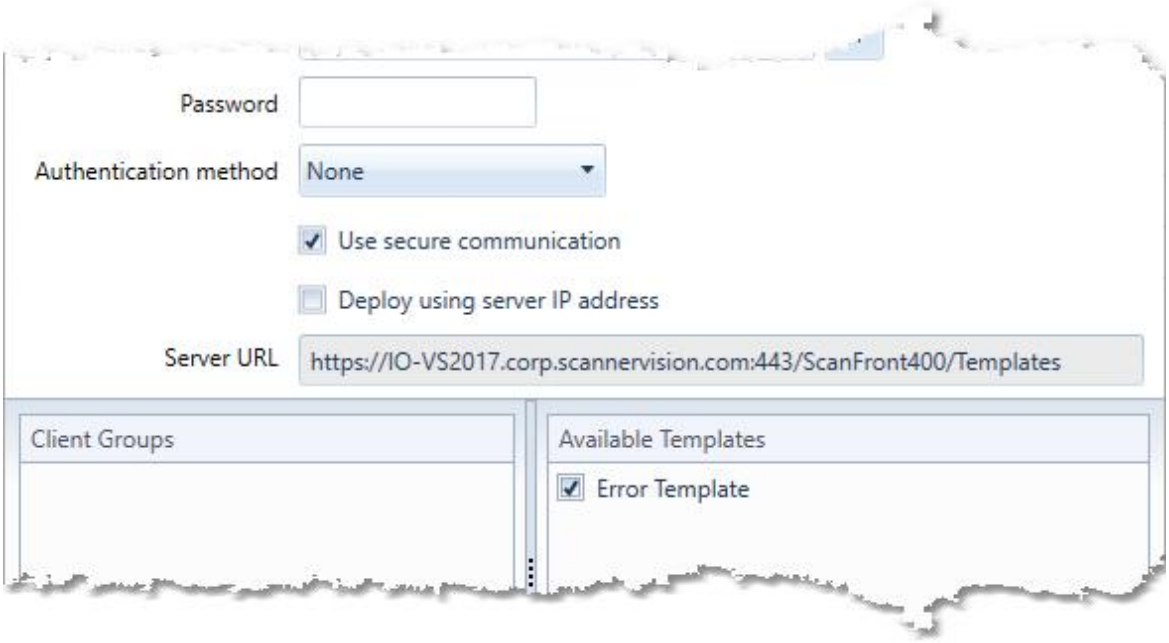
Available Templates

A list of all templates in ScannerVision that have client capture enabled. Select the templates in the list which you want to be available on clients in the group. You can re-arrange the order of the templates which will be the order in which they are displayed on the client. To change the order of a template, select the template and right click on it to show the context menu. From the context menu you can move the selected template up and down.

It is advisable to select a template icon for every template that will appear on a client as it makes identification of templates easier.

Server URL options

Some clients such as Canon, FujiXerox, Epson and HP offer additional options as shown below.



Use secure communication

When this option is selected the URL that is registered on the device will have the "https" prefix and it will include the SSL port. Some brands of MFPs such as FujiXerox may have a different prefix.

This option is only available to be set when non-secure network communication is enabled on the Network Server Settings screen. In this case you can choose whether you want the client to communicate over SSL or not. If non-secure network communication is disabled you don't have this choice and thus the check box is checked automatically and disabled.

Some MFP brands may not allow connection to a SSL enabled web server when the certificate that is used by the web server was not signed by a trusted Certificate Authority (CA). Depending on the brand and model of your MFP, it may or may not come pre-installed with the certificates of trusted CAs such as Comodo, GlobalSign, Go Daddy, Thawte etc. If you have a private CA on your network that issues certificates for use by internal web services and a certificate that was issued by your internal CA is used for the ScannerVision Networking Server, you may need to install the root certificate of the CA on the MFP. Each brand/model of MFP will have its own way of doing this. Please refer to your device's manual for instructions on how to do this.

Deploy using server IP address

When this option is selected the server URL that is registered on the device will include the IP address instead of the host name.

This option is only available to be set when non-secure network communication is enabled on the Network Server Settings screen and use of secure communication is disabled for the client.

Server URL

This read only edit box shows the address that will be registered on the device. The host name component of the URL comes from the select you have made on the Network Server Settings screen from the "Host name" drop down. If you make any change to the host name you have to redeploy or re-configure the client. Strictly speaking this is only necessary if you are using secure (SSL) communication.

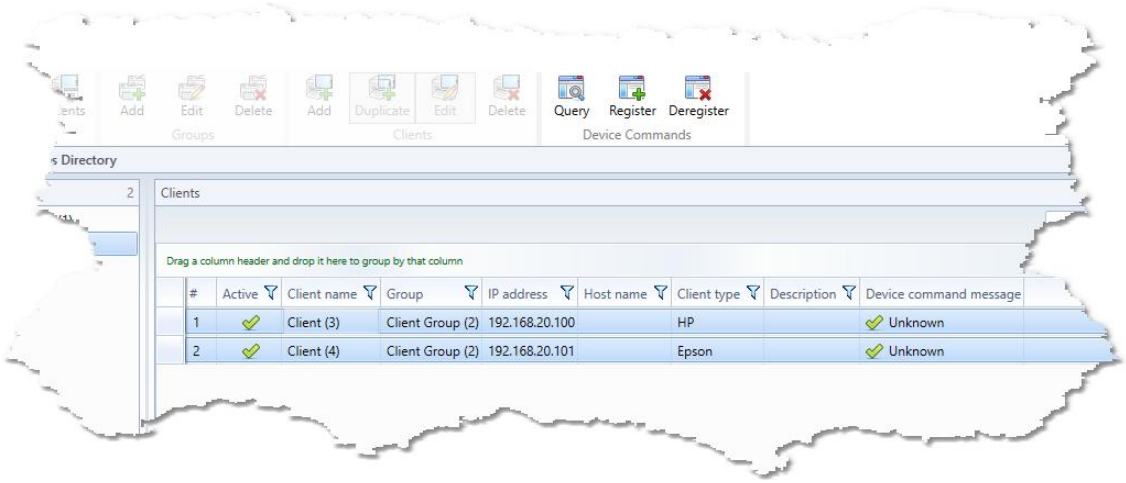
6.4.2.1.1 Remote Deployment

Most MFP clients need to be installed on the physical device and each brand of MFP has its own unique way of doing this. Some brands like Fuji Xerox requires configuration of a URL in the built in web browser only, while others such as Ricoh requires manual installation of the client application. Certain brands offer the ability to remote deploy client to the device.

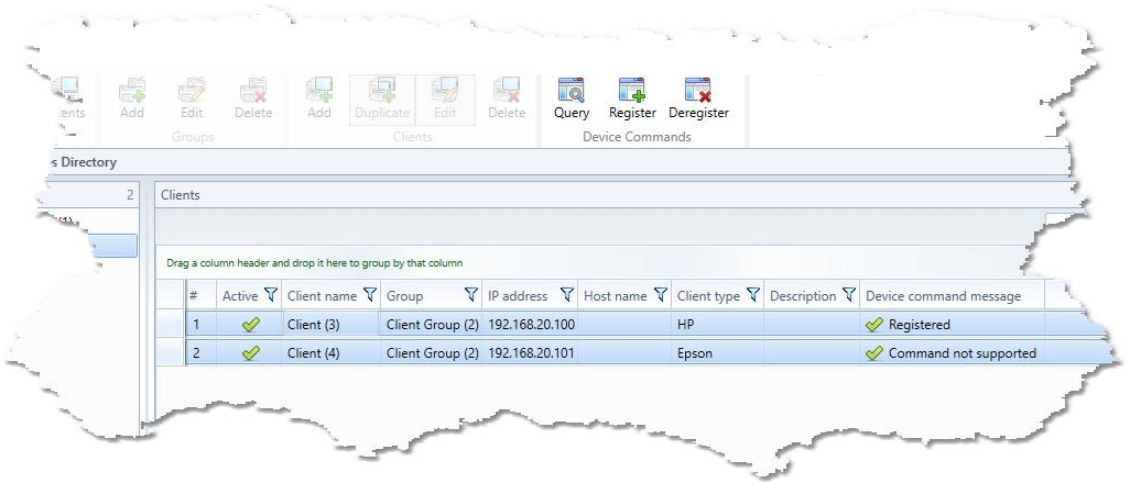
ScannerVision currently supports remote deployment to the following brands:

1. Epson
2. FujiXerox
3. HP
4. Canon

When you select one ore more clients that support remote deployment the buttons in the "Device Commands" sections of the toolbar and context menu enable. The commands can be performed on multiple clients concurrently.



Note the "Device command message" column of each client on which commands are performed to see the result. Below is the the result of the "Query" command:



Query	Interrogates the MFP to determine if the client is installed.
Register	Register (installs) the client on the MFP.
Deregister	Deregisters (removes) the client from the MFP.

Note

When deploying to an Epson or Canon ScanFront 400 device the network interface of the device or the device itself is restarted automatically which could take several minutes to complete. Please allow enough time for this before executing another command against the device.

6.4.2.2 Context Menu

The "Clients" context menu is shown below:

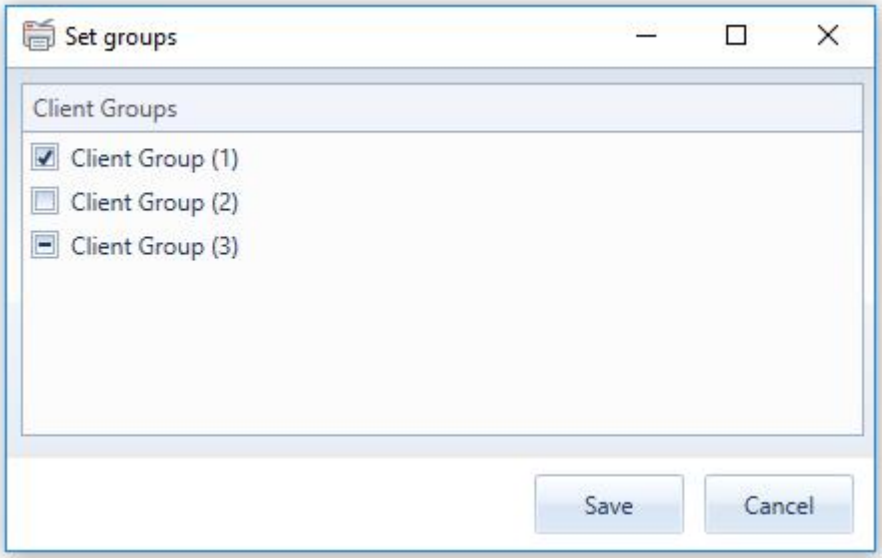
Clients

Add/Edit/Delete

Adds a new client or edits/deletes the selected client(s) respectively.

Set groups

Allows you to bulk assign the groups to which all the selected clients belong. The "Set Groups" window is shown below:



The check box that appears next to a client group when the window first appears could be in one of three states as represented by groups 1, 2 and 3 above:

Checked E.g. "Client Group (1)" indicates that **all** the selected clients belong to this group.

Unchecked E.g. "Client Group (2)" indicates that **none** of the selected clients belong to this group.

Combination E.g. "Client Group (3)" indicates that **some** of the selected clients belong to this group.

The states of the check box when saving have the following meaning:

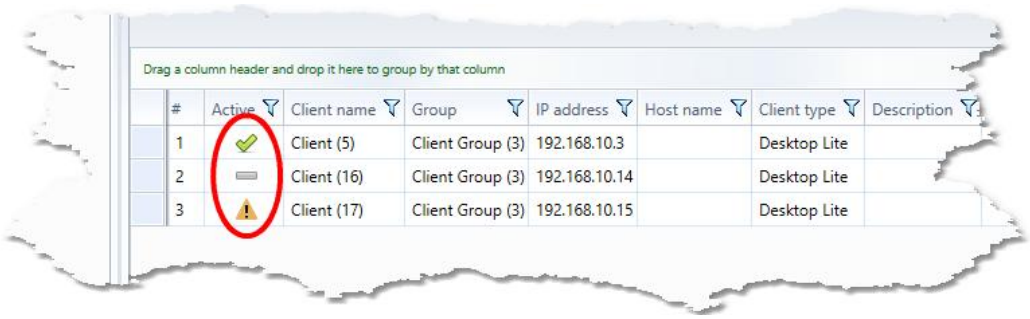
Checked All the selected clients will be assigned to this group.

Unchecked All the selected clients will be removed from this group.

Unchanged No changes will occur that affect the clients in this group.

Enable/Disable selection

Enables/Disables all selected clients. Enabled and licensed clients have a check mark in the "Active" column. Disabled clients have a horizontal gray bar while enabled and unlicensed clients have an exclamation mark:



Reload

Reloads all clients from the database.

Device Commands

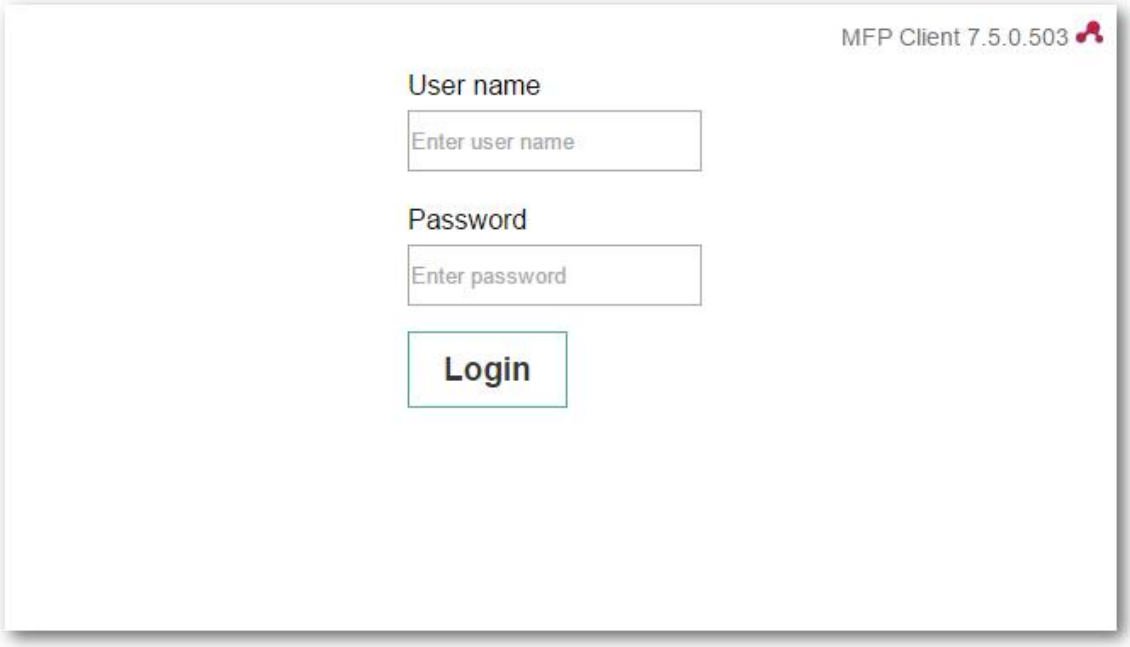
Query	Interrogates the MFP to determine if the client is installed.
Register	Register (installs) the client on the MFP.
Deregister	Deregisters (removes) the client from the MFP.

Grid Columns

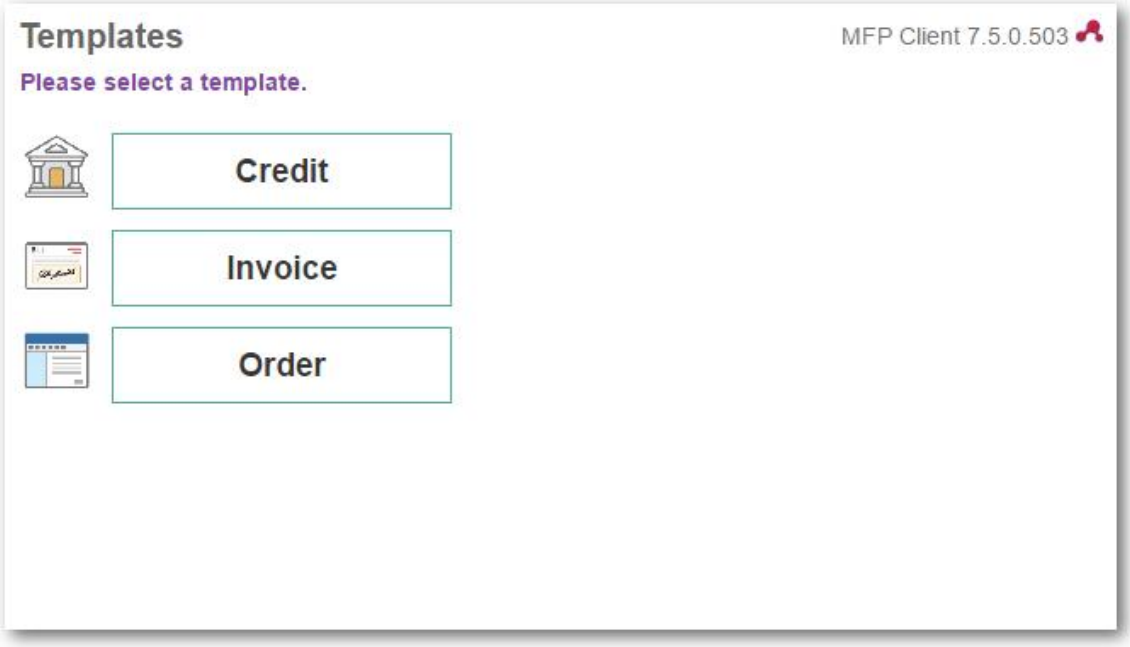
The Grid Columns section allows you to show or hide columns in the grid.

6.4.2.3 Web Client Operation

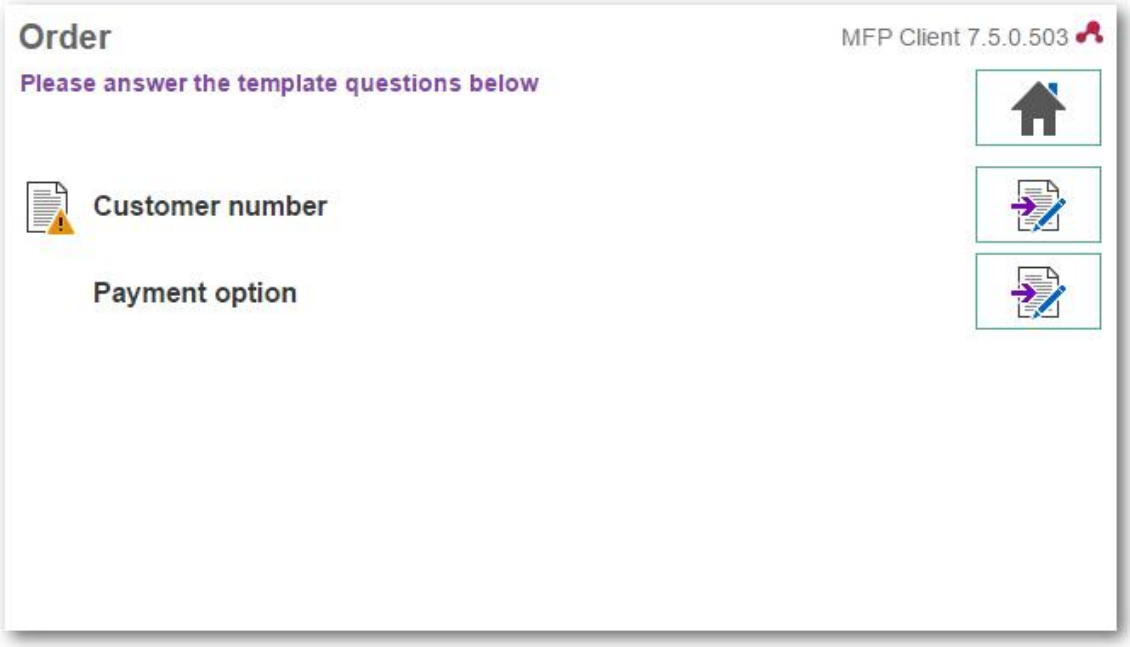
All ScannerVision web clients work in exactly the same way and the look exactly the same on all supported MFP brands. Once you have launched the ScannerVision web client, you will see a screen similar to this (if authentication is enabled on the client):



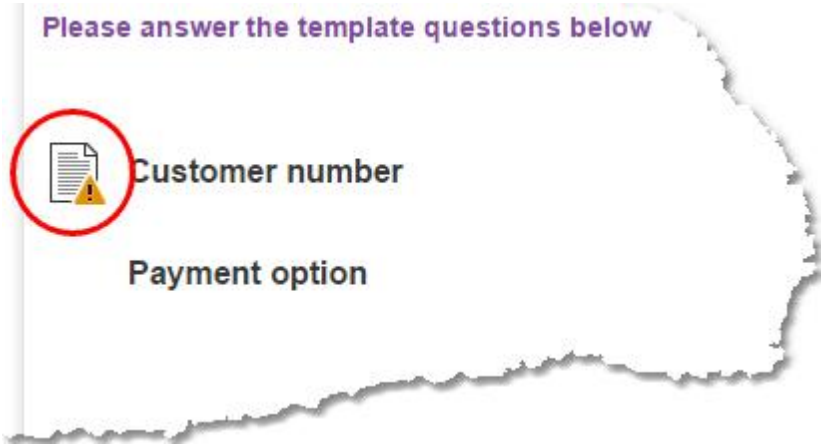
Enter your user name and password here and tap the "Login" button. This will take you to the Templates screen show below:



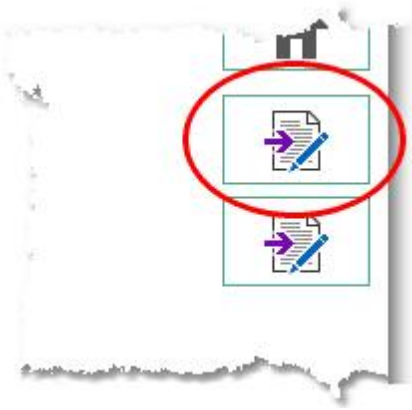
On the templates screen select the template you want to use for your scan. You may be required to complete one or more questions depending on the configuration of the template:



If a template question is required to be filled in you will see the icon with the exclamation mark circled in red below:



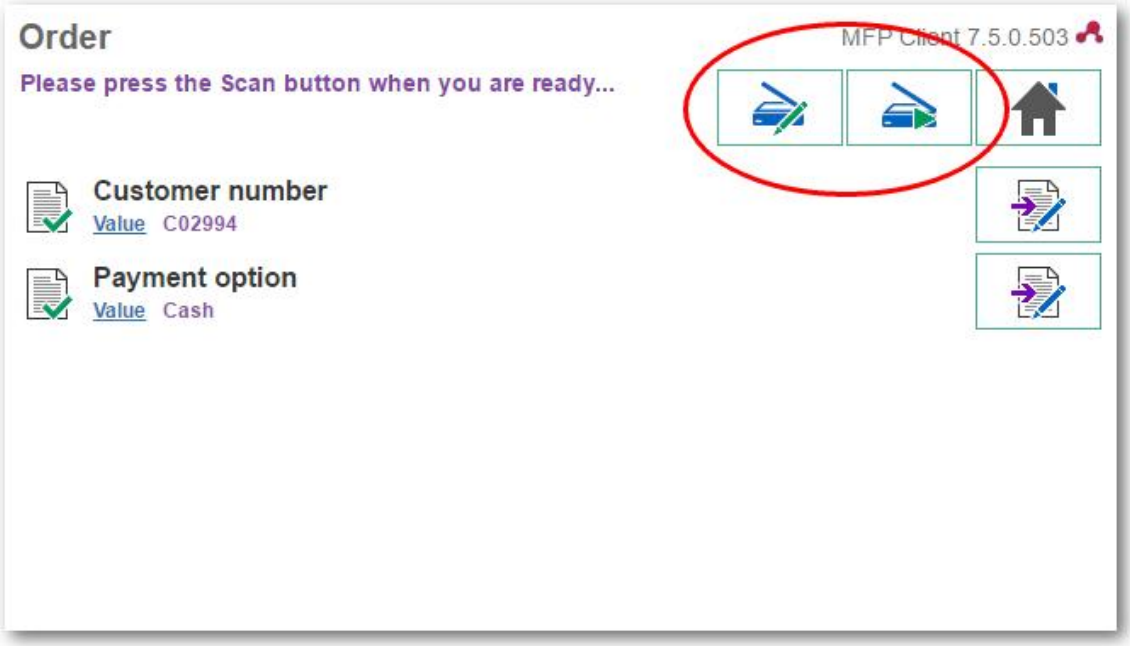
To answer a question, tap the edit button circled in red below:



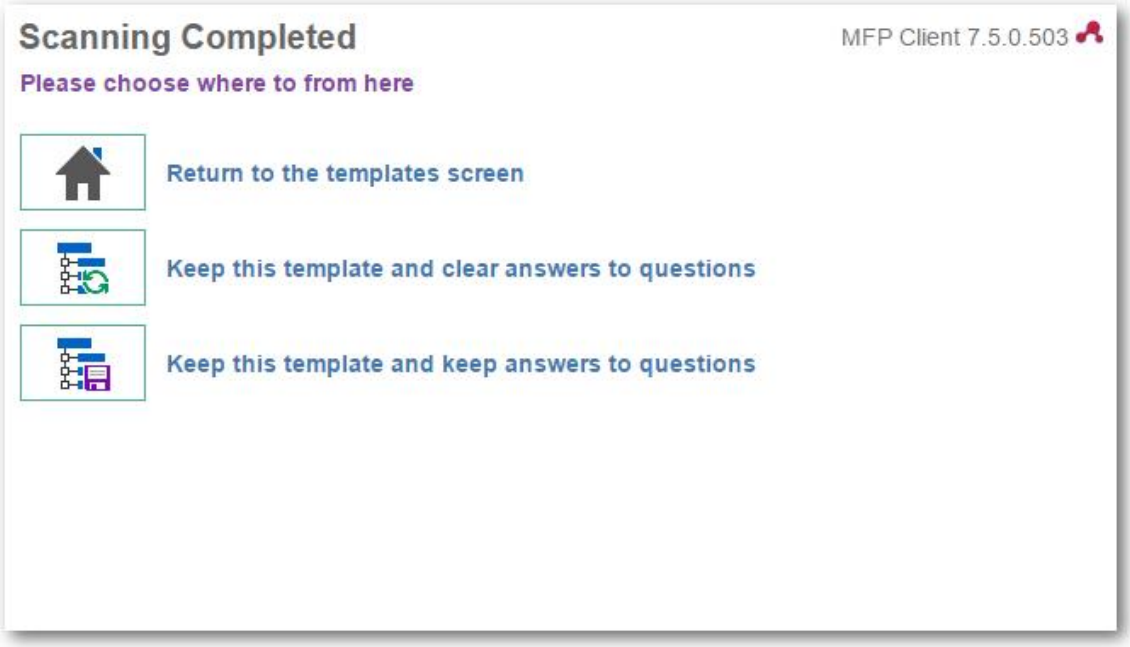
If you have to type a value for the question you will be taken to a page where you can enter a value. Tap in the edit box to show the MFP's built in soft keyboard and enter a value. If you are satisfied with the value that you have entered, tap the the "Continue" button circled in red below:

If you have to select a value from a list, tap the drop down edit box to show a list of options and tap the option you desire then click the "Continue" button.

Once you have completed all required questions the "Edit Scan Options" (left) and "Scan" (right) buttons, circled in red below, will appear:



After your scan has been completed you will be taken to the "Scanning Completed" page:



You are presented with three options here:

Return to the templates screen	Returns to the templates screen from where you can select a new template.
Keep this template and clear answers to questions	This option takes you to the template questions screen of the currently selected template. The answers to all questions have been **cleared**.
Keep this template and keep answers to questions	This option takes you to the template questions screen of the currently selected template. The answers to all questions have been **retained**.

6.4.3 Connection Checklist

If you experience difficulty getting a client to connect to the ScannerVision Networking Server please ensure the following:

1. That all ports you have configured on the Network Server Settings and FTP-ES Server Settings

screens have been opened in your firewall.

2. That you have configured a Client and that the client is Active.

3. That the ScannerVision Networking Server has been started.

If you have confirmed the above open the Log window and monitor the log output when a client is trying to connect.

Connecting with an embedded MFP client or the Desktop Client

If you see something like the following:

```
Thread #10152, 10/30/15 09:07:55, [Detailed]: Received CONNECT command
Thread #10152, 10/30/15 09:07:55, [Detailed]: Client address (host name): 127.0.0.1 (IO8), Client ID: 1, Protocol version: 9
Thread #10152, 10/30/15 09:07:55, [Detailed]: Client "IO8" is not allowed to connect
Thread #10152, 10/30/15 09:08:04, [Detailed]: Client "127.0.0.1:57281" disconnected
```

There could be one of 2 reasons for this:

1. The (Desktop Client - indicated by Client ID 1) client you configured has not been configured with an IP address of "127.0.0.1" nor with a host name of "IO8".

2. You don't have a license for the Desktop client.

Connecting with a Web MFP client

If you see the following in the log:

"Desktop" client "127.0.0.1" is not licensed to connect.

it means that you don't have a license for the Desktop Web client.

If you see the following:

No "Desktop" client with address "127.0.0.1" has been configured.

The Desktop client you configured has not been configured with an IP address of "127.0.0.1".

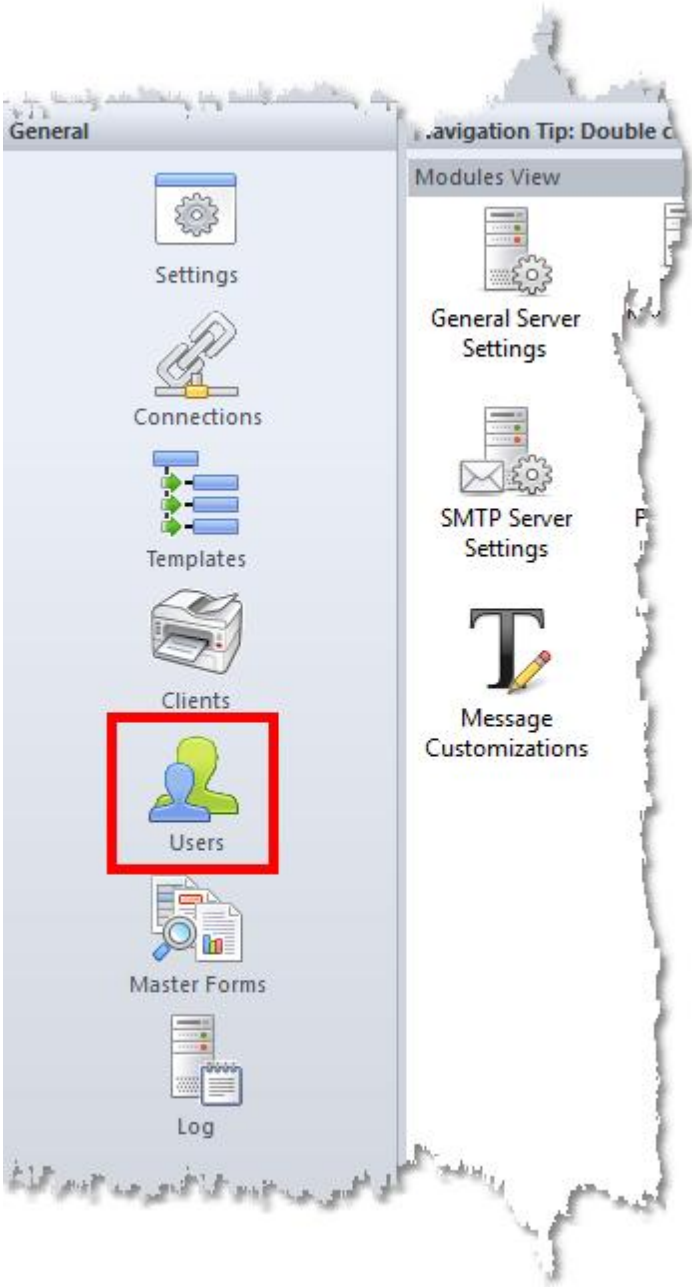
6.5 Configuring Users

Users in ScannerVision represent people who are allowed to make use of Desktop or MFP Clients. By configuring users you have the ability to restrict access to clients to authorized personnel only. You can also select the templates a user sees when he/she logs in to the client on a per user or user group basis.

Important Note

Users only come into play if user authentication has been enabled on the particular client they are using or the group the client belongs to.

To configure users press the "Users" button in the navigation pane shown below:



6.5.1 User Groups

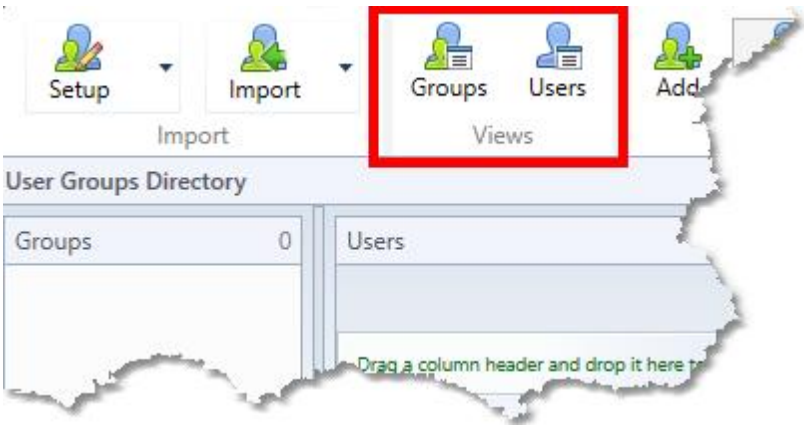
Users can belong to one or more groups or to none at all. Putting users in groups means that you can manage them on a group level. There are two aspects which can be controlled on a group level:

1. Activation/Deactivation

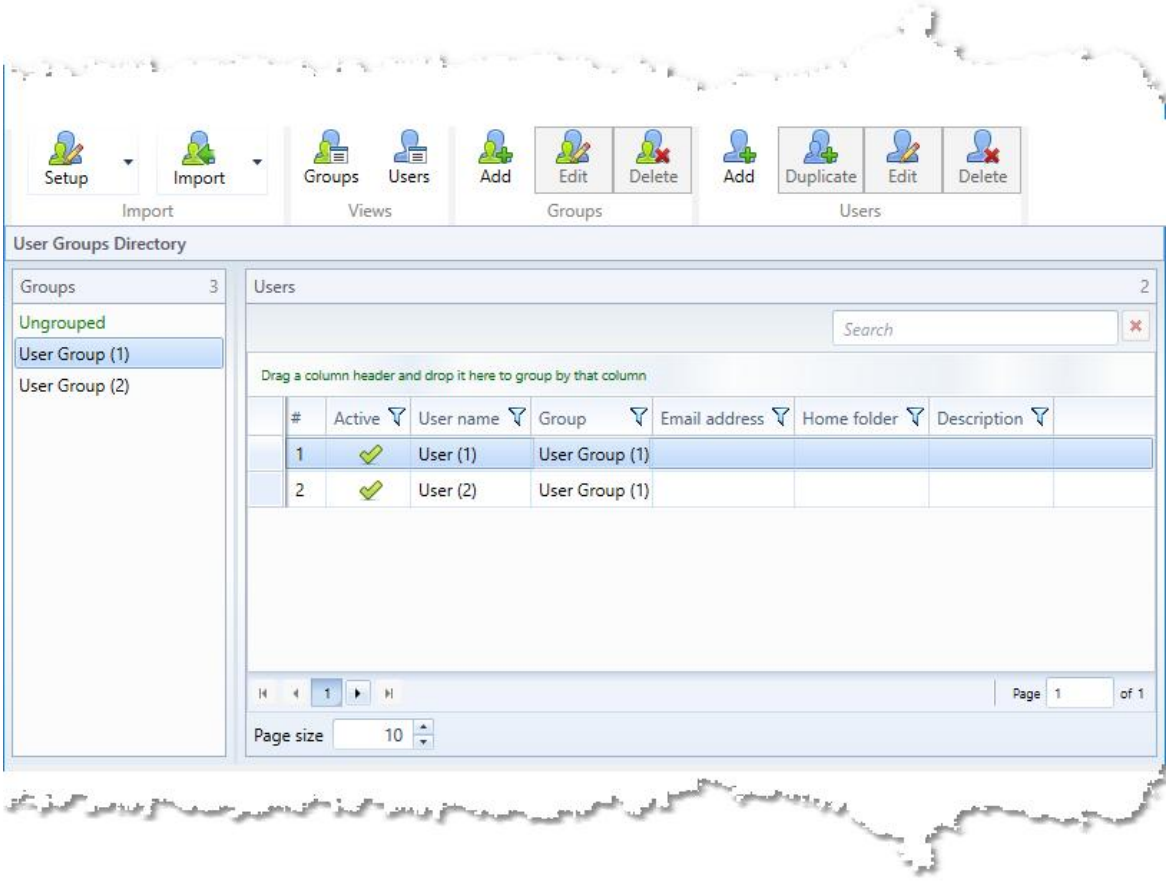
2. Templates

You can choose the sequence in which you want to configure users. You could create the groups first and then the users or you could create the users first and then the groups and then assign the users to groups afterward.

The user settings screen can be viewed in 2 modes namely "Groups" and "Users". The mode can be set using the respective buttons in the "Views" section of the toolbar:

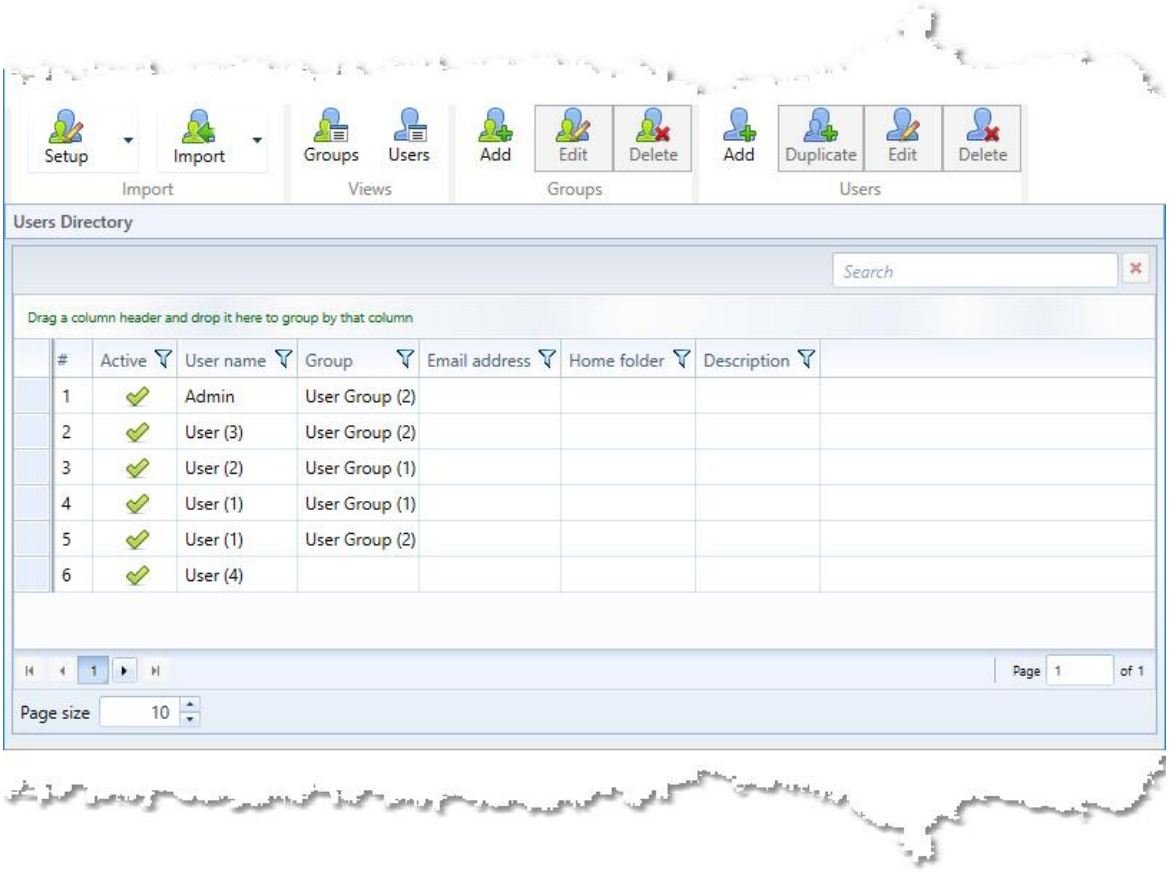


The "Group" view - which is the default - is shown below.



In the "Groups" view only the users that belong the groups you have selected in the "Groups" list are shown.

The "Users" view looks like this:

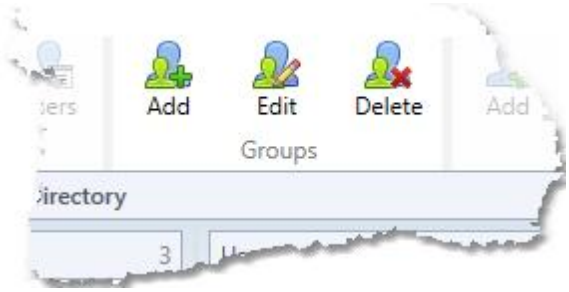


The "Users" view shows all users. Note that users who belong to multiple groups are duplicated in the view for each group to which they belong such as "User (1)" in the screen shot above.

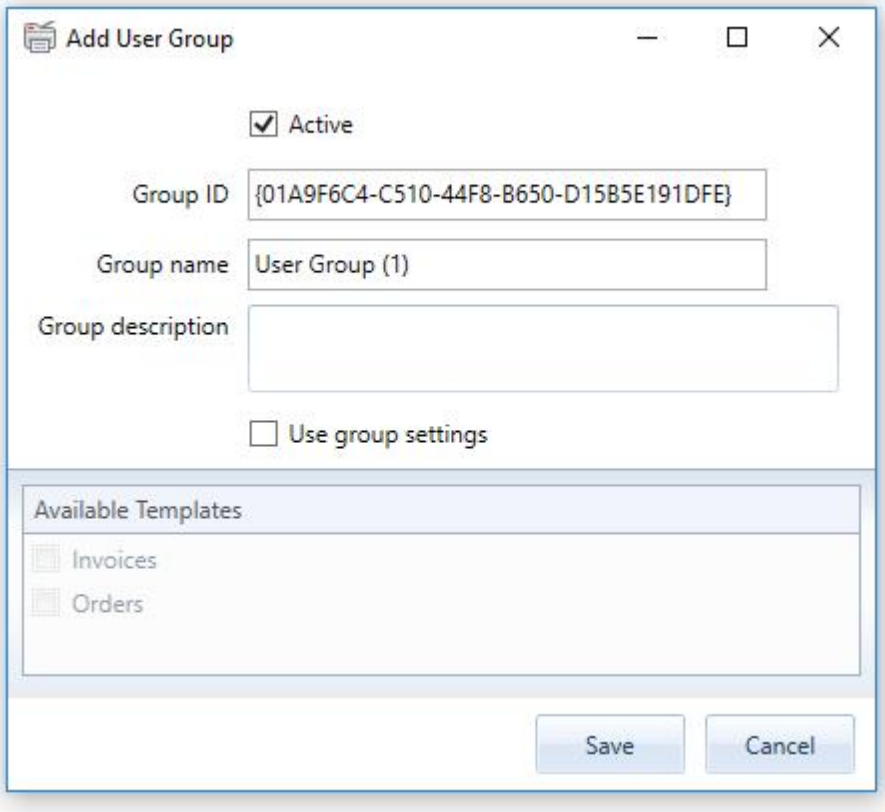
Users who do not belong to a group will appear in the virtual group called "Ungrouped" as shown in the screen shot below.

The "Ungrouped" group is a visual grouping mechanism and it cannot be edited. If all users belong to a group the "Ungrouped" group will not appear in the list.

To Add, Edit or Delete a group you can click the respective button in the "Groups" section of the toolbar or in the groups context menu (Context Menu):



When you click the "Add" button the following screen appears:



When a new group is created it is given the name "User Group" with a number in brackets. This number is normally one higher than the number of existing groups with the same name, unless there

is a gap in the numbering in which case the first available number is used.

Active

Enables/disables the group. If a group is disabled no user belonging to the group will be able to connect to the ScannerVision Networking Server.

Group ID

The identifier of the group which is assigned automatically when the group is created. This is a read only field.

Group name

A descriptive name for the group.

Group description

A short description of the group. This could be used to provide further information about the group to other ScannerVision administrators.

Use group settings

When this option is enabled users' templates are configured on the group level. In other words, all users belonging to the group will share the list of templates you select here. If this option is disabled templates have to be configured on a per user basis.

Available Templates

A list of all templates in ScannerVision that have user capture enabled. Select the templates in the list which you want to be available on users in the group. You can re-arrange the order of the templates which will be the order in which they are displayed on the user. To change the order of a template select the template and right click on it to show the context menu. From the context menu you can move the selected template up or down.

It is advisable to select a template icon for every template that will appear on a user as it makes identification of template easier.

6.5.1.1 Context Menu

The "Groups" context menu is shown below:

Views

The views section lets you change the view to "Groups" or "Users". See User Groups.

Groups

Add/Edit

Adds a new user group or edits the selected group respectively.

Delete

Deletes all the groups. Deleting a group **does not** delete the users who are assigned to the group also. If all the groups that users belong to are deleted the users will become un-grouped.

Clear selection

Clears the selection of the "Groups" list.

Enable/Disable selection

Enables/Disables all selected groups. Disabled groups appear grayed out in the Groups list:

Reload

Reloads all groups from the database.

Users

Add/Edit/Delete

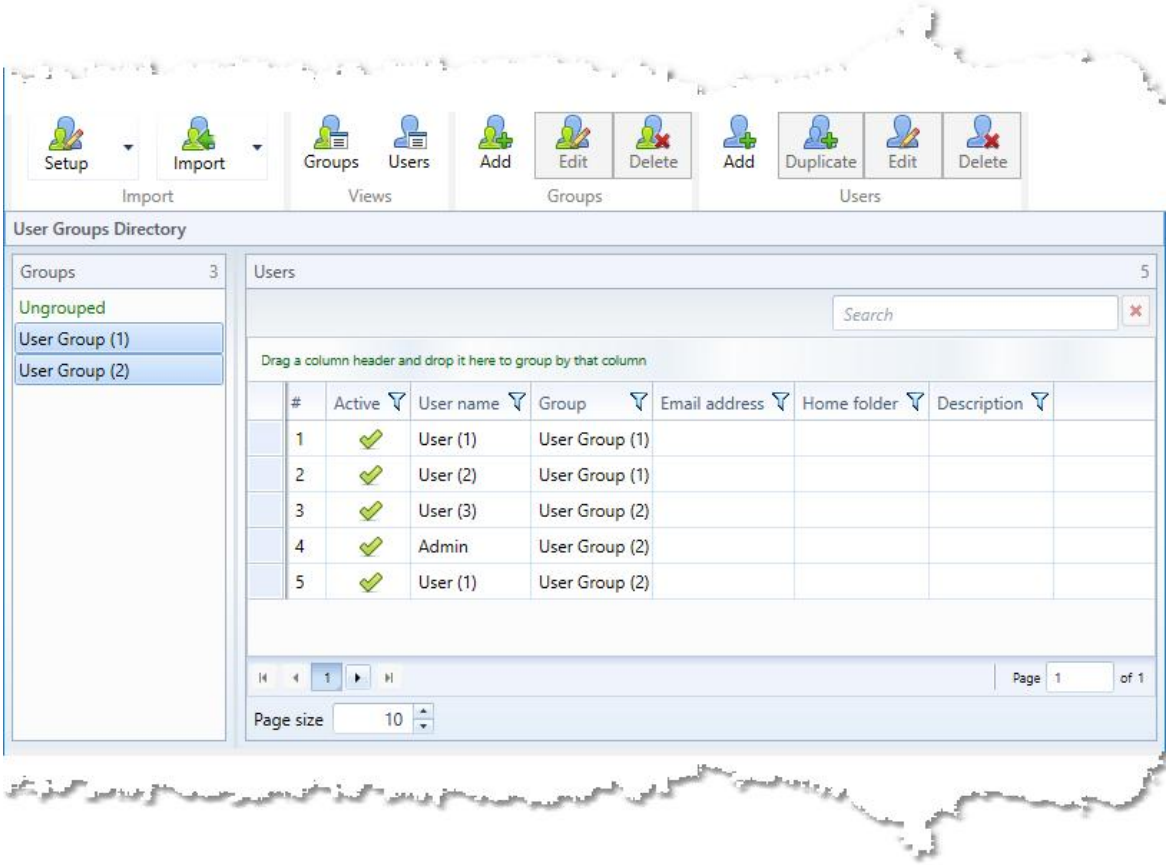
Adds a new user or edits/deletes the selected users respectively.

Reload

Reloads all users in the selected group(s) from the database.

6.5.2 Users

The User settings screen is shown below. To add a user, select the group(s) to which the user should belong and press the "Add" button on the toolbar or in the context menu (Context Menu).

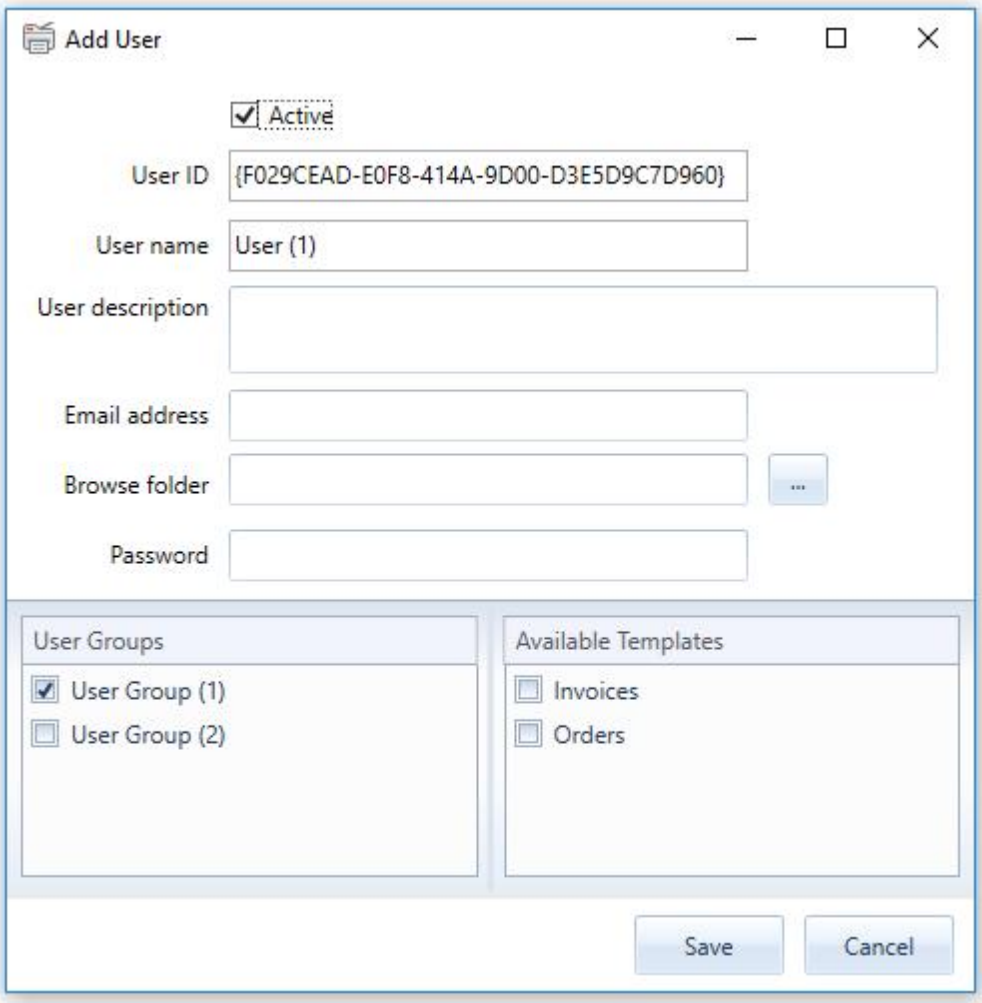


When multiple groups are selected all the users that belong to the selected groups are shown as is seen above. Users that appear multiple times do so only because they belong to multiple groups and not because there are multiple instances of the users in the database. The number of entries in the grid can be controlled by the value in the "Page size" field below the grid to the left. If you have more users than can fit on one page, additional pages are added that you can navigate to by clicking the respective page number (above the "Page size" field) or by entering a page number in the "Page" field to the right below the grid.

You can also search for a user by name by entering a search term in the "Search" edit box above the grid to the right.

When you have multiple groups selected and then add a user, the user will be assigned to the selected groups automatically.

When you user the "Add" button the following screen appears:



Active

Enables/disables the user. If a user is disabled he/she will not be able to connect to the ScannerVision Networking Server.

User ID

The identifier of the user which is assigned automatically when the user is created. This is a read only field.

User name

A descriptive name for the user.

User description

A short description of the user. This could be used to provide further information about the user to other ScannerVision administrators.

Email address

The email address of the user. This could be used in the Email Connector to send the user an email containing relevant information including the scanned document, extracted metadata etc. This value is available in the [USERMAIL] metadata tag.

Browse folder

A folder on the server or a network share that is specific to the user. Documents and captured metadata could be routed here for later retrieval by the user. This value is available in the [USERHOMEFOLDER] metadata tag. This is also the folder which content is served in template questions of type "Browse" (see Capturing Document Metadata).

Available Templates

A list of all templates in ScannerVision that have user capture enabled. Select the templates in the list which you want to be available on users in the group. You can re-arrange the order of the templates which will be the order in which they are displayed on the user. To change the order of a template, select the template and right click on it to show the context menu. From the context menu you can move the selected template up and down.

It is advisable to select a template icon for every template that will appear on a user as it makes identification of template easier.

6.5.3 User Import

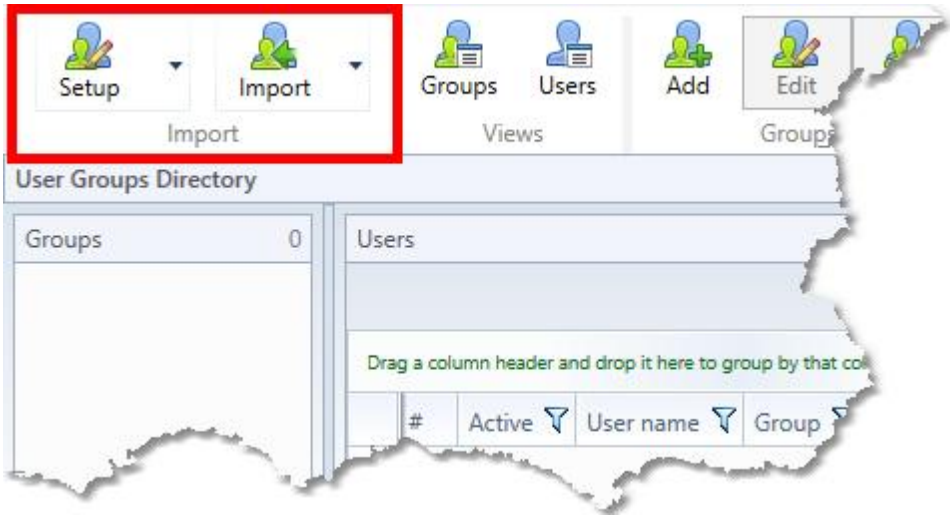
ScannerVision allows you to import users from Microsoft Active Directory (AD), CSV (CSV) files and databases (DB). You can import users from one or more of these sources and imports can be scheduled using the Windows Task Scheduler.

For further details please refer to these topics:

- AD, CSV, DB
- Windows/LDAP

6.5.3.1 AD, CSV, DB

ScannerVision offers advanced user import facilities which can be configured and executed from the section highlighted below:



You can import users from Microsoft Active Directory (AD), CSV (CSV) files and databases (DB). You can configure as many of these import sources as you want and schedule them using the Windows Task Scheduler.

When users are imported into ScannerVision they are first imported into a staging table in the database. This table contains the raw data coming from the import source as well as additional information about the status of each record. Users in the staging area are not active in ScannerVision until they have been saved from the staging area to the User table by the ScannerVision User Importer utility.

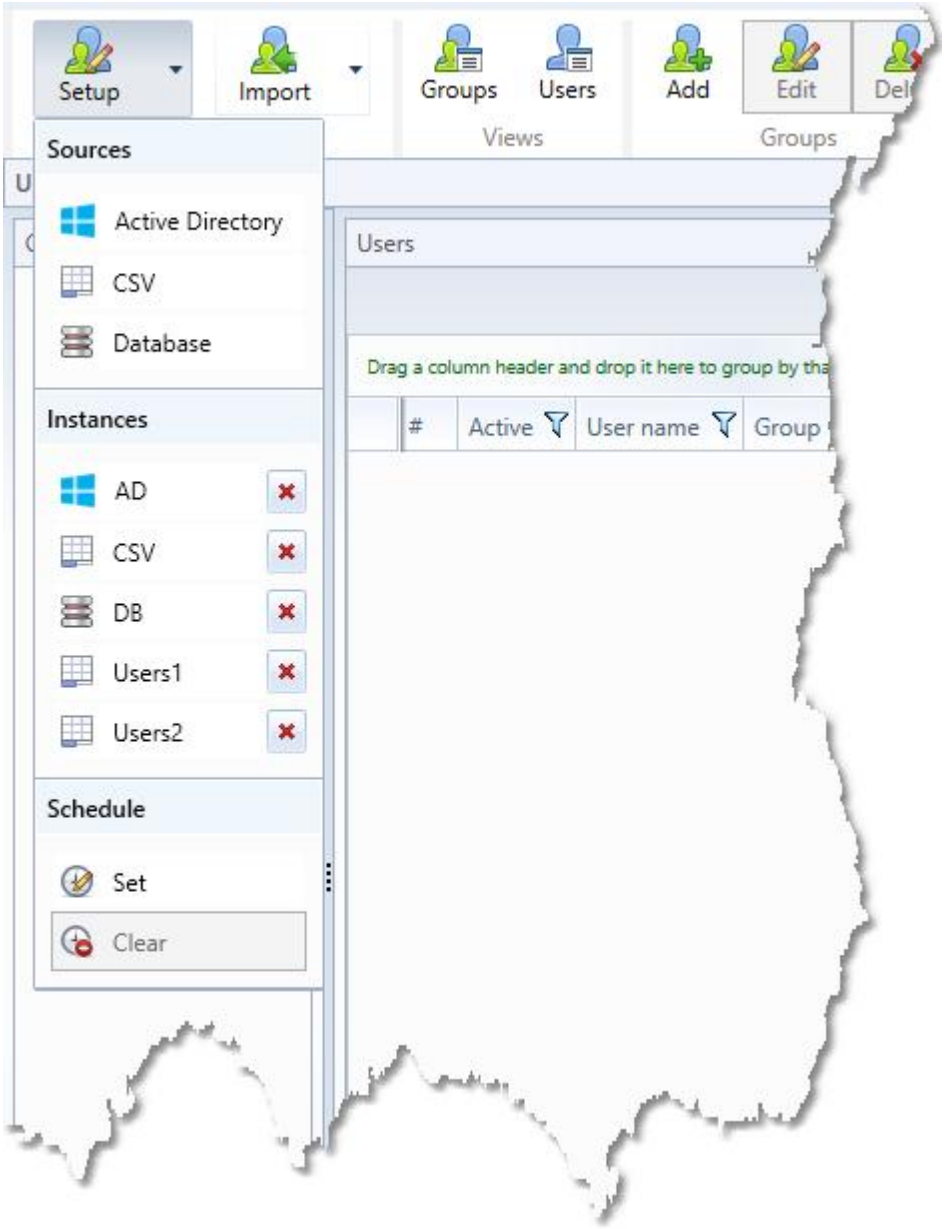
When you click the "Setup" button a popup window appears from where you can:

1. Add new import instances

2. Edit existing import instances

3. Configure the import schedule

Options 2. and 3. above will only appear after you have added at least one import source.



Sources

Press the relevant button under the "Sources" section to create a new import source instance.

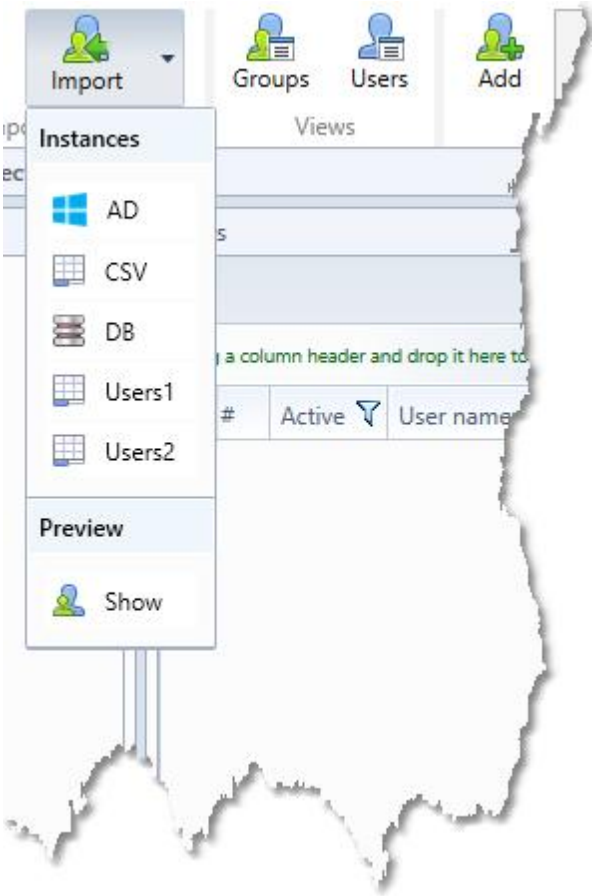
Instances

Existing import sources are listed under the "Instances" section. Since you can configure multiple import instances you have to name them. Named instances can be configured by clicking the relevant button. To delete an instance click the "x" button next to it.

Schedule

Under the "Schedule" section you can configure the schedule of when all active import instances are executed. An import instance can be scheduled individually from its respective configuration screen.

To perform an on-demand import of an import instance or to show the User Import Preview screen press the relevant button in the "Import" list shown below:



Instances

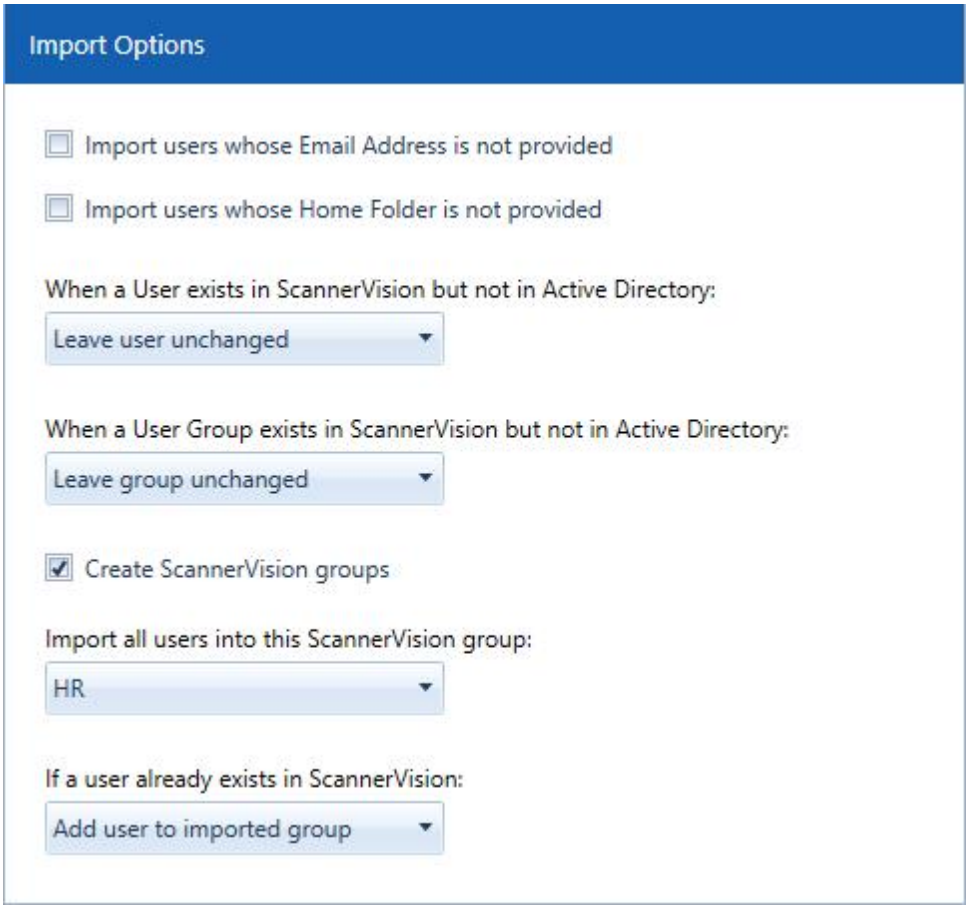
Press the relevant button in the "Instances" section to perform an on-demand import of the selected import source. An on-demand import from this section performs a complete import which is to say users a imported into the staging area and then saved to the Users table.

Preview

To display the User Import Preview screen press the "Show" button.

6.5.3.1.1 Import Options

All import sources offer a set of options that allow you to define the behavior of the particular import source.



Import users whose Email Address is not provided

Select this option to import users who do not have an email address specified in the import source data.

Import users whose Home Folder is not provided

Select this option to import users who do not have a home folder specified in the import source data.

When a User exists in ScannerVision but not in "X"

When a user exists in the ScannerVision Users table that does not exist in the import source data you have three options:

Delete user	Delete the existing user.
Disable user	Disable the existing user.
Leave user unchanged	The existing user data is not modified.

See notes below.

When a User Group exists in ScannerVision but not in "X":

When a user group exists in the ScannerVision User Group table that does not exist in the import source data you have three options:

Delete group	Delete the existing group.
Disable group	Disable the existing group.
Leave group unchanged	The existing group data is not modified.

See notes below.

Create ScannerVision groups

Select this option if you want user groups to be created automatically. If you disable this option and you also do not provide a group under the "Import all users into this ScannerVision group" option, users will be imported in the "Ungrouped" group (unless the group exists already).

Import all users into this ScannerVision group

If you provide a group and disable the "Create ScannerVision groups" option, all users will be imported in this group unless the group they belong to exists already and it is different from this one. In this case the group as it appears in the import data is applied. This group could therefore be seen as the "fallback" group when the imported user does not have a group specified.

If a user already exists in ScannerVision

If the group specified in the import data differs from that of an existing user in the Users table you have 3 options:

Add user to imported group	Add the user to the imported group in addition to the group(s) he/she belongs to already.
Move user to imported group	Remove the user from the groups he/she currently

	belongs to and move them to the group specified in the import data.
Leave user in current group(s)	Leave the user in the group(s) he/she currently belong and ignore the group specified in the import data.

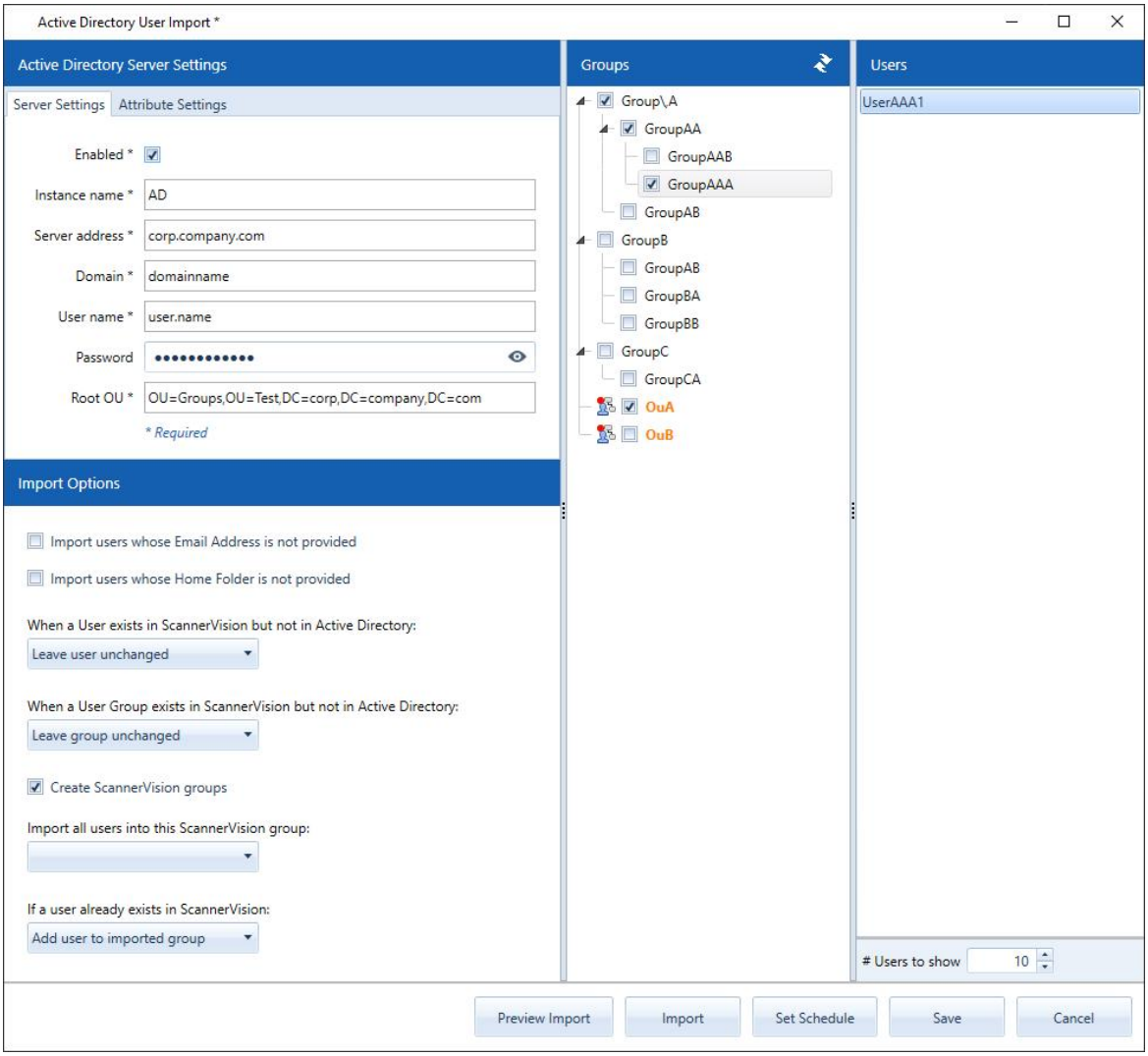
Notes

When importing users when there are already users in the ScannerVision database you should take note of a few things. Let's say you are importing users from 2 CSV files that contain users from the Human Resources and Sales departments. Further, let's assume you have configured the Sales import to delete users who don't exist in the CSV file. What is going to happen when you execute the HR import and then the Sales import? All users in HR will be deleted - assuming no users belong to both departments. Since the Sales import executes after the HR import it is going to find all the HR users existing in the Users table but none of them existing in the sales CSV file. Therefore all HR users will be deleted. If you have selected non-existing users to be disabled, all HR users will be disabled.

To avoid unintended consequences it is recommended to leave user and groups unchanged when you have multiple import sources or when you create users manually and also import users. It is only when you want to mirror a single import source such as AD that you would select either the delete or disable options for users and groups.

6.5.3.1.2 Configure AD

The Microsoft Active Directory user import setup screen is shown below. For information about import option please refer to the Import Options section.



Enabled

Enables or disable the import instance.

Instance name

The name describing this import instance. This is useful when you have multiple import sources configured.

Server address

The address of your AD server

Domain

Domain name.

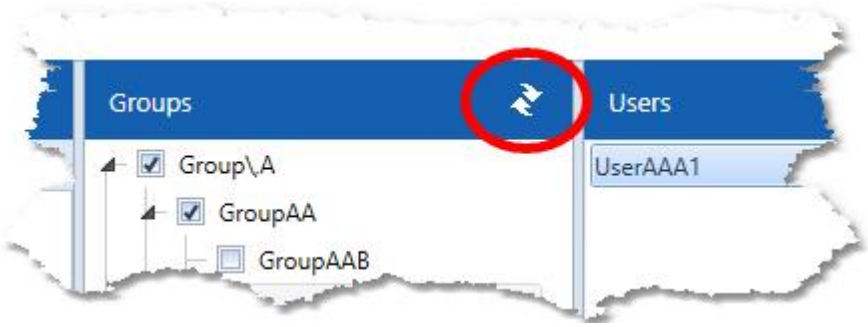
User name / Password

The credentials of a user who has permission to connect to and query the domain controller.

Root OU

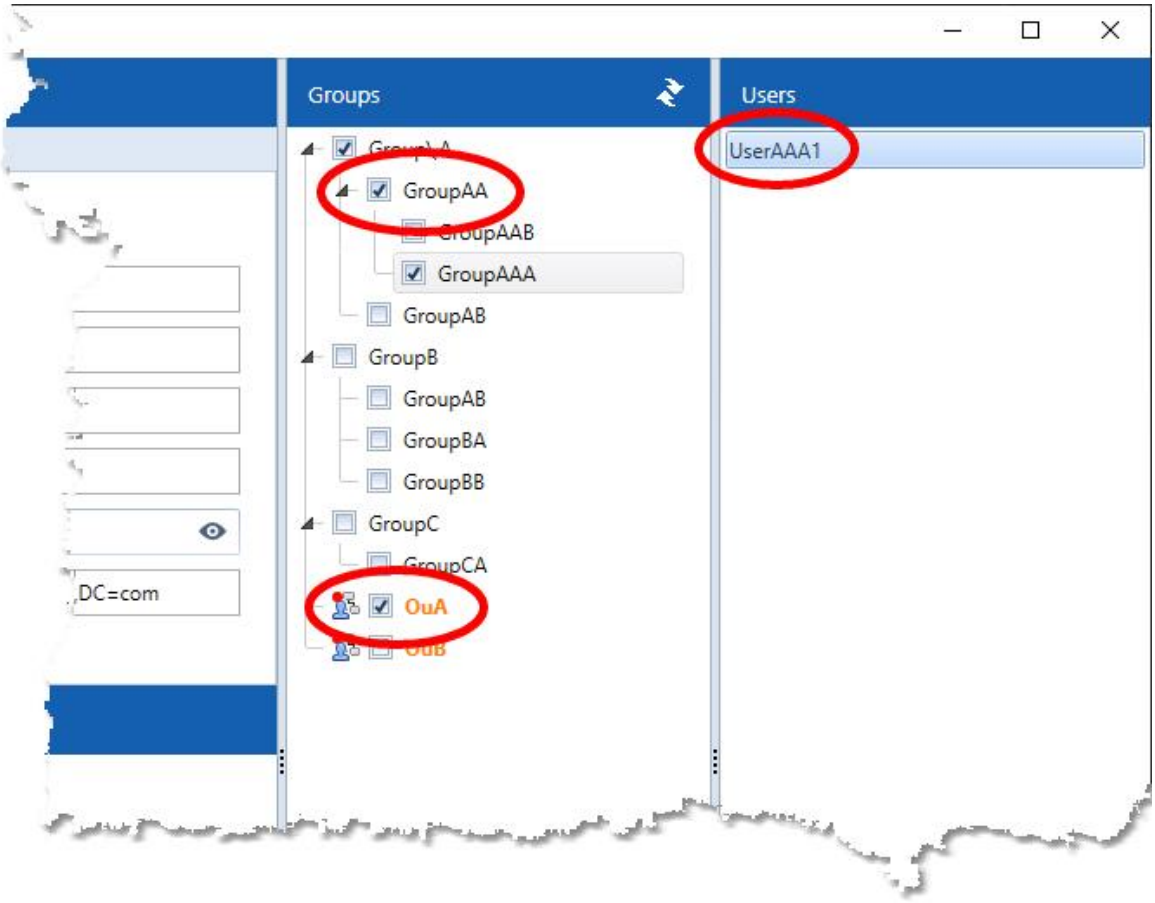
The root organizational unit from where users have to be imported.

Once you have configured your AD server settings you can press the refresh button in the top right hand corner of the "Groups" section shown below:

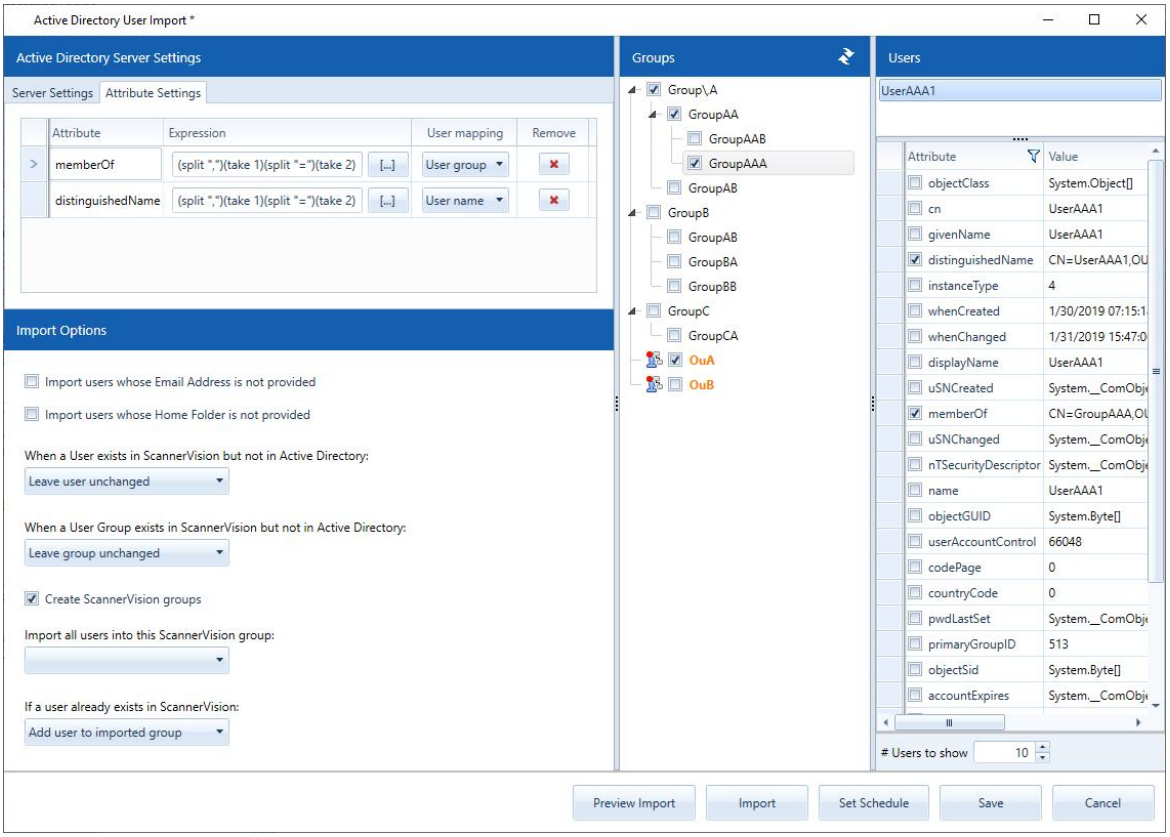


You should see a list of the groups and organizational units. When you select a group or OU the users in that group or OU are shown in the Users section. You can specify how many users must be loaded in the "# Users to show" edit box. OUs are shown in an orange color and has the OU icon to the left of the OU's name.

Only the groups and OUs which are selected, i.e. have a check mark next to the name, are imported. A context menu is available to facilitate expanding, collapsing, selection and deselection of all items and to refresh the list.

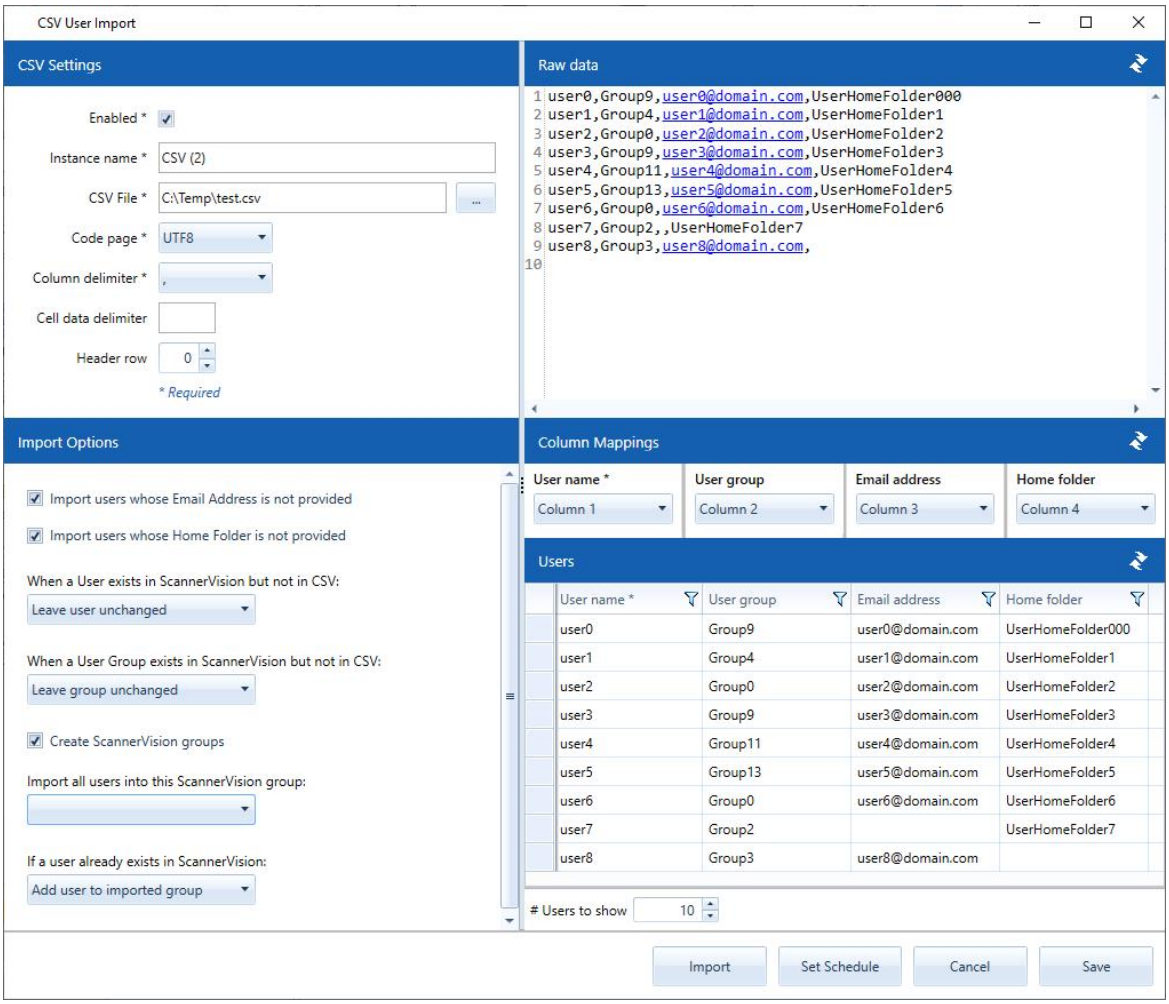


When you open the "Attribute Settings" tab shown below and then select a user, a list is populated with user attributes obtained from AD. If you so choose you can select up to 4 of these attributes and map them to the relative user fields in ScannerVision. You can also specify a ScannerVision Expression if you want to select portions from the attribute value as shown below. If you do not map an attribute to a user field the value of the respective field of the user object returned by AD will be used. In other words, you can override only the fields you want.



6.5.3.1.3 Configure CSV

The CSV user import setup screen is shown below. For information about import option please refer to the Import Options section.



Enabled

Enables or disable the import instance.

Instance name

The name describing this import instance. This is useful when you have multiple import sources configured.

CSV File

The path to the CSV file to import.

Code page

The character encoding scheme that is used by the CSV file.

Column delimiter

The character that delimits columns in the CSV file.

Cell data delimiter

If column values contain multiple items, e.g. the email address contains multiple email addresses separated by for example a comma then enter a comma in the "Cell data delimiter" edit box. If there are different cell data delimiters in the CSV file, enter them all without any spaces e.g. ",|." (the " characters are not cell data delimiters). Only the first item of the cell is imported.

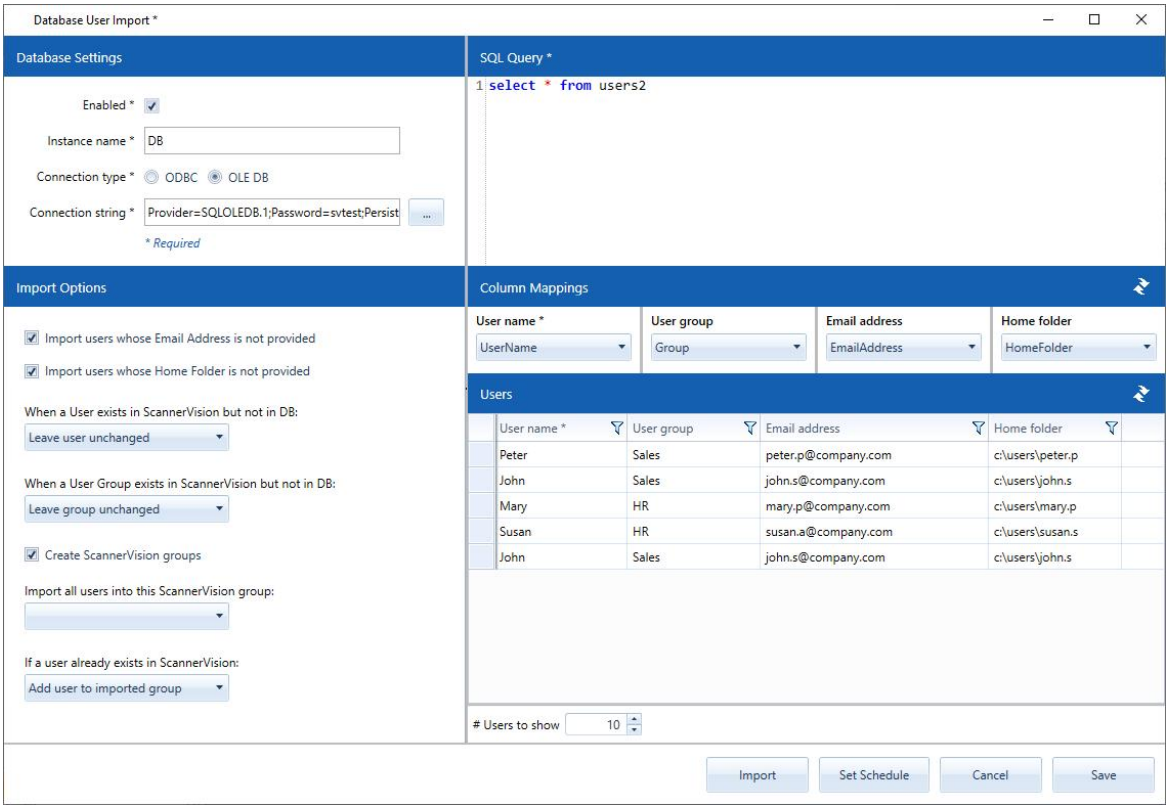
Header row

If the CSV file contains a row that specifies header information, enter the row number here.

Once you have configured the CSV settings you can click the refresh button in the top right hand corner of the respective sections. Configure column mappings to the respective user fields by using the drop down boxes in the "Column Mappings" section. You can then press the refresh button in the "Users" section to confirm that the mapping is correct. The number of user records that is loaded can be set in the "# Users to show" exit box.

6.5.3.1.4 Configure DB

The CSV user import setup screen is shown below. For information about import option please refer to the Import Options section.



Enabled

Enables or disable the import instance.

Instance name

The name describing this import instance. This is useful when you have multiple import sources configured.

Connection type

The type of connection to the database. You can connect with an ODBC or OLE DB data source.

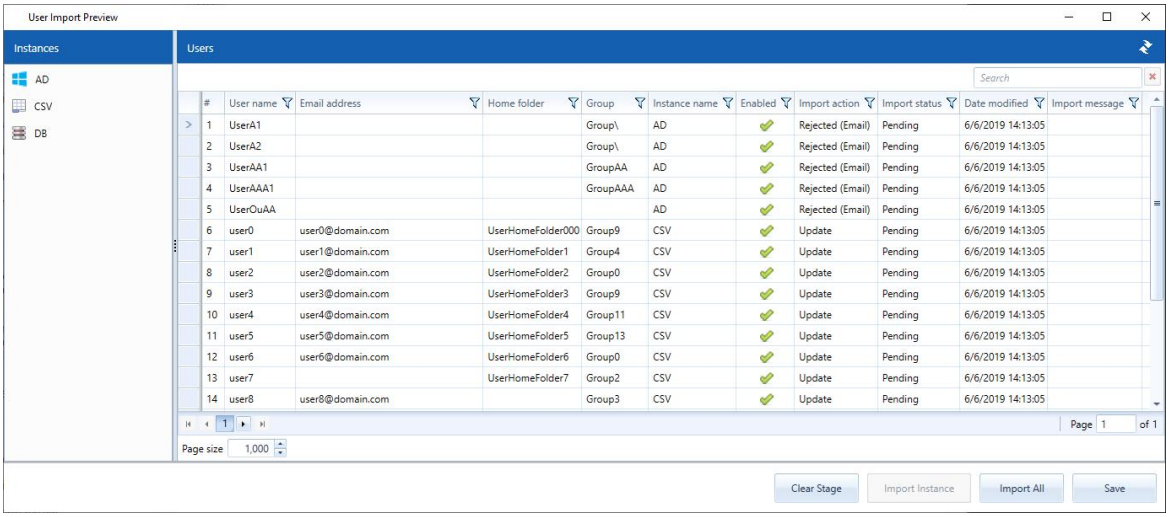
Connection string

The connection string to connect to the database. Press the "..." button to display the connection string creation wizard. Please ensure that your password is persisted in the connection string. The connection string is encrypted in the settings file.

Once you have configured the database settings you can click the refresh button in the top right hand corner of the respective sections. Configure column mappings to the respective user fields by using the drop down boxes in the "Column Mappings" section. You can then press the refresh button in the "Users" section to confirm that the mapping is correct. The number of user records that is loaded can be set in the "# Users to show" exit box.

6.5.3.1.5 User Import Preview

The User Import Preview screen is shown below. With this screen you have a view into the user import staging area. The information in this table is only cleared before an import instance is executed or when you have pressed the "Clear Stage" button. You can therefore open this view after a scheduled import was performed to see the raw data and what the status is of every record.



Clear Stage

Press this button to clear all data from the staging area.

Import Instance

Press this button to import users for the selected import instance. Data is imported to the staging area only and not to the User table.

Import All

Press this button to import users for all import instances. Data is imported to the staging area only and not to the User table.

Save

Press this button to save ALL data in the staging area to the User table. The staging are may contain more data than what is visible on the screen. You could for example have selected a particular instance from the "Instances" list in which case only the records for that instance are displayed. Or

you could have entered a search filter. So take note that when the "Save" button is pressed all records in the staging area are processed.

Staging table columns

User name, Email address, Home folder, Group

These columns contain the raw data as it existed in the import source at the time of the import.

Instance name

The import instance that resulted in the creation of this user record.

Enabled

This column indicates if the user was marked as enabled or disabled in the import source.

Import action

This column shows what action will be taken by the importer when records are saved. Possible values are:

Insert	A new user will be created.
Update	An existing user will be updated.
Delete	An existing user will be deleted.
Disable	An existing user will be disabled.
RejectedEmail	User will not be imported because he/she has no email address.
RejectedHomeFolder	User will not be imported because he/she has no home folder.

Import status

This column shows the status of the record. Possible values are:

Pending	The user has not been imported (saved) into the User table yet.
Imported	The user has been imported (saved) into the User table.
NotImported	The record was processed but the user was not imported (saved) into the User table.

Import message

This column may contain a message that explains why a user was not imported.

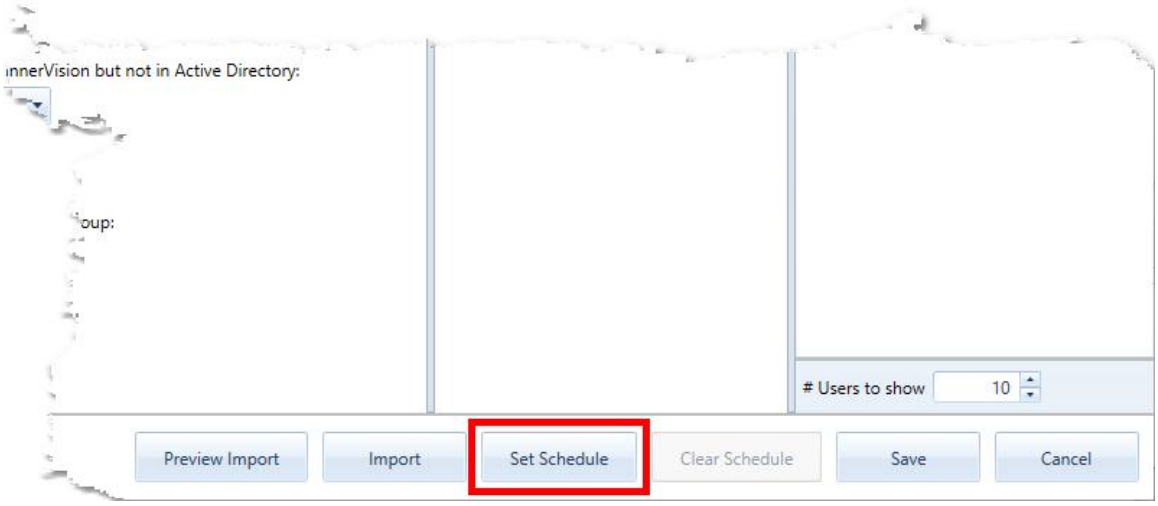
Processing order

Records are processed in the order that they appear in the "Users" grid. Even though the "AD" instance may be the first in the "Instances" list, "CSV" second and "DB" third, it is not necessarily the case that their records will be imported in this order. For example, if you clear the staging area, import from "DB" first, then "AD" and lastly from "CSV", this would be the order in which the records will appear in the staging area and consequently saved to the User table when you press the save button.

When users are imported by the "all instances" scheduled import, i.e. when no instance name is specified as a command line parameter to the ScannerVisionUserImporter.exe utility, users are imported according to the order of instances as they appear in the "Instances" list.

6.5.3.1.6 Scheduling

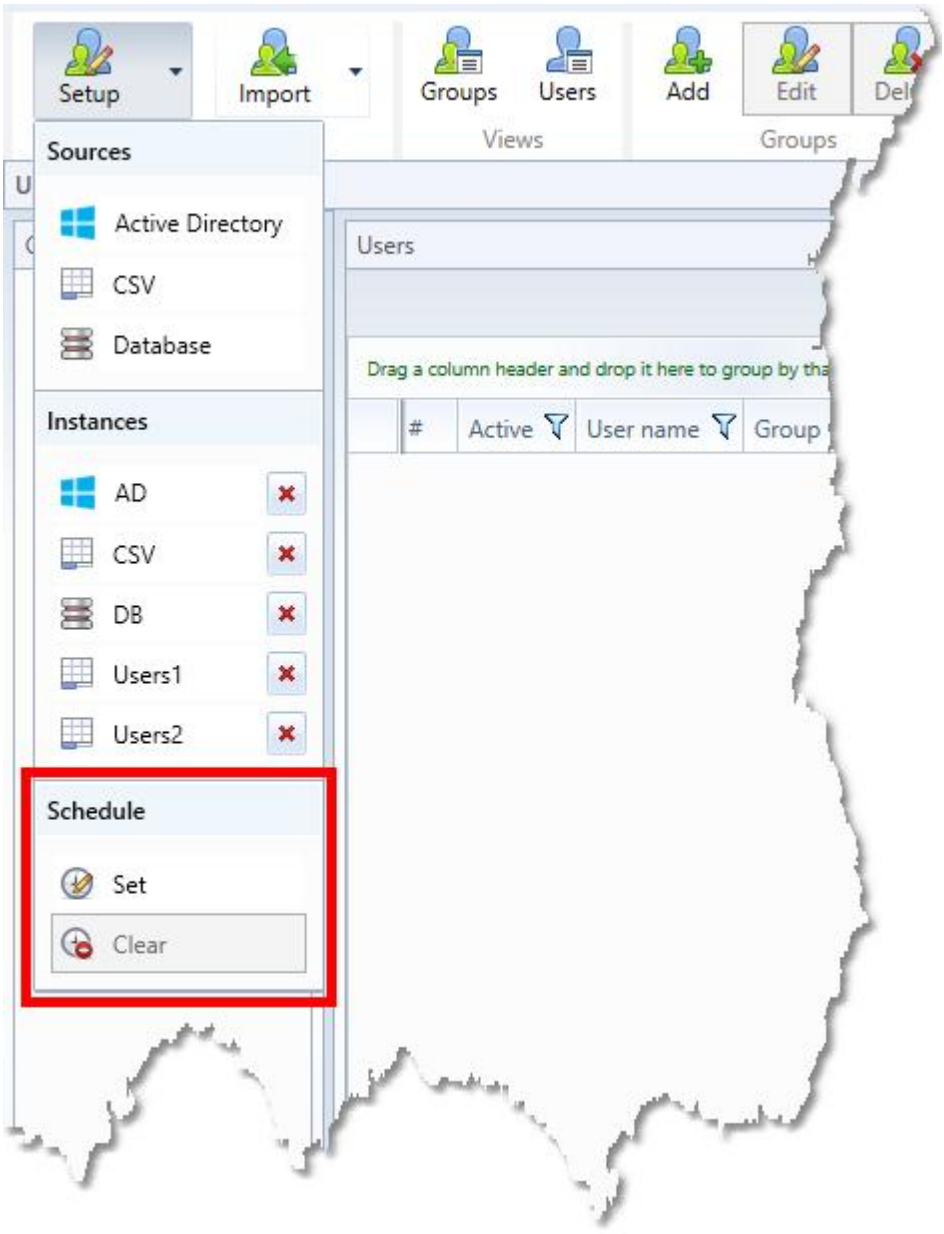
The Windows Task Scheduler is used for the scheduling of user imports. If you want to configure the schedule yourself refer to the ScannerVisionUserImporter section for details on how to call the import utility. If you have multiple import sources you can schedule them individually by clicking the "Set Schedule" button on the respective instance's configuration screen as shown below:



Only one instance of the import utility can run at a time so when you configure schedules manually make sure you allow enough time between them so that they don't overlap. The clear a schedule press the "Clear Schedule" button. This button will be enabled when a schedule has been set for this

instance.

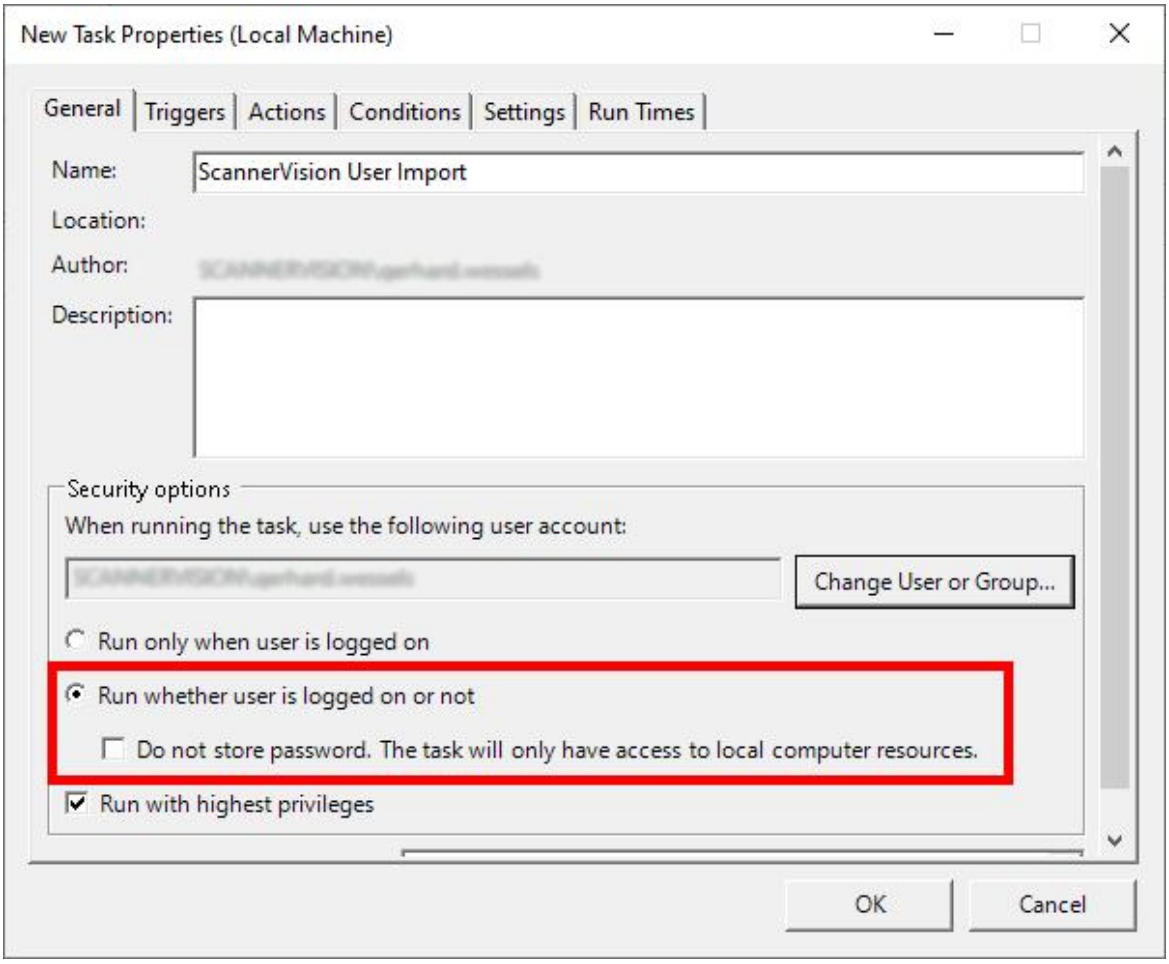
To set or clear the import schedule for all instances press the relevant button under the "Schedule" section of the "Setup" button list.



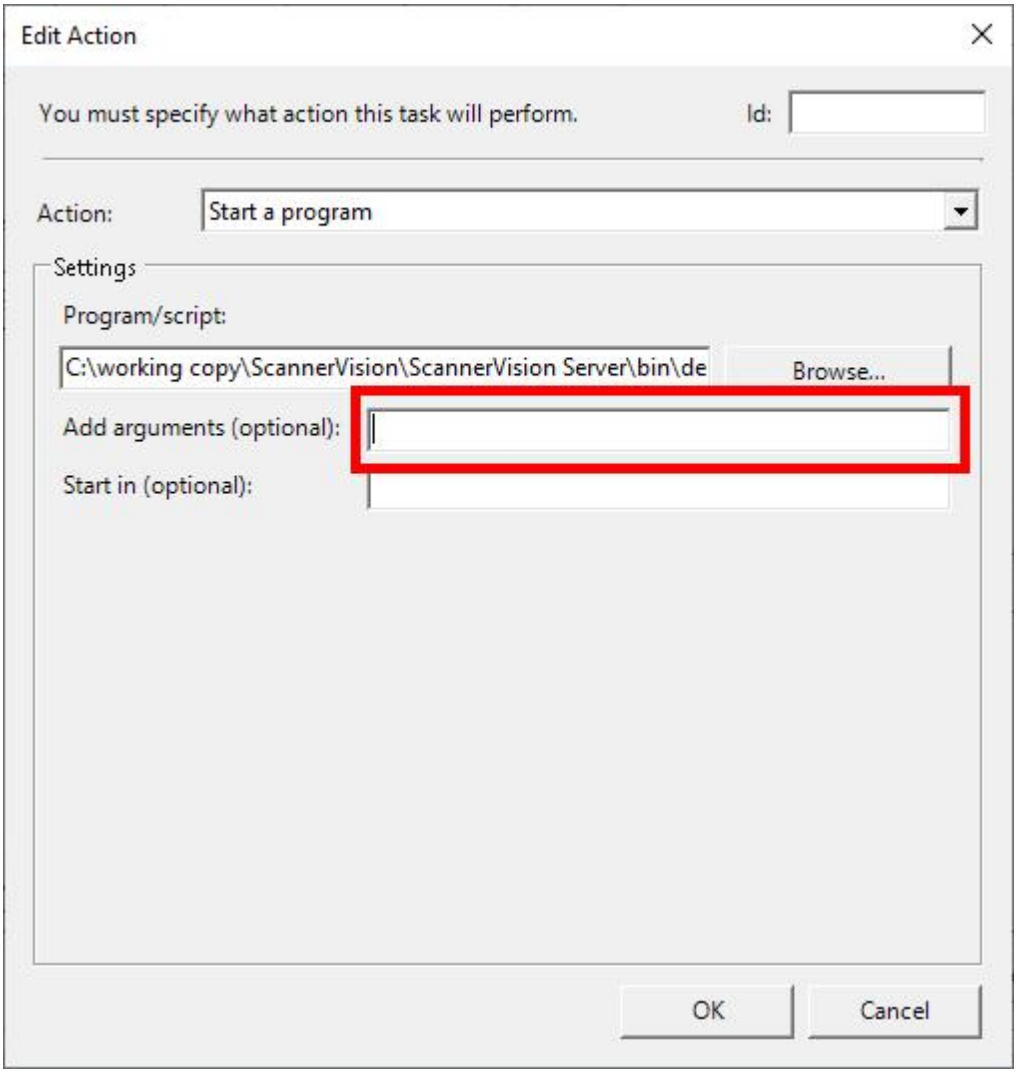
When configuring the user import schedule it is important to ensure that the import is performed with a user account that has the necessary privileges to access the ScannerVision settings database. This is usually the same user that configures templates and perform other tasks through the

ScannerVision user interface.

To ensure that a console window is not displayed when the scheduled import executes and that imports are performed while no user is logged on ensure your configuration looks like this:



Ensure that the "-c" command line parameter which keeps the console window open when the import is complete is not specified for the action:



6.5.3.1.7 ScannerVisionUserImporter.exe

The ScannerVisionUserImporter.exe is a console application that is installed in your ScannerVision installation folder. The utility is responsible for the importing of user data into the staging area as well as saving that data to the User table. The utility takes several command line parameters. If it is executed without any command line parameters all enabled import sources are executed and user data is saved to the User table.

Command line parameters are:

-c	Keeps the console window open when the import is complete.
-s	Imports user data into the staging area only.
-i	Saved user data that is in the staging area to the User table.
Instance name	The instance name you want to execute. If your instance name

contains spaces put it in quotation marks.

Examples

Import instance "DB" into the staging area only:

```
ScannerVisionUserImporter.exe BD -s
```

Import all instances into the User table:

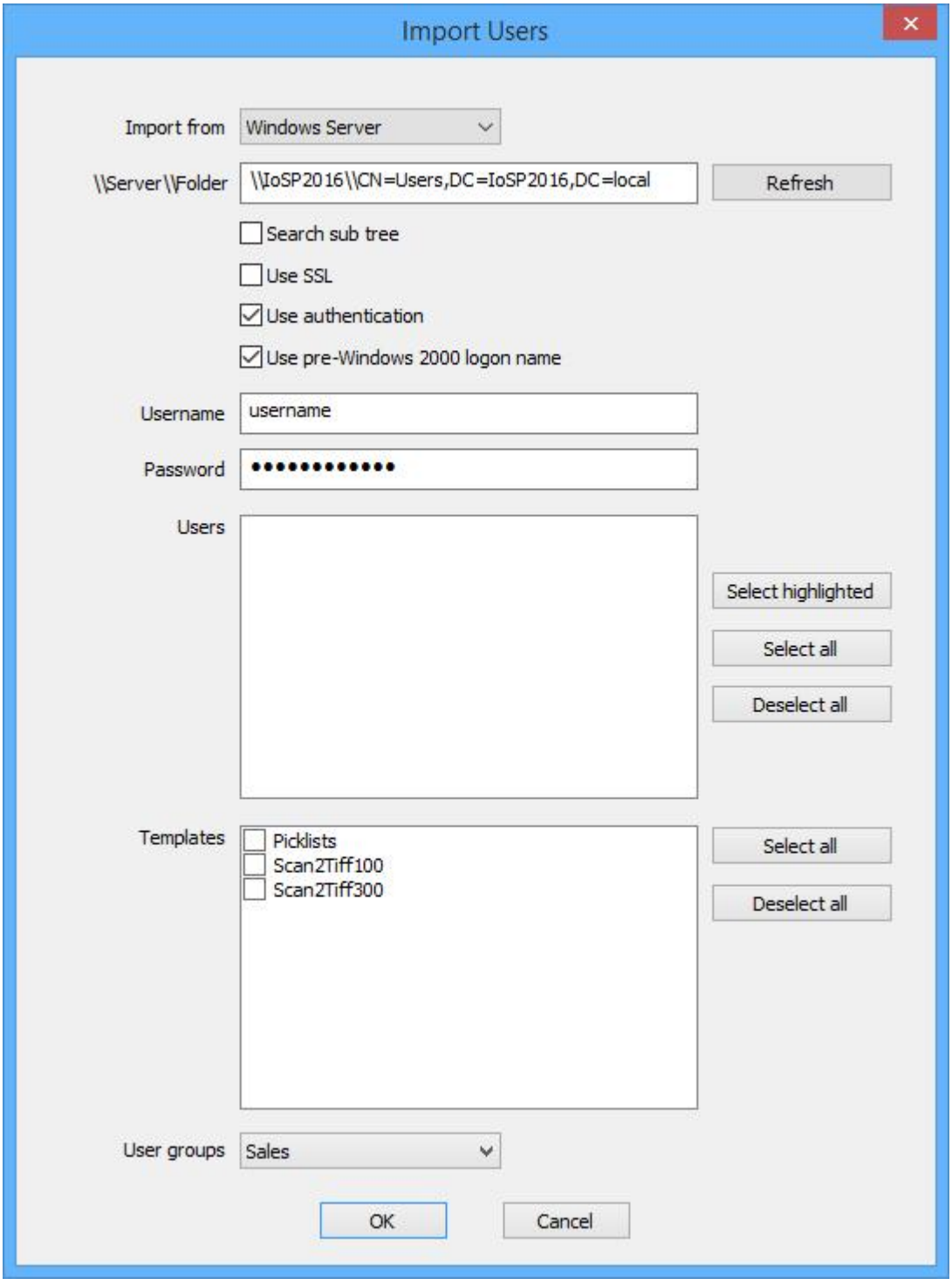
```
ScannerVisionUserImporter.exe -i
```

Perform complete import and keep the console open:

```
ScannerVisionUserImporter.exe -c
```

6.5.3.2 Windows/LDAP

With the Import Users tool you can import users from a Windows or LDAP server. The tool can be launched from the "Users->Import..." menu bar option.



Import from

The type of server from which you want to import users.

- Windows Server

- LDAP Server

\\Server\\Folder

The IP address or host name of the server to import from including any sub-folders if applicable e.g.

\\LDAPSERVER\\CN=Employees,DC=Company,DC=com

\LDAPSERVER\CN=Employees,DC=Company,DC=com

//LDAPSERVER//CN=Employees,DC=Company,DC=com

/LDAPSERVER/CN=Employees,DC=Company,DC=com

Search sub tree

Find users in sub directories.

Use SSL

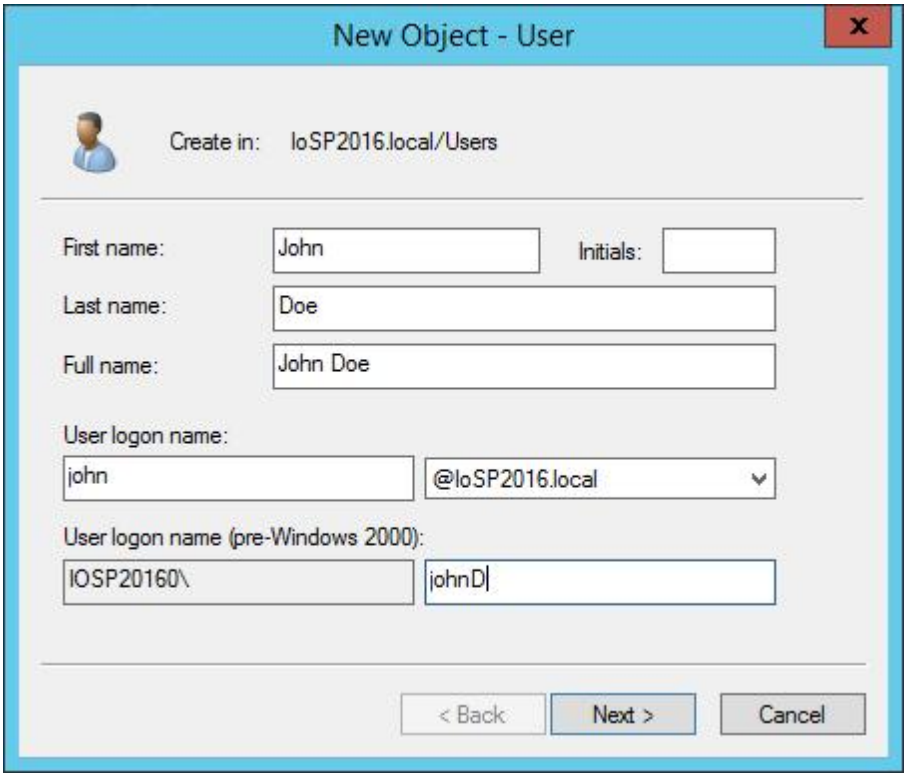
Communicate with the server using Secure Socket Layer communication.

Use authentication

Select this option to specify a Username/Password to log in to the server.

Use pre-Windows 2000 logon name

When importing from Microsoft Active Directory the "Use pre-Windows 2000 logon name" option determines which logon name is used as the ScannerVision user name as shown in the screen shot below of the new user creation dialog in Windows.



Username/Password

Login credentials for the import server.

Note

In the case of import from AD or LDAP, authentication against users IN the directory is not supported, i.e. you cannot use **OU=UserName,CN=Employees,DC=Company,DC=com** as the user name.

Users

Available users on the import server. Select the users you want to import if you have selected the "Only selected users" option.

Templates

The templates that have been configured in ScannerVision. Select templates you want to make available to the imported users.

User groups

The user group which you want users to be imported to.

6.5.4 Resetting Passwords

Users can change their password by navigating to the following URL:

https://[SERVER]:[PORT]/SetPassword where [SERVER] is the IP address or host name of the ScannerVision Networking Server and [PORT] is the SSL port you have configured in the Network Server Settings.

Forgotten passwords have to be reset by the ScannerVision administrator in ScannerVision itself.

6.5.5 User Based Licensing

User Based Licensing (UBL) is an add-on license option that allows you to bypass the configuration of Aivika Capture Pro clients in ScannerVision. This is useful in situations where many users are imported into ScannerVision. The license includes the number of concurrent and active users who are allowed to connect. You do however still have to configure MFP clients since ScannerVision needs to know to what type of device it is connecting. Each manufacturer has its own specification for communicating with the device.

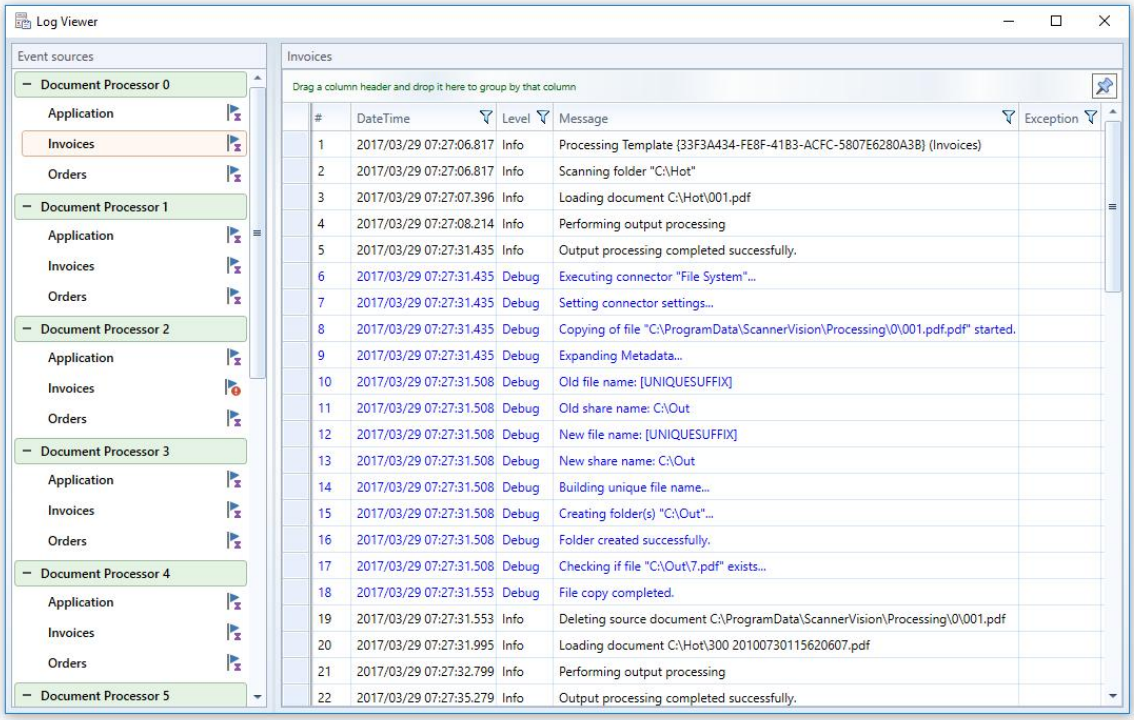
The users in the UBL are not named but concurrent which means that any user that is active in ScannerVision can connect provided that there are slots open. So if you have a 10 user UBL license and you have 20 users in your ScannerVision database, the first 10 of any of the 20 users in the database (provided the user is Active) will be able to connect. When a user logs out of the Aivika Capture Pro or Mobile client or closes the application, the connection slot on the server is freed up. The connection slot is also freed up after 5 minutes of inactivity from the client.

When Aivika Capture Pro is installed in Microsoft Terminal Server environment you have to have a UBL that includes the number of Terminal Server users. Aivika Capture Pro does not support client based licensing when it is running on Terminal Server.

6.6 Log

The ScannerVision log viewer is a stand-alone application that can be launched by clicking the "Log" button in the Navigation Pane or by selecting the "Log Viewer" menu option in the Windows start menu.

The log viewer give you a run time overview of the ScannerVision system with a little bit of history included - see below. To view detail historical log messages please refer to the "C:\ProgramData \ScannerVision\Logs" folder. The log files roll over on a daily basis and they are in CSV format so that you can open them in a spreadsheet application such as Microsoft Excel which allows you to filter, group and sort the messages.

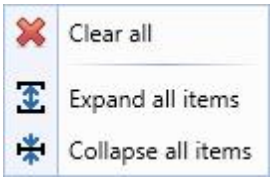


The "Event sources" list to the left of the screen shows the sub system from which the events have originated. In the screen shot above you can see several "Document Processor" entries - one for each Document Processor instance that you have licensed on your system. The entries underneath the root category represent the context within the sub system from where the messages originated. In the example above there are 2 contexts namely "Application", "Invoices" and "Orders". The Application context is standard in all sub systems and represents general operational messages such as during start-up, shut down etc. "Invoices" and "Orders" in the example above represent specific templates and would naturally be different in your installation. When you select a context you will see only the messages that were generated from the specific context.

Each context maintains a rolling buffer of the last 100 messages. When you select a particular context you will see all these messages immediately while new ones are appended to the bottom of the grid. The context you are viewing maintains a buffer of up to 5000 messages.

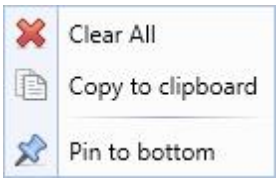
The messages view does not bring new message into view automatically. To do this you can select the  button. When you select a log message line the "pinning" stops.

The "Event sources" context menu gives you the following options:



Clear all Clears all entries in the Event sources and messages views. This only effects the log viewer and not the log files on disk.

Expand/Collapse all items Expands/collapses the root nodes of the "Event sources" tree view respectively.



Clear all Clears all log messages. This only effects the log viewer and not the log files on disk.

Copy to clipboard Copies the selected log messages to the clipboard.

Pin to bottom Keeps messages view pinned to the bottom so that new messages are always in view.

6.6.1 Error Codes

General

0	Function not successful
2	Successfully aborting
3	Function successful. Delete file before quitting
4	Function successful. Retry doing the failed op
8	Function called recursively. A least one call was successful, but some failed
-13	Invalid parameter passed
-789	The structure size is invalid
-27	Invalid bits/pixel

-11	Invalid width/height
-2	Invalid bitmap handle
-814	A NULL pointer was passed
-1360	Signed image data not supported
-290	The buffer size is too small
-149	An unspecified exception occurred - could be memory access violations
-351	There is nothing to do. No items was found
-144	Image is empty
-21	Invalid QFactor specified
-261	Invalid handle
-1364	32-bit Grayscale bitmaps not supported
-1486	The LUT length is invalid. LUTLength should be 1 << (HighBit - LowBit + 1)
-1491	The DIB is invalid
-100	User has aborted operation
-29	Escape key pressed - or user aborted operation
-1	Not enough memory available
-3	Not enough memory available (too low)
-16	Feature not supported
-143	Document capability is required
-804	Extended grayscale support is required
-315	JBIG capability is required
-10	File not found

-14	Not able to open file
-5	Error writing file
-6	File not present - abort
-7	Error reading file
-4	Error seeking to position
-295	not enough disk space
-805	File is read-only. Cannot open file with write access
-9	Invalid file format
-8	Invalid filename specified
-1490	Error creating file
-309	The DXF filter is required
-319	The JBIG filter is required
-349	WMF filter is required
-355	CMW filter is required
-356	CMP filter is required
-357	FAX filter is required
-358	PDF filter is required
-359	TIF filter is required
-819	DCR filter is required
-820	KDC filter is required
-821	DCS filter is required
-822	ABC filter is required
-823	ABI (ABIC) filter is required

-824	JB2 (JBIG2) filter is required
-541	J2K filter is required
-825	PNG filter is required
-807	LTCLR.DLL is missing
-353	LVKRN.DLL is required
-818	LTSGM.DLL is missing
-1479	LTANN.DLL is missing
-1480	LTIMG.DLL is missing
-1482	LTKRN.DLL is missing
-1483	LTFIL.DLL is missing
-1484	LTDIS.DLL is missing
-1485	LTDIC.DLL is missing
-815	The overlay does not exist
-816	Something is wrong with the overlay index

File Handling

-28	Invalid window size
-1353	Invalid filter name
-310	Page not found
-311	You cannot delete a page from a file containing only one page
-12	Image format recognized, but sub-type not supported
-77	Premature end-of-line code
-800	The PSD Layer is missing

-23	Invalid compression format
-78	Bad version number
-520	Thumbnail not found
-1492	The PSD Channel is missing
-530	Error in JP2 Box values
-531	The header does not match the JP2 signature - not a JP2 file
-532	JP2 file has a feature that is usupported
-533	Invalid save options were specified to the encoder
-534	File header does not contain SOC marker
-535	File contains complete header but no compressed image data
-536	Invalid save options were specified or file includes invalid encoded values
-537	Compression ratio, target file size, or tile size was too small for encoder
-538	Specified Wavelet decomposition level was too high
-539	Decoder could not translate J2K marker - file is corrupt or invalid
-540	J2K file has image with more than 30 bits per component
-65	stamp not found
-141	Stamp size is too big or invalid bits/pixel, etc
-142	Stamp is present but data is corrupted
-801	Bad JPEG marker
-806	Bad JPEG Resync marker
-808	The size of the JPEG marker cannot exceed 64K
-809	The required JPEG marker is missing

-811	The marker index is invalid (too big)
-810	This file does not contain Exif extensions

Color Profiles

-918	Invalid colorspace
-785	Invalid format
-786	Method not supported
-787	Error opening profile
-788	Invalid color profile
-791	U and V not multiples
-792	No non planar vertical sub sampling supported
-793	Planar alignment not supported
-794	Unsupported conversion
-795	Truncate height
-796	Truncate width
-797	Truncate width and height
-1150	Unknown ICC profile Tag
-1151	Unknown ICC profile Type
-1152	Unknown ICC profile Tag and Type
-1153	Invalid ICC profile
-1154	Invalid ICC profile
-812	The ICC profile was not found
-817	The ICC profile was invalid

-813	An error has occured while decoding the profile

PDF

-721	The PDF file is encrypted but no password has been specified.
-722	Invalid password specified
-723	FAX is required for this function
-724	CMP is required for this function
-725	Invalid Document Structuring comments (Ps and EPS)
-726	Could not find Fonts directory
-727	Cannot insert, delete, append or replace pages
-728	File is corrupted
-729	Either the files required for initializing the PDF engine were not not found or they were found but they incorrect

Dialogs

-150	An error occurred during creation of the common dialog
-151	You have already initialized the dialogs
-152	You did not initialize the dialogs
-153	You did not initialize the dialogs
-154	The dialogs could not load the LTCLR.dll
-155	The dialogs could not load the LTDLGRES.dll
-900	LTCLR.DLL is not loaded
-901	LDLGRES.DLL is not loaded

-902	Dialogs not correctly initialized

Annotation

-202	Annotation DLL not loaded
-200	LTDIS.dll not loaded
-201	LTFIL.dll not loaded
-148	One or more annotation objects are still locked

TIFF

-15	Unknown compression format
-917	The compact function was finished without copying all the pages
-140	Bad TIFF tag
-404	Tag not found
-1365	Cannot add tags to a IFD selected

Document

-1200	DLLs are not loaded
-1201	Specified index is not corrected
-1202	invalid fill method
-1203	Invalid recognition module
-1204	invalid character filter
-1205	invalid zone type
-1206	invalid language Id

-1207	invalid spell language
-1208	invalid enumeration section
-1209	invalid page count
-1210	Illegal internal code as a parameter
-1211	Set parameter is not acceptable
-1212	End of list of the available Code Pages
-1213	Error while loading the Code Page Definition file
-1214	The length of the exported code exceeds the buffer size
-1215	Character conversion is not available for the given character
-1216	Conflict: The selected Code Page does not support some characters in the selected languages. There is no exact code in the Code Page for them
-1217	Character Set and Code Pages module initialization error
-1218	OCR engine initialization error
-1219	OCR engine termination error
-1220	Module initialization warning
-1221	Application has aborted the current process
-1222	Application has terminated the current recognition process without losing the recognized text
-1223	Module is not present
-1224	OS could not load a module
-1225	Missing entry in a module
-1226	Invalid module
-1227	Module initialization error

-1228	The requested function is not available, or there is no appropriate license
-1229	General error in the engine
-1230	General Protection Fault in the engine
-1231	Not supported operational system
-1232	Syntax error in the specified engine Settings file
-1233	Invalid setting
-1234	The engine is busy
-1235	Recognition process Time Out
-1236	Internal error in image module
-1237	Not enough memory during image processing
-1238	Invalid rectangle dimensions
-1239	Non-supported resolution
-1240	Cannot process compressed image
-1241	Invalid bitmap address
-1242	Unsupported BitsPerPixel value
-1243	Internal error in image module
-1244	Invalid image handle
-1245	Buffer overflow during processing the image
-1246	Image operation denied
-1247	No more lines in defined image area
-1248	Non-supported image size
-1249	Engine manager module error

-1250	No recognized text available, either because the zone is empty or the required recognition module has not been initialized properly
-1251	There is no selected recognition module
-1252	There is no zone in the zone list
-1253	Invalid zone index
-1254	Invalid zone coordinates
-1255	MOR recognition module initialization error
-1256	MOR recognition module's knowledge base file not found
-1257	MOR recognition module's knowledge base file corrupted
-1258	Incorrect knowledge base file version
-1259	Size of image is too large
-1260	File is corrupted
-1261	Internal error in the MOR recognition module
-1262	General Protection Fault in the MOR recognition module
-1263	Dot-matrix recognition error
-1264	Checkmark recognition error
-1265	HNR module's knowledge base file is corrupted
-1266	Parameter is out of valid range
-1267	Internal error in the RER module
-1268	RER module's knowledge base file not found
-1269	Character Set is empty for the RER module
-1270	The specified Character Set not fully supported by the RER recognition module
-1271	RER specific file not found

-1272	Not enough memory for the SPL module
-1273	Uninitialized spell object
-1274	Spell file open error
-1275	Spell file read error
-1276	User dictionary write error
-1277	Invalid file format
-1278	Module initialization error
-1279	User dictionary close error
-1280	Illegal language setting
-1281	No more suggestions/items available
-1282	Internal checking error
-1283	Item already exists in the User dictionary
-1284	Item doesn't exist in the User dictionary
-1285	Item wasn't inserted into the User dictionary
-1286	Section does not exist in the User dictionary
-1287	Regular expression syntax error
-1288	Output format conversion subsystem was not initialized
-1289	No more converters available
-1290	There is no converter DLL file
-1291	Wrong parameter or parameter not found
-1292	TMP file is corrupted
-1293	TMP file not found
-1294	Cannot create TMP file

-1295	Cannot seek in TMP file
-1296	TMP file read error
-1297	TMP file write error
-1298	Cannot close TMP file
-1299	Cannot load the DLL file
-1300	Internal error in ATMTXT module. (End-Of-File detected)
-1301	Not enough memory for ATMTXT/TMP module
-1302	Image is not available for inserting into the output
-1303	Not enough memory to register an info item
-1304	Output file open error
-1305	Output file read error
-1306	Output file not found
-1307	Output file write error
-1308	Invalid command sequence in the ICF file
-1309	Cannot create the output file
-1310	Not enough memory for ATMTXT/TXT module
-1311	Invalid document type
-1312	File too large
-1313	There was a WARNING during the output file conversion
-1314	Operation cancelled
-1315	Illegal option specified
-1316	Encrypted source
-1317	Output file seek error

-1318	Braille recognition module initialization error
-1319	Matrix matching recognition module initialization error
-1320	M/TEXT recognition module initialization error
-1321	Non-implemented feature
-1322	couldn't find specific zone
-1323	No available selected zone
-1324	can't export the specified page
-1325	can't lock the specified page
-1326	can't set the specified page to be activated.
-1332	Too many regions result when layout analysis
-1333	Top or bottom horizontal line not found
-1334	The slope of line too large
-1335	Frame has not been detected, so can not add virtualline
-1336	Too many lines
-1337	Too many pages in batch processing list
-1338	This region is not a table region
-1339	This region is a table region but contains no data
-1340	This line is not valid (neither horizontal nor vertical or type is not match)
-1341	Angle is too large, can not rotate image
-1342	Unable to determine orientation
-1343	Invalid RDF file name
-1344	PDF DLLs are not loaded

-1346	ICR Module is missing
-1347	OMR Module is missing
-1348	Languages files are missing
-1349	OCR engine initialization error
-1350	ICR Module is missing
-1351	OMR Module is missing
-1352	Languages files are missing

TWAIN

-80	TWAIN Failure due to unknown causes
-81	TWAIN Not enough memory to perform operation
-82	TWAIN No Data Source
-83	TWAIN DS is connected to max possible apps
-84	TWAIN DS or DSM reported error
-86	TWAIN Unrecognized MSG DG DAT combination
-90	TWAIN Operation was cancelled
-560	No TWAIN Library
-561	Invalid DLL
-562	TWAIN is not initialized
-564	Check status
-565	End of list
-566	Cap is not supported
-567	Source is not open

-568	Bad value
-569	Invalid state
-570	Caps neg not ended
-571	Open file
-572	Invalid handle
-573	Write to file
-574	Invalid version number
-575	Read from file
-576	File is not valid
-577	Invalid access right
-578	Custom base
-579	Denied
-580	File exists
-581	File not found
-582	Not empty
-583	Paper jam
-584	Paper double feed
-585	File write error
-586	Check device on line
-587	Stop scan

Barcode

-410	Invalid error check digit

-411	Invalid bar code type
-412	Invalid bar code text out option
-413	Invalid bar code width
-414	Invalid bar code height
-415	Bar code string is too small
-416	Invalid bar code string for a specified bar code type
-417	No bar code recognition
-418	Invalid bar code measurement unit
-419	Invalid multiple max count
-420	Invalid bar code group
-421	Invalid bar code data structure
-422	No duplicated bar code
-423	Reached the last duplicated bar code
-424	Invalid bar code data string length
-425	Invalid bar code area location
-431	Barcode PDF417 symbol is found but cannot read successfully
-1382	Bar Code Engine DLL not found
-1481	Incorrect module value it should be >= 0

Automation

-690	Invalid handle
-691	Invalid state

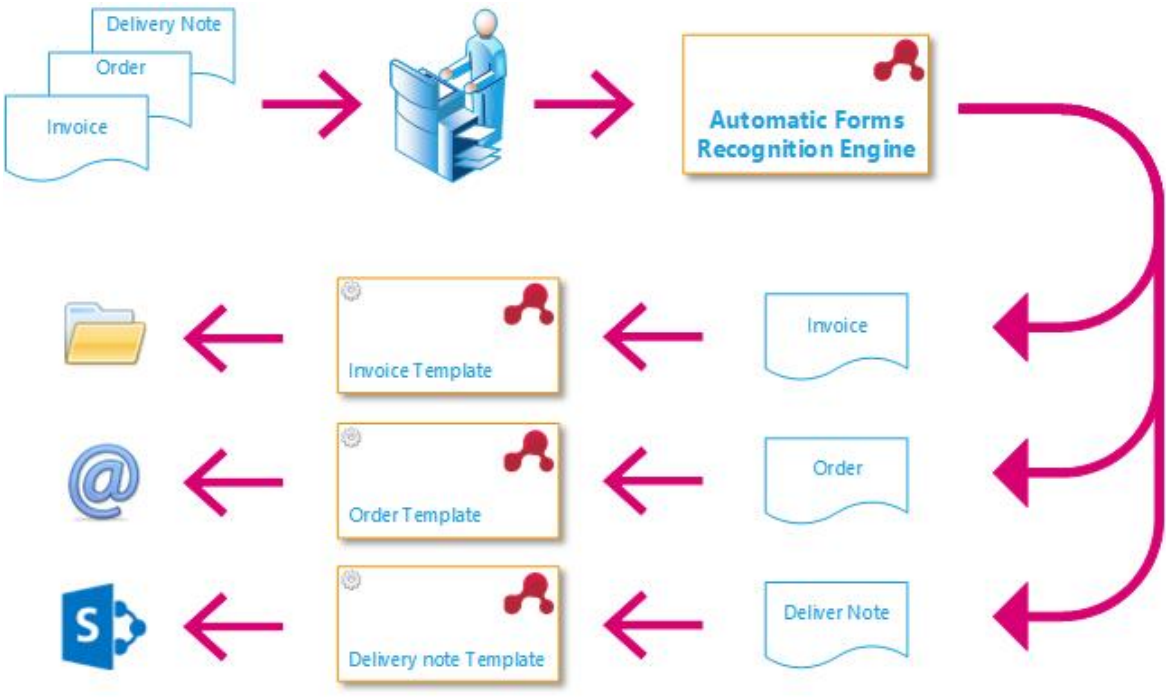
Container

-630	Invalid state/handle
-631	Invalid operation
-632	No resources available

7 Automatic Forms Recognition

The ScannerVision Automatic Forms Recognition system allows you to recognize documents based on the contents of the document such as delivery notes, order and leave forms. In fact, any document that contains a significant amount of static information can be saved in ScannerVision as a Master Form. Master Forms are the reference documents against which the Automatic Forms Recognition Engine compares incoming documents and they can contain multiple pages. You can scan a batch of documents of different types and when ScannerVision Processing Server processes the batch each type of document is recognized and the batch is split at that position. If a document cannot be recognized it is copied to the rejected folder. Recognized documents are dropped in the Hot Folder of a pre-configured Workflow Template for processing. Each active Master Form must be assigned to a Workflow Template for processing.

The graphic below depicts the Automatic Forms Recognition processing work flow.



Automatic Forms Recognition Template vs Workflow Template

Automatic Forms Recognition Templates are, as the name suggests, templates that pass captured documents through the Automatic Forms Recognition Engine. Forms that have been processed by a Automatic Forms Recognition Template are analyzed by the Automatic Forms Recognition Engine and when recognized dropped into the Hot Folder of the Workflow Template that is assigned to the Master Form. A Automatic Forms Recognition Template provides a subset of the configuration options of a regular Workflow Template namely:

1. Capture
2. Xml Import
3. Notification Settings

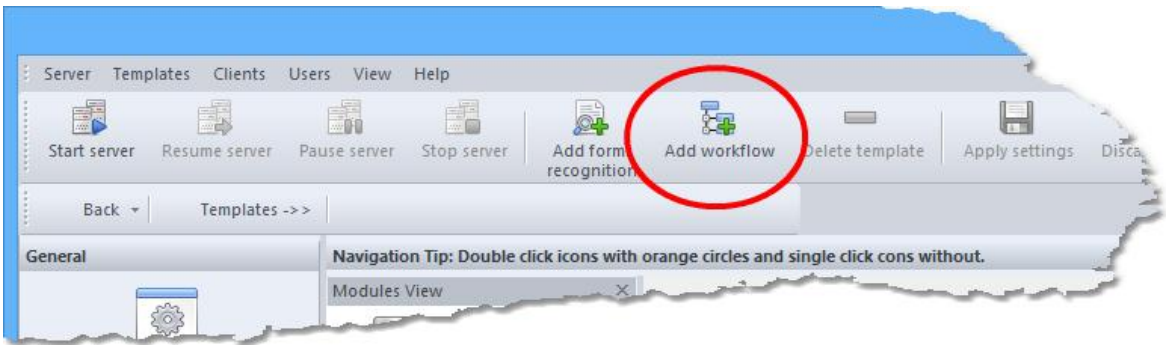
These functions are 100% equivalent to those in the Workflow Templates in terms of both configuration and processing. Although you have the ability to create as many Automatic Forms Recognition Templates as you want one can take care of all your Automatic Forms Recognition processing.

Configuring Automatic Forms Recognition

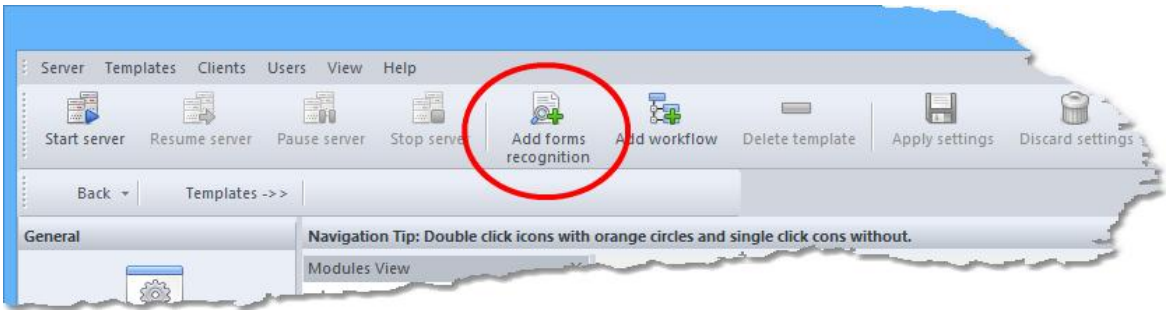
Configuring Automatic Forms Recognition involves 3 steps:

1. Configure Workflow Template for processing recognized documents. You should do this first because you won't be able to activate a Master Form without assigning a Workflow Template.
2. Configure Automatic Forms Recognition Template for capturing documents. This step can be performed first, second or last as it is not dependent on the existence of Master Forms or Workflow Templates.
3. Configure Master Forms to compare documents against.

To configure a Workflow Template click on the "Add workflow" button shown below:



To configure a Automatic Forms Recognition Template click on the "Add forms recognition" button shown below:



For both Automatic Forms Recognition Templates and Workflow Templates follow the relevant steps described in the Creating Templates section.

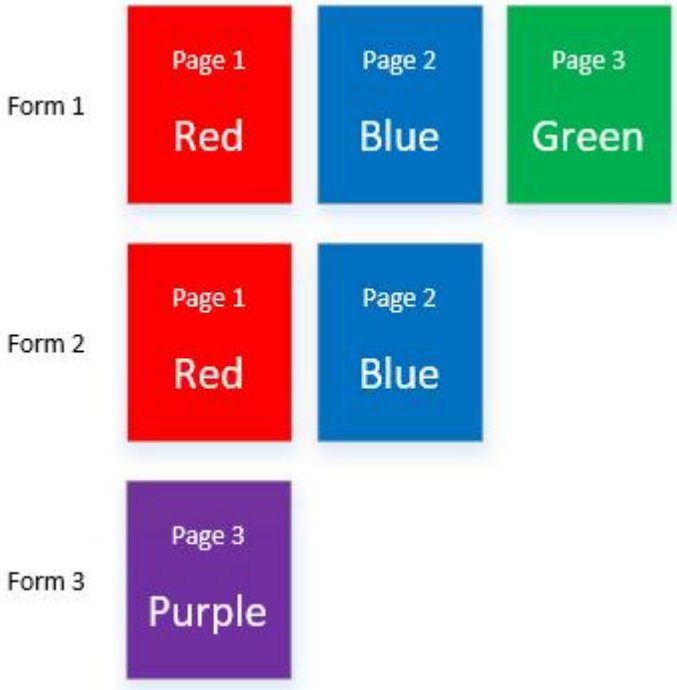
To configure Master Forms follow the steps described in the Creating Master Forms section.

7.1 How does it work?

The Automatic Forms Recognition Engine operates against a set of Master Forms which are the signatures of the forms that the engine can recognize. Master Form can be single-paged or multi-paged.

When an empty/blank form is loaded into the Master Forms Creator the contents of each page is analyzed to determine the content and its location on the form. This information is stored as the signature for the page. The signatures of all the pages in the form together constitute the Master Form.

In the diagram below we represent the signature of each page in the Master Form as a color. We have 3 Master Forms with 3, 2 and 1 page(s) respectively with the signatures Red/Blue/Green, Red/Blue and Purple respectively.



Now, let us assume we are scanning a batch of documents that have the following composition:

When the Automatic Forms Recognition Engine starts its analysis of the scanned document it know nothing about its composition and therefore it can contain any of our Master Forms. So, what the engine does is it creates a list of candidate Master Forms and add all the known Master Forms to the list. In our example this list would therefore contain Form 1, Form 2 and Form 3. The engine now compares Page 1 of the scanned document against the first page of all the Master Forms in the candidates list. When the engine does this comparison it calculates a number that represents the certainty that the signatures of the 2 pages are the same. When we created the Master Forms we had to specify the level of certainty we want the engine to apply. So if we specified a certainty of 90% and the engine calculated a certainty of say 94% for the 2 pages it is comparing then it will deem the pages to have the same signature. If the certainty was 89% then the engine will say the signatures are not the same.

If we tabularize the result of the analysis of Page 1 in the scanned document is may look something like this:

Master Form	Certainty Page 1

1	99%
2	99%
3	3%

With a certainty of only 3% Form 3 falls out of the candidates list. So we are left with Form 1 and Form 2 in the candidates list. The engine now moves to page 2 of the scanned document and repeats the process with the following results:

Master Form	Certainty Page 2
1	99%
2	99%

Both Form 1 and Form 2 remains in the candidates list and the engine moves to page 3 of the scanned document. The results are as follows:

Master Form	Certainty Page 3
1	99%

Since Form 2 only has 2 pages there is nothing to compare against for page 3 of the scanned document. Since a match was found for page 3 of Form 1, Form 2 falls out of the candidates list because there is another Master Form with more matching pages. We are therefore left with only one candidate and since matches were found in the scanned document for all the pages of Form 1 we have found a match of Form 1.

What remains now is to determine where the scanned document needs to be split. By convention, the Automatic Forms Recognition Engine splits a document only when a new match is found of a Master Form. All pages in the scanned document that follows after the Master Form up to the start of the new form will be part of the first document. The Automatic Forms Recognition Engine will therefore split the scanned document between pages 5 and 6 as follows:

Now that a successful match and split operation has been completed the whole process starts over again. The Automatic Forms Recognition Engine treats the remaining pages (page 6 and 7) of the scanned document as if is a completely new document as follows:

All the Master Forms are added back to the candidates list and the analysis starts again. For this "new" scanned document we end up with the following results:

Master Form	Certainty Page 1	Certainty Page 2	Certainty Page 3
1	99%	99%	99%
2	99%	99%	

We are left with 2 candidates and no more pages to compare with. The Automatic Forms Recognition Engine now has to decide which of the 2 candidates is the best match. Since Form 1 contains 3 pages and Form 2 only 2 it means that **all** of the pages of Form 2 have been found and only 2 out of 3 for Form 1 and therefore Form 2 is the better match.

Other scenarios

Scenario 1

We have 2 Master Forms each with 2 pages and a scanned documents is analyzed yielding the following results:

Master Form	Certainty Page 1	Certainty Page 2

1	99%	99%
2	98%	99%

Technically we have 2 candidates with matches found for all their respective pages. What the Automatic Forms Recognition Engine does now is to add the certainty values of the signatures of each page as follows:

Form 1: 99 + 99 = 198

Form 2: 98 + 99 = 197

Form 1 has a higher total and is therefore chosen to be the best match.

If the total of the 2 candidates were the same we technically have a duplicate Master Form and the Automatic Forms Recognition Engine selects the first one.

Scenario 2

We have 2 Master Forms each with 2 and 3 pages respectively and a scanned documents is analyzed yielding the following results:

Master Form	Certainty Page 1	Certainty Page 2	Certainty Page 3
1	95%	95%	95%
2	99%	99%	

In this scenario we have 2 candidates. The first candidate has a higher number of matching pages but at a lower level of certainty than the other candidate with matches found for all its pages and at a higher level of certainty.

The Automatic Forms Recognition Engine will select Form 1 since more pages match - regardless of their level of certainty (assuming of course the certainty is above the minimum level set for the Master Form).

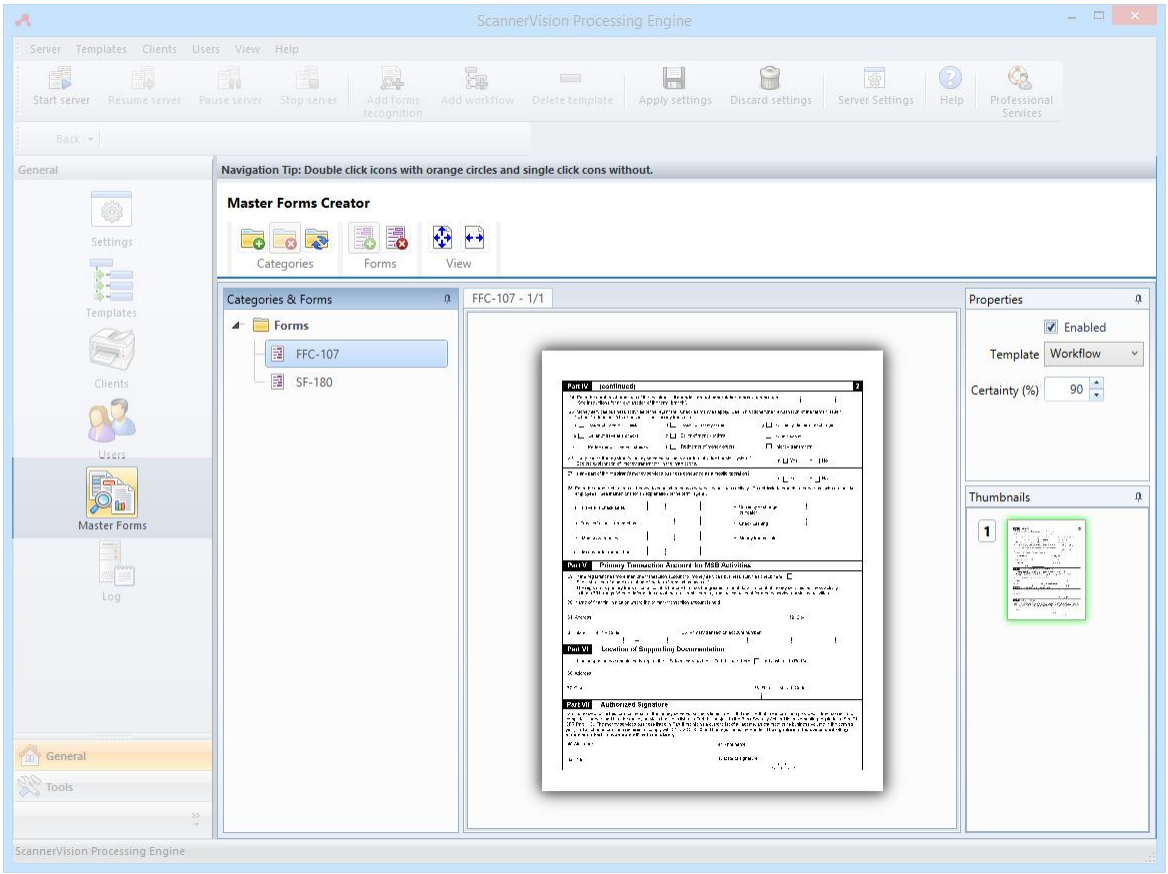
Scenario 3

We have a scanned document with the following composition:

What happens to pages 1 and 2? The convention is that pages *following* a recognized Master Form form part of the document. Pages *before* a Master Form are discarded.

7.2 Creating Master Forms

To create Master Form click on the "Master Form" button in the Navigation Pane on the left as shown below. This will open the Master Forms Creator screen.



To create a Master Form follow these steps:

1. Create a new category or select a the category to which you want to add the master form.
2. Click the  button and navigate to the document you want to be recognized. This should be a clean, uncompleted version of the document.
3. Select the Workflow Template in the "Templates" drop down box in the Properties area.
4. Select the level of certainty that you want the Automatic Forms Recognition Engine to apply that a form is the correct one.
5. Press the "Apply settings" button on the toolbar to save your changes.

Thumbnails

The "Thumbnails" area shows thumbnails of the pages in the Master Form. When a Master Form contains multiple pages you can preview individual pages by selecting its thumbnail.

Toolbar

The functions of the button on toolbar at the top of the Master Form editor are described below:

	Create a Master Form category.
	Delete a Master Form category and all children.
	Reload categories and forms.
	Add a Master Form.
	Delete a Master Form.
	Fit Master Form preview.
	Fit Master Form preview width.

8 Creating Templates

ScannerVision contains 2 types of templates namely Workflow Templates and Automatic Forms Recognition Templates. A Automatic Forms Recognition Template provides a subset of the configuration options of a regular Workflow Template namely:

1. Capture
2. Xml Import
3. Notification Settings

The discussion that follows is based on Workflow Template but the functions above are 100% equivalent to those in the Workflow Templates in terms of both configuration and processing.

Templates determine what processing ScannerVision performs. They determine where documents come from, what processing is done on them and where they ultimately end up. You can define as many templates as you want and they don't have to function in isolation. You could for example configure a template to read a barcode and based on the value of the barcode route the document to a specific folder from where it will be picked up by another template which knows how to process that particular type of document.

There are 3 primary functions every template performs. They are:

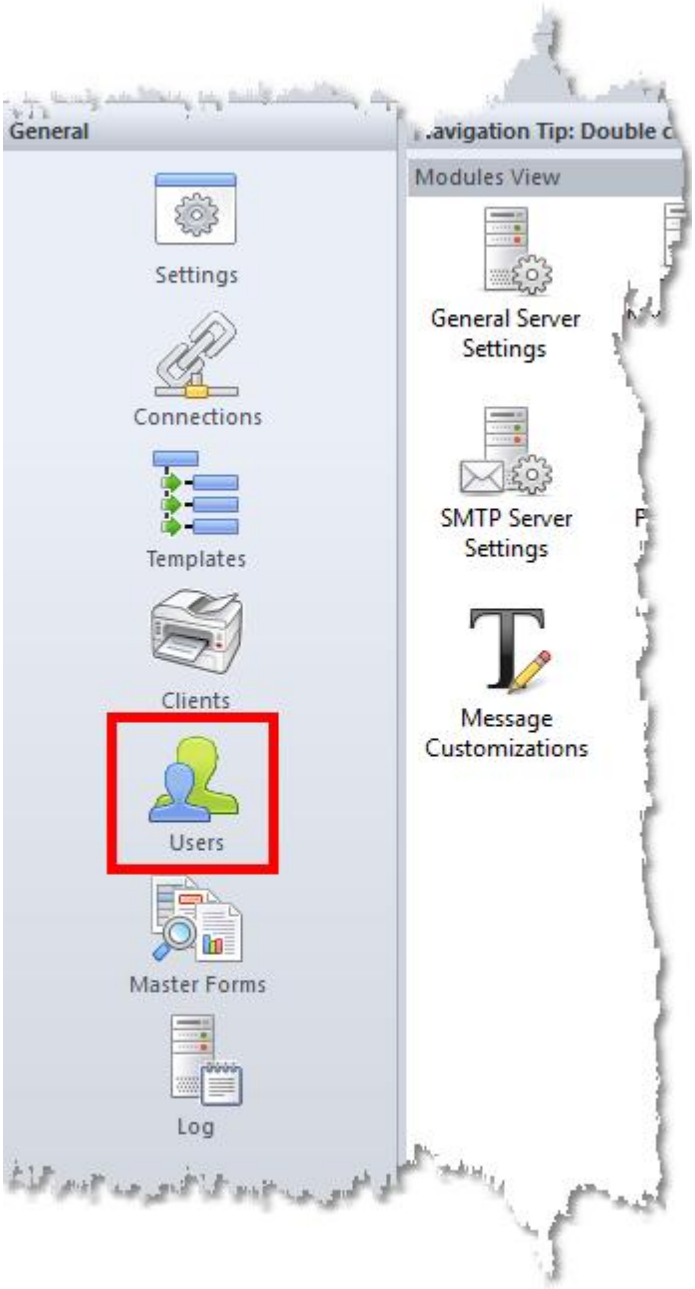
- Capture

  This determines where documents come from such as MFP or Desktop Clients, Ftp, Hot Folder etc.
- Process

  This determines what is done to the document. This could include barcode reading, OCRing, splitting, conversion to a different format etc. and much more.
- Store

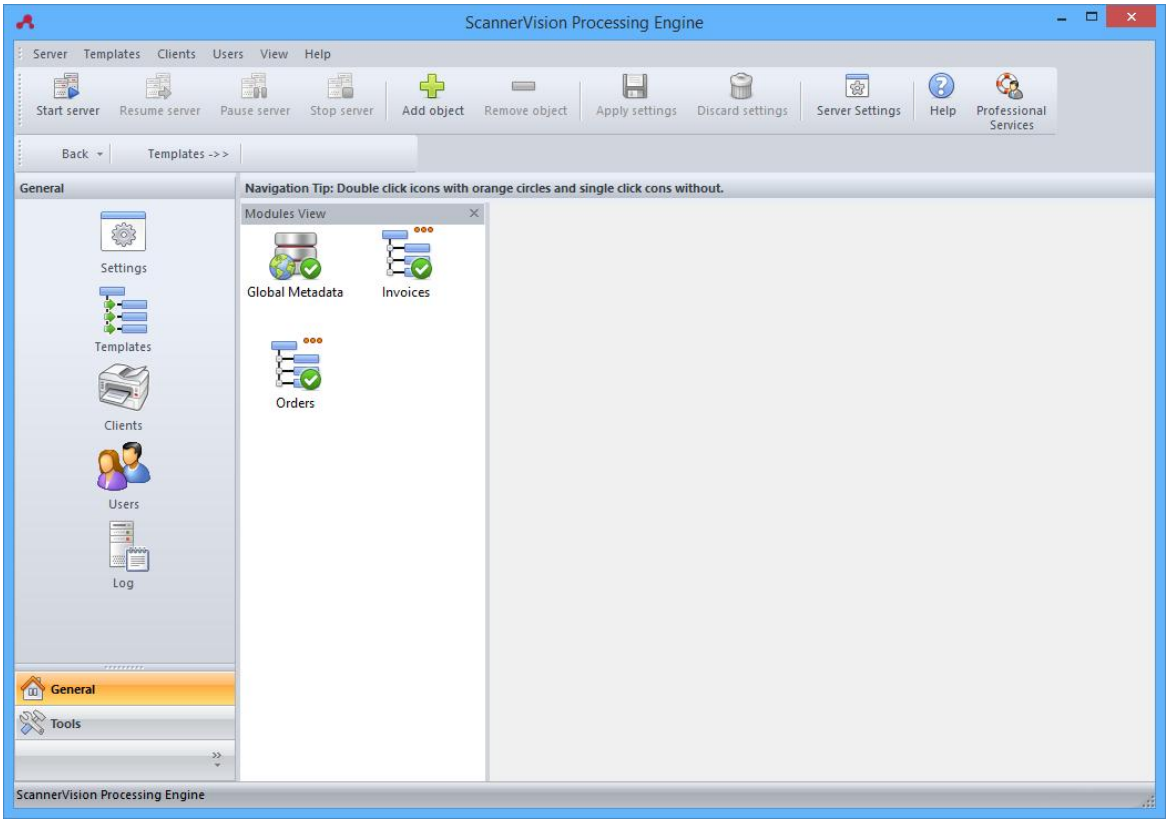
  This determines where documents finally end up which could be another folder on the server or network, a database, email or a document management system such as SharePoint, Laserfiche or AivikaOne.

These functions are discussed in detail in the relevant sections that follow. To access the template configuration area click the "Templates" icon in the Navigation Pane:



This will open the templates configuration screen. Initially you will only see the Global Metadata icon.

To add a new template click the "Add object" button in the toolbar.



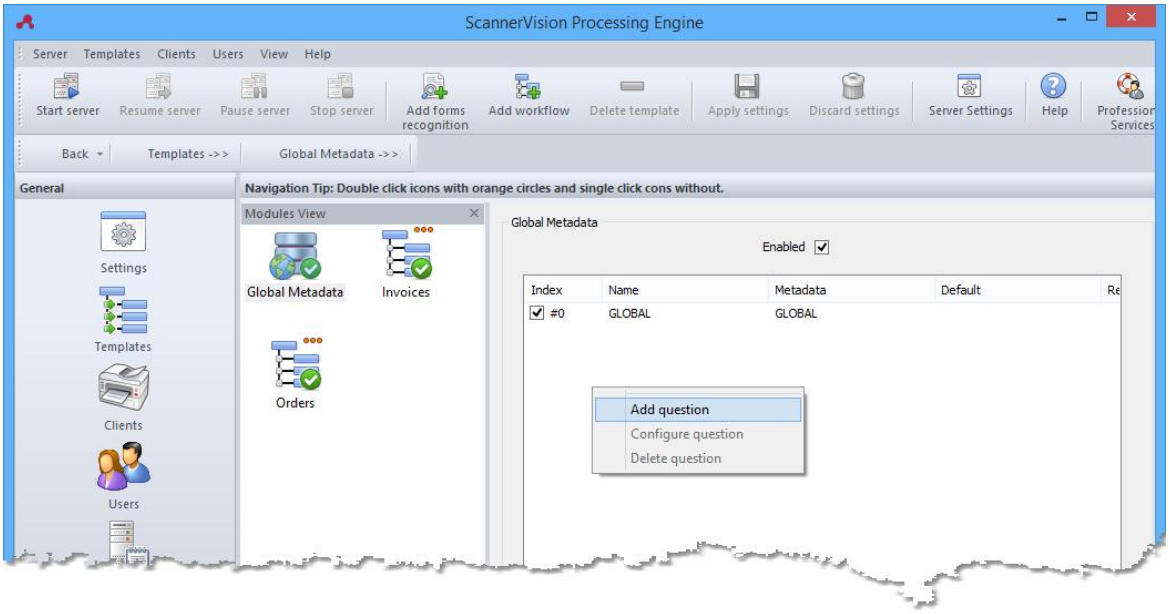
To configure a template double click the respective icon in the modules view.

8.1 Global Metadata

Please refer to the Capturing Document Metadata section for an overview of document metadata and how it can be configured.

Questions that are configured on the "Global Metadata" screen apply to all templates.

To add a global metadata question right click in the list window of the "Global Metadata" screen and select the "Add question" menu option:



To configure or delete an existing question, right click on the relevant question and choose the desired option.

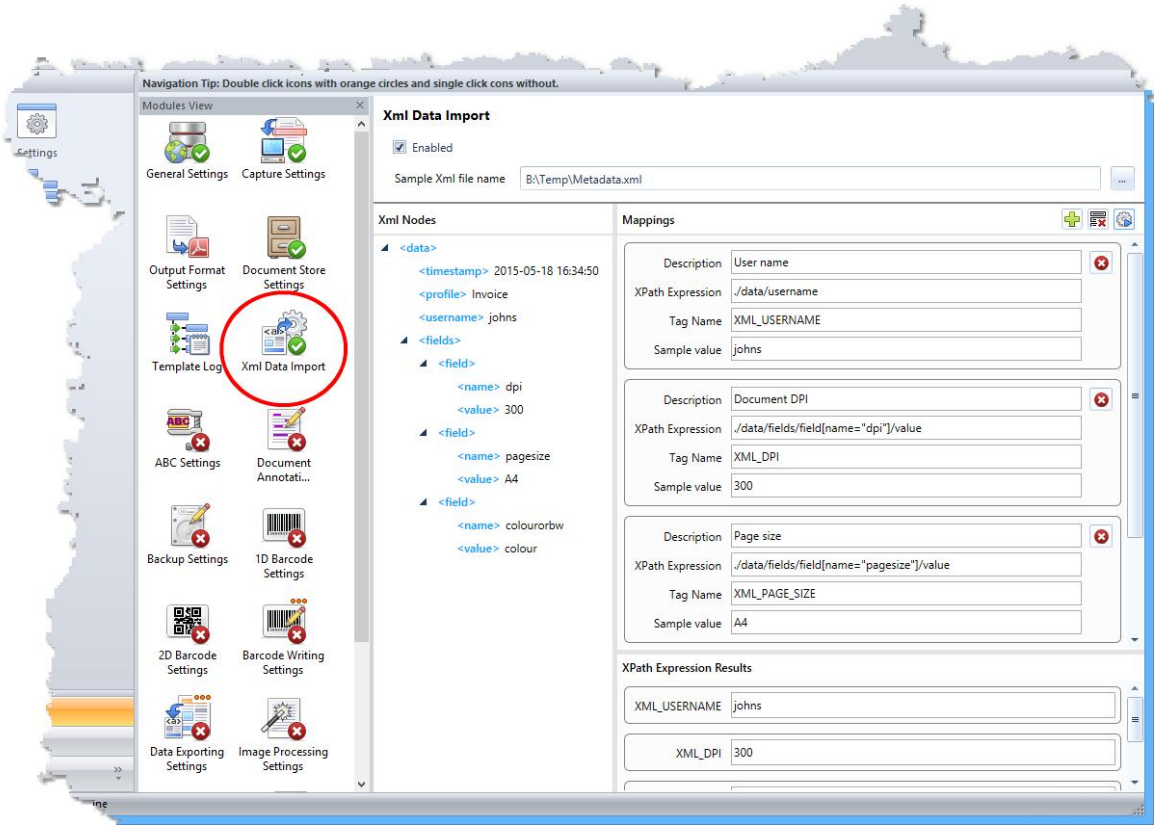
8.2 Importing From Xml file

ScannerVision allows you to import metadata from an Xml file when capturing from Hot Folder. The name of the Xml file should be the same as that of the document being processed but with a ".xml" extension.

In order to import metadata from an Xml file you have to map each node in the Xml file that you want to import to a unique metadata tag. You have to specify an XPath Expression that will return exactly zero or one Xml element.

Xml metadata is imported before any document processing is done so the metadata will be available for use from the start of the processing.

Below is a screen shot of the Xml Data Import screen.



Enabled

Enables/disables the importing of XML metadata.

Sample Xml File name

The fully qualified path to a sample Xml metadata file with which to configure mappings and it should remain in the location pointed to. It becomes part of the template. When the template executes the XML file name of the file to import is determined automatically. The name of the Xml file should be the same as that of the document being processed but with a ".xml" extension.

Xml Nodes

This is a tree view of the XML data contained in the sample file. The tree shows the name of the XML node in blue font and between angle ("<", ">") brackets with the value of the node in black text next to it.

Mappings

The Mappings section contains a list of the mappings you have configured.

XPath Expression Results

Shows the results of your XPath expressions. Click the  button in the top right corner of the "Mappings" area to test your expressions.

Opening a sample file

To open a sample XML file click the "..." button next to the "File name:" edit box and navigate to the sample file. The path to file is saved in the template and it will be loaded automatically when the template is selected. It is therefore important to ensure that the sample file remains in the location specified.

When a sample file is opened you will be prompted to remove previously configured mappings if there are any.

To add a mapping

You have 3 ways of adding a mapping:

1. Double click on the XML element in the "Xml Nodes" tree.

2. Drag the XML element from the "Xml Nodes" tree into the "Mappings" area.

3. Select an element in "Xml Nodes" tree and then press the  button in the top right corner of the "Mappings" area. The  will be disabled if no element is selected in the "Xml Nodes" tree.

When you have added a new mapping the mapped element is populated with default data obtained from the sample XML file. The 2 most important properties of the mapped element are "XPath Expression" and "Tag Name". The XPath Expression property contains the XPath expression that will be executed during document processing to obtain a value for the ScannerVision metadata. The Tag Name property, as the name suggests, contains the name of the tag that is created. This tag name has to be unique in the template. The "Description" and "Sample value" properties are used by the ScannerVision Expression Editor. The Description property represents the tool tip when you hover over the tag while the Sample value represents the value that is given to the tag when you select it.

To remove a mapping(s)

To remove a single mapping click the  button of the element you want to remove.

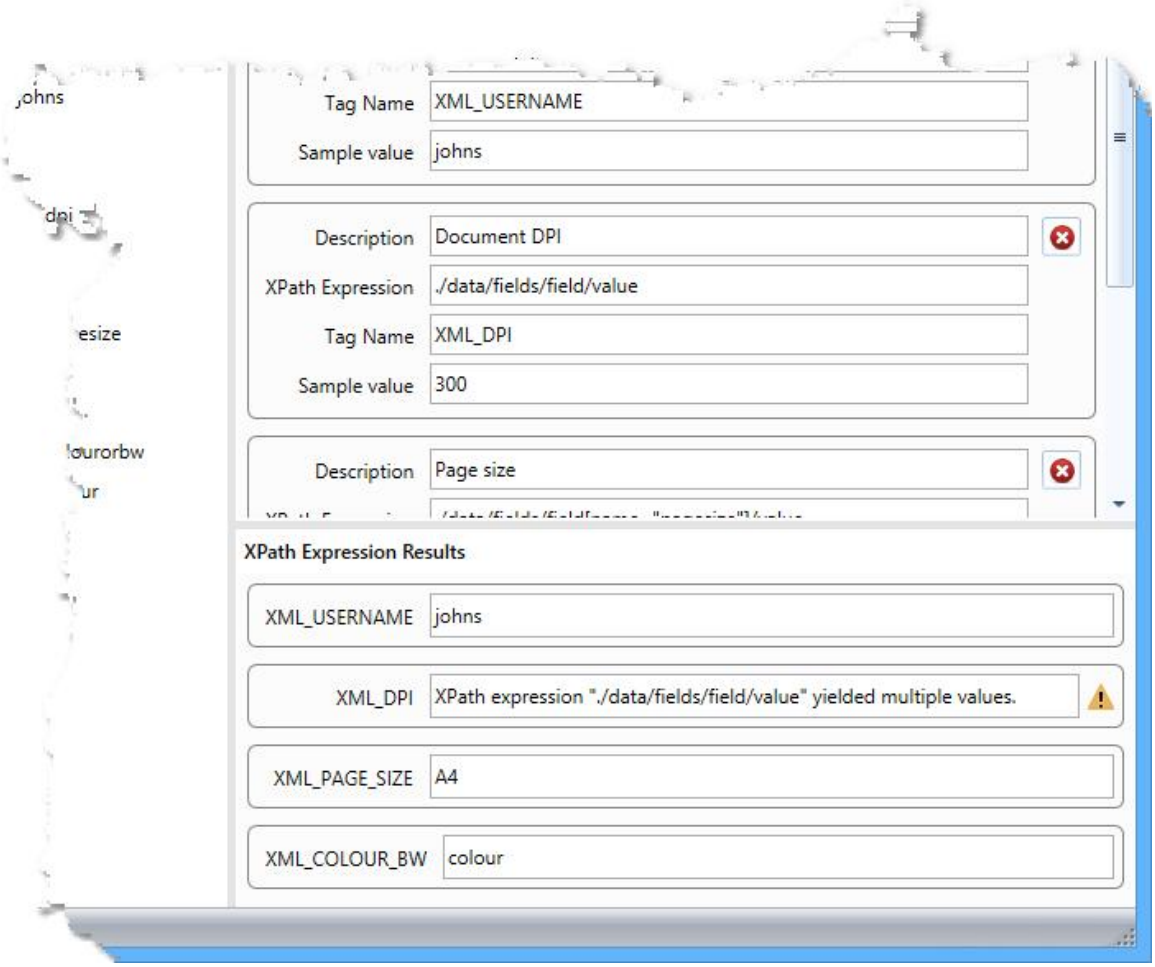
To remove all mappings click the  button in the top right corner of the "Mappings" area.

To test your XPath expressions

To test if you XPath expressions return the values you expect click the  button in the top right

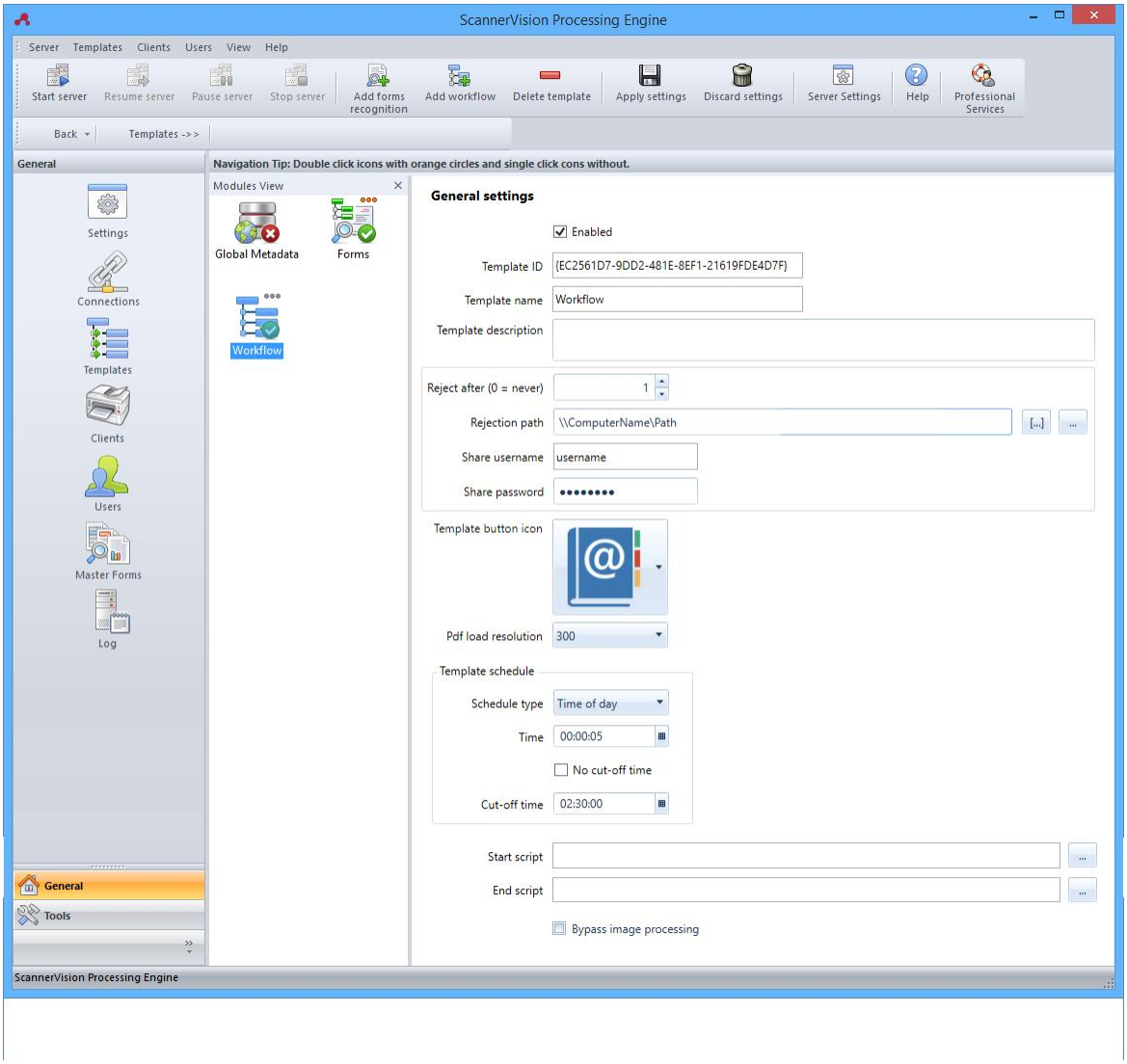
corner of the "Mappings" area. The results are shown in the "XPath Expression Results" area shown below.

Pay particular attention to elements that have an exclamation mark next to them. These expressions did not succeed. The reason for the failure is shown next to the tag name.



8.3 Template General Settings

On the General Settings screen you configure general template information such as the template name, the icon to show on clients and the template execution schedule.



Active

Enables/disables the template. Disabled templates are not executed.

Template ID

The identifier of the template.

Template name

You can (and should) give your template a meaningful name. This will make it easier to identify templates when you have many.

Template description

This is where you can put any additional information about the template to aid yourself or fellow administrators. The value you enter here is for your own use and has no effect on the functioning of the template.

Template button icon

Bitmap images can be assigned to a template to aid in their identification on desktop and MFP clients. You can create your own images 96x96 pixels in size. Put them in the "C:\ProgramData \ScannerVision\Buttons" folder. They will be available after you have restarted the application. The color of the bottom left pixel is used as the transparency color and every pixel in the image with this color will be transparent. Images have to be 24bit, RGB bitmaps.

Reject after (applies to Workflow Templates only)

If ScannerVision encounters a problem with the processing of a document it will re-process the document the next time the template is scheduled for execution. The number of retries can be limited to the number specified here. If the retry limit is reached the document is moved to the rejected folder ("C:\ProgramData\ScannerVision\Rejected").

If you set this value to 0 ScannerVision will re-process the document forever or until it succeeds. value to 0 ScannerVision will re-process the document forever or until it succeeds.

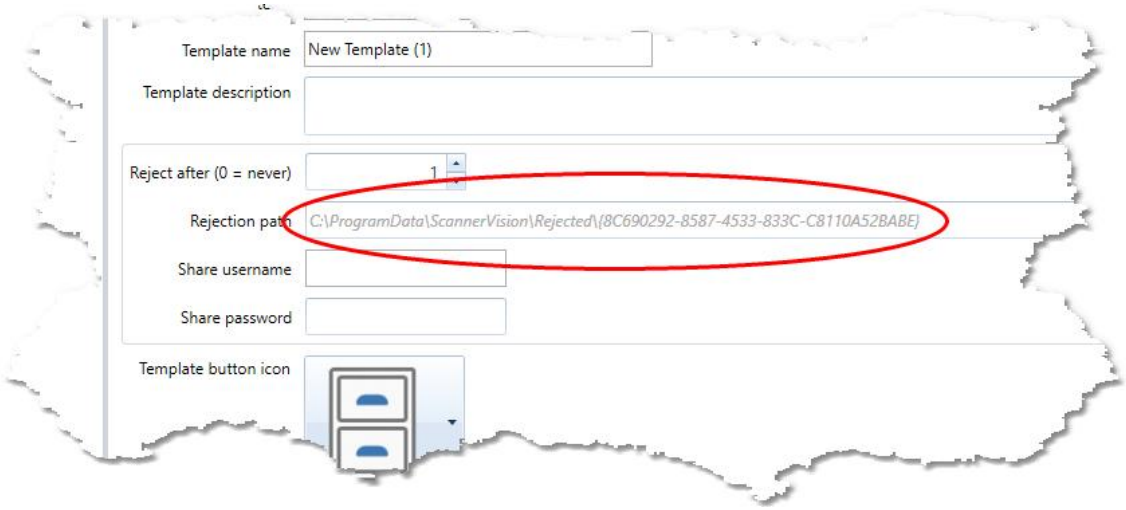
Rejection path

The output path of rejected documents. This setting will only be available when the "Reject after" value is greater than 0.

If you provide no rejection path the default path is shown as in the screen shot below.

The rejection path is published to the [REJECTEDFOLDERPATH] metadata tag. The name of the rejected document is published to the [REJECTEDFILENAME] tag and the fully qualified path to the rejected document is published to the [REJECTEDFILENAMEWITHPATH] tag.

Since the rejection path could contain metadata tags the resolving of the rejection path is only done when the actual rejection of the document happens at the very end of the processing work flow prior to the execution of the notifications and the template's end script. If the document is not rejected the metadata tags mentioned above will be undefined.



Share username/password

The credentials to connect to a protected network share. These settings will only be available when the "Reject after" value is greater than 1 and the rejection path is a network share.

PDF load resolution

The resolution of the bitmap to which PDF documents are rasterized. For more information refer to Appendix D - PDF Input Documents.

Template Schedule

Templates can be configured to run at an interval or on a fixed time every day.

Schedule type	Set the type of schedule. Options are "Interval" or "Time of day".
Time	Sets the interval between runs when the schedule type is set to "Interval" or the time at which the template should run when the schedule type is "Time of day".
No cut-off time	This option becomes available when the "Time of day" option is selected. When selected the template will continue to execute until the respective capture sources which are configured for the template do not contain any more documents to process.
Cut-off time	This option becomes available when the "Time of day" option is selected. This setting allows you to configure a window withing which the template will execute. The template will continue to process until the respective capture sources which are configured for the template do not contain any more documents to process or the cut off time is reached.

Start/End script

A Visual Basic or JScript script to be executed prior to and/or at the end of the template execution.

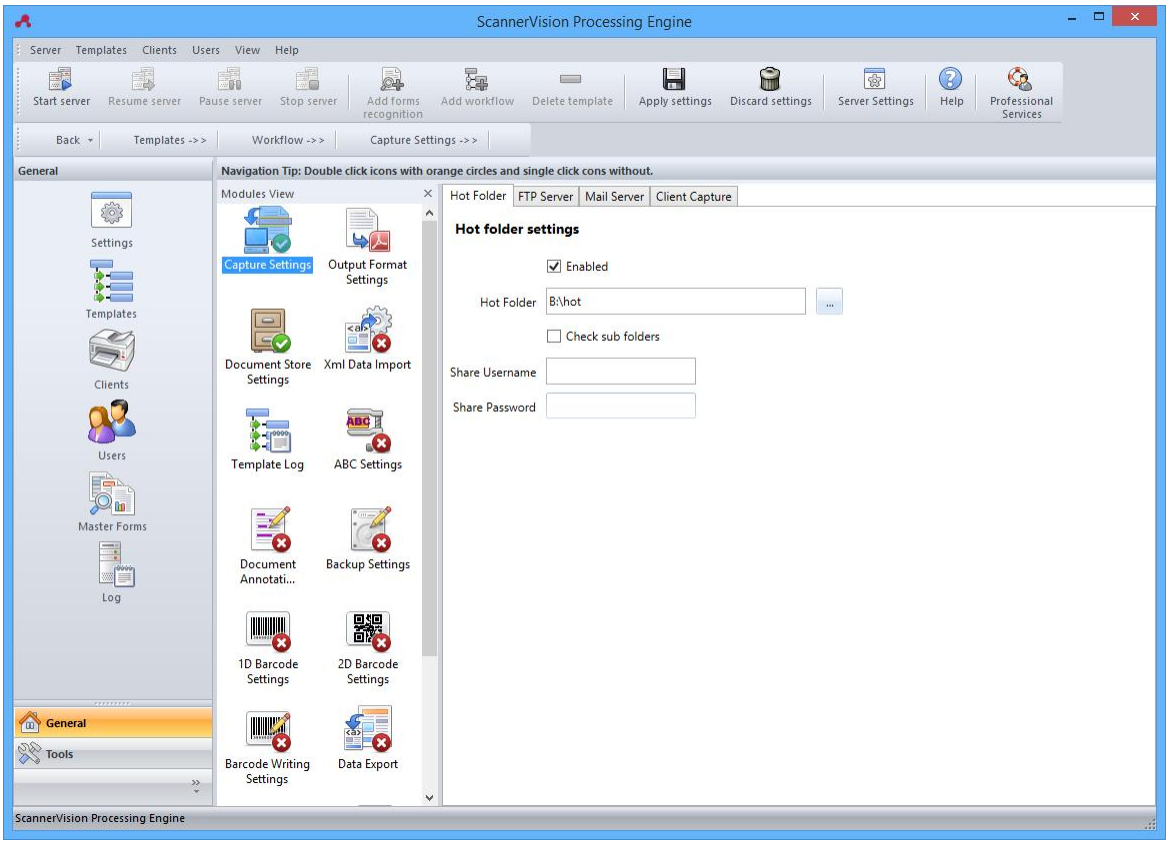
Bypass image processing

When this option is enabled no image processing is done. The document is passed straight on to the connectors which speeds up processing significantly.

8.4 Capture

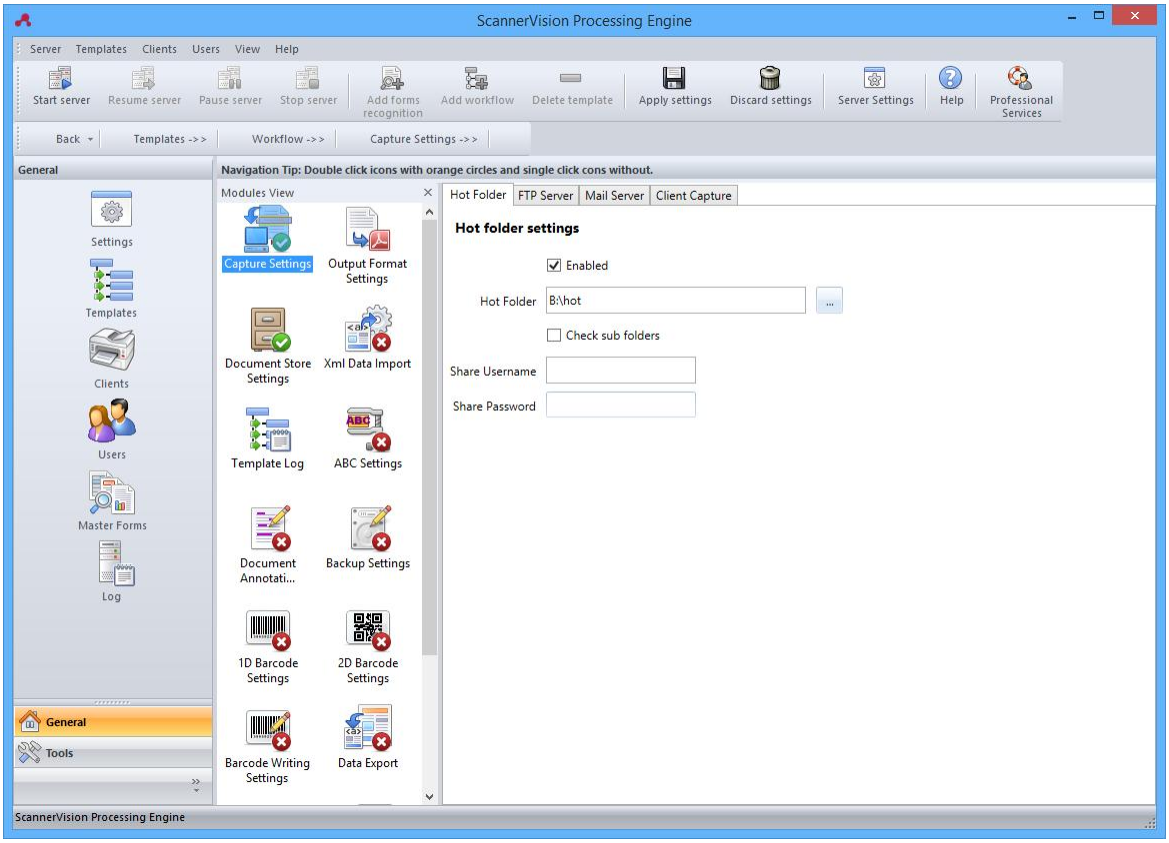
Template processing starts with the capturing of a document. ScannerVision supports four document capture sources which are discussed in the sections that follow. The source are:

- **Hot Folder**
- **FTP Server**
- **POP3 Server**
- **Client Capture**



8.4.1 Hot Folder

The Hot Folder capture source is a file system directory - local or network shared - into which documents are copied from any external source.



Active

Enables/disables hot folder capture.

Hot Folder

The fully qualified path of the hot folder. You can configure more than one hot folder by separating folder paths with a semi-colon " ; ".

Check sub-folders

Select this option if you want sub folders to be included in the monitoring for new documents.

Share Username/Password

Login credentials if the hot folder is a protected shared network paths.

8.4.2 FTP

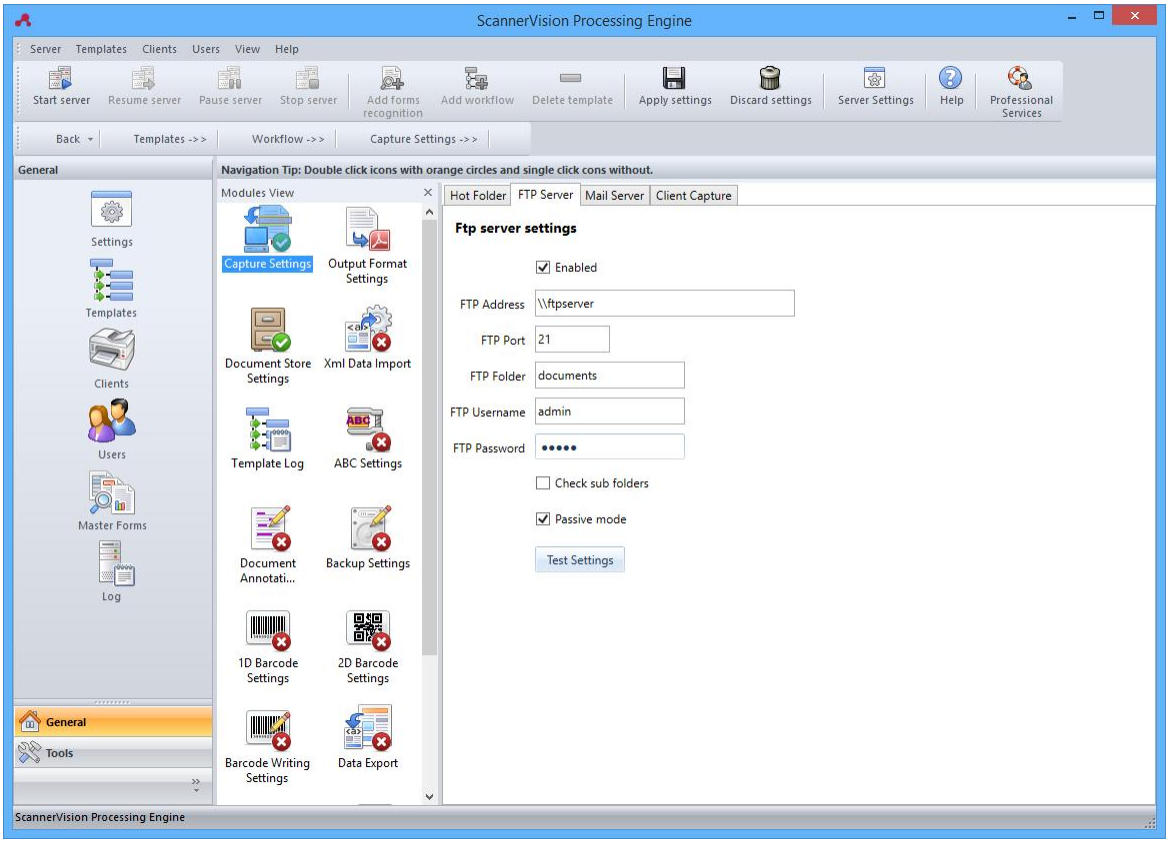
FTP capturing allows for the capturing of documents from any accessible ftp server. When documents are processed that have been captured from an ftp server the following metadata tags will be populated:

[FTPADDRESS]

[FTPPORT]

[FTPFOLDER]

[FTPFILENAME]



Active

Enables/disables ftp capture.

FTP Address

The ftp server address.

FTP Port

The port on which the ftp server listens.

FTP User/Password

Login credentials of the ftp server.

FTP Folder

The folder on the ftp server to capture from.

Check sub-folders

Select this option if you want sub folders to be included in the monitoring for new documents.

Passive mode

Select this check box to connect in passive mode. A discussion of passive mode is beyond the scope of this document. For more information please see Passive mode.

Test Settings

Press this button to verify that the ftp server can be contacted.

8.4.3 Email

Email capturing allows for the capturing of documents from IMAP and POP3 email servers. All email attachments that are of a supported document type are processed while unsupported attachments are discarded.

Attachments that are of a supported format are validated before they are processed and if any of them fail the validation for whatever reason the email is flagged (or starred) and marked as unread and no attachments are processed. If you want ScannerVision to retry to process the email you can clear the flag using an email client. All flagged/starred emails are ignored - even if they were not flagged/starred by ScannerVision.

Supported document formats are .tif, .tiff, .pdf, .bmp, .jpg, .gif, .pcx, .png, .doc, .docx and .rtf.

When documents are processed that have been captured from an email server the following metadata tags will be populated:

[EMAILFROM]

[EMAILTO]

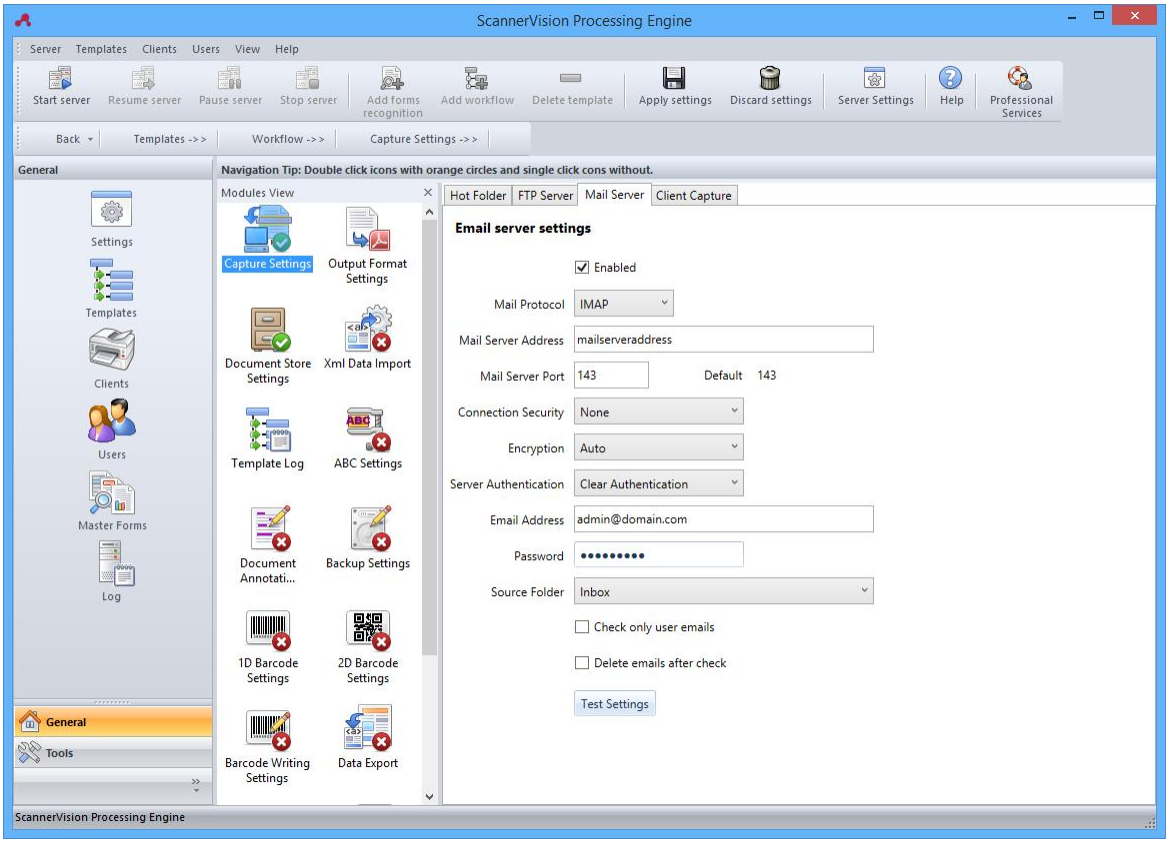
[EMAILCC]

[EMAILBCC]

[EMAILSUBJECT]

[EMAILBODY]

[EMAILFILENAME]



Active

Enables/disables email capture.

Mail Protocol

Specifies the type of email server you are capturing from. Options are IMAP and POP3.

Mail Server Address

The email server address.

Email Server Port

The port of the mail server to which ScannerVision connects. The port changes to the default for the particular type of security protocol (shown to the right of the "Default" caption) that is selected with the "Connection Security" setting. You can modify the port regardless of the connection security choice you have made.

Connection Security

Specifies whether communication with the email server should be encrypted. The choice made here applies to all communication other than authentication which is determined by the "Server Authentication" option.

None	No encryption. All communication is unencrypted.
STARTTLS	Communication is initially unencrypted to allow for connection to regular POP3 or IMAP port. Encryption is enabled automatically prior to authentication.
SSL/TLS	All communication is encrypted from the start and connection must be done on a dedicated SSL/TLS port (usually 995 for POP3 and 993 for IMAP).

Encryption

When "Connection Security" is set to STARTTLS or SSL/TLS this option specifies the encryption protocol that is used when communicating with the server.

Auto	The strongest supported encryption algorithm supported by the mail server is used. If the mail server does not support encryption communication will fail. The "Auto" option does not allow no encryption. If your email server does not support encryption please choose "None" under "Connection Security".
SSL 2	Secure Socket Layer V2
SSL 3	Secure Socket Layer V3
TLS	Transport Layer Security

Server Authentication

The authentication method to use to connect to the email server. "POP3 Clear Authentication" and "APOP Authentication" are supported.

Clear Authentication	Passwords are transmitted as clear text.
Secure Authentication	The strongest supported encryption algorithm supported by the mail server is used. If the mail server does not support encryption communication will fail. If you email server does not support encryption please choose "Clear Authentication".
APOP Authentication	Secure APOP authentication. Only supported on POP3 servers.

Email address

The email account which should be monitored.

Password

The password to connect to the email server.

Source Folder (IMAP only)

The IAMP folder which should be monitored.

Check only user emails

Select this option to process only mails received from Users.

Delete emails after check

Select this option if you email messages to be deleted after they have been processed.

Note

Certain email servers don't physically delete emails that have been marked for deletion for as long as a client is connected to the particular mailbox. When all clients have disconnected, emails that were marked for deletion are then physically deleted. ScannerVision maintains a list of email IDs of deleted emails so that emails that have been marked for deletion are not processed again.

Test Settings

Press this button to verify that the POP 3 server can be contacted.

8.4.4 Client Capture

Client capture allows for the capturing of documents from desktop or MFP Clients. When documents are processed that have been captured from a ScannerVision client the following metadata tags will be populated:

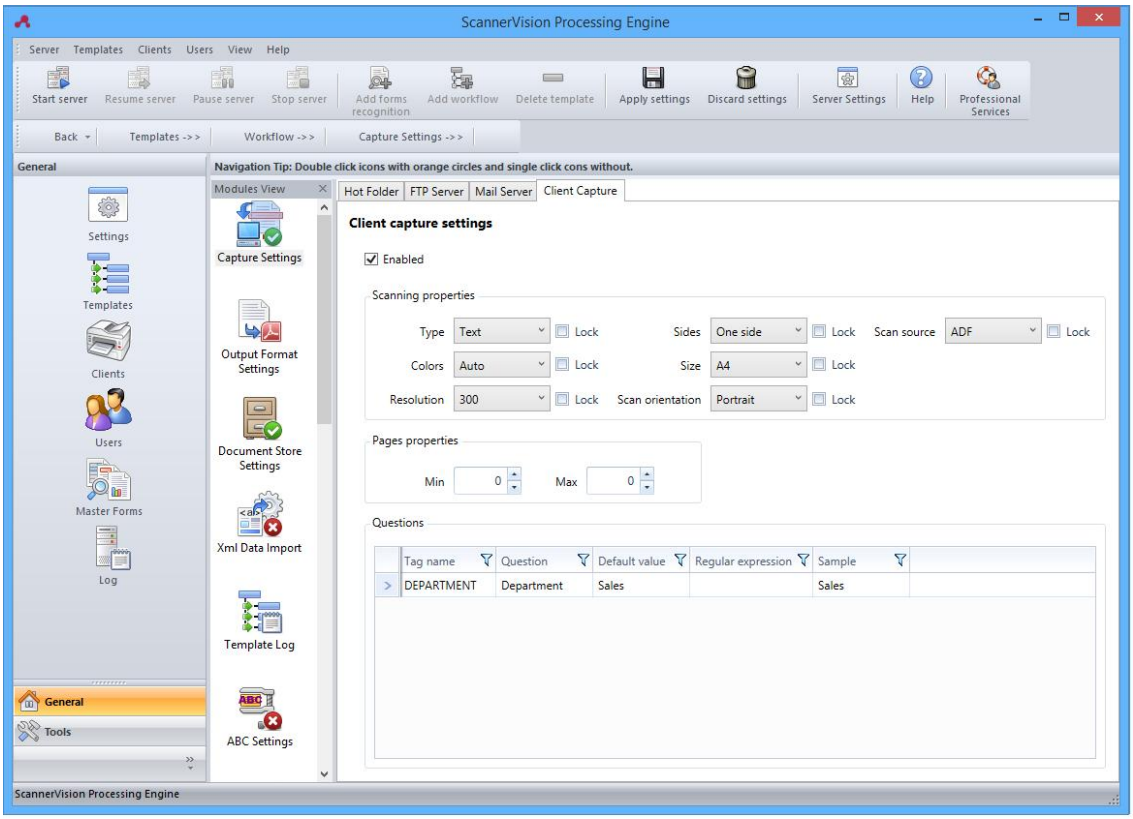
[USERNAME]

[USERMAIL]

[USERHOMEFOLDER]

[CLIENTIPADDRESS]

[MFDMODELNR]



Active

Enables/disables client capture.

Scanning properties

Scan properties configure the scan properties of the MFP device or TWAIN source in the case of a desktop client. The properties defined here can be locked by selecting the "Lock" check box after each property. This will prevent the user from modifying the property on the client. These options do not apply when existing documents are loaded into the Desktop Client.

Type	The type of document that is scanned. Options are text, photo and text and photo.
Color	Color option of the scan. Options are Auto, B&W (black and white), Greyscale and Colors.
Resolution	The scan resolution. Options are 100dpi, 200dpi, 300dpi, 400dpi and 600dpi.
Sides	Select if images are scanned on one side only or both sides. Options are One side and Duplex (both sides).
Size	The size of the image to be scanned, options are Auto, A3, A4, A5, B4, B5, Letter, Legal, Executive or Folio.
Orientation	The orientation of the pages to be scanned. Options are Portrait and Landscape.
Scan source	The source of the document to scan. Options are ADF (Automatic Document Feeder) and Glass.
Lock	Prevents users from changing settings. If the box is selected, users can see which setting is set for the process, but cannot change the scan setting.

Pages Properties

The minimum and the maximum number of pages that a document is allowed to have. If the number of pages of a document are out of bounds it will not be submitted to the ScannerVision Networking Server. If the minimum and maximum number of pages are the same value (larger than zero) the document must contain exactly that number of pages. To allow any number of pages, select zero for both the minimum and maximum value.

Questions

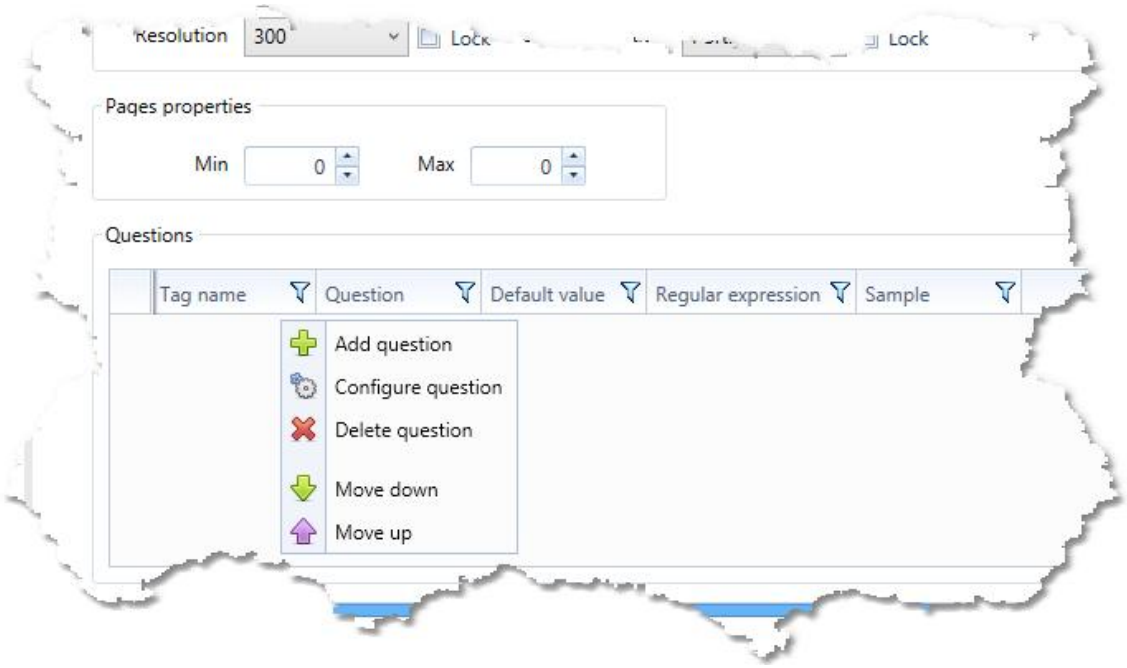
Please refer to the Capturing Document Metadata section for details.

8.4.4.1 Capturing Document Metadata

Document metadata is information that describes a document or enriches the information about a document. Metadata is captured through a set of questions or prompts that are presented to a user once a template has been selected on one of the ScannerVision clients (Desktop or Embedded MFP

clients). The metadata that is captured this way accompanies the document through the whole ScannerVision processing work flow and it can be used in various ways. It could for example be used to determine the final output destination of the document or it could be used as a key into a database table from where additional information could be obtained. The metadata can also be stored along with the scanned document in a document management system such as SharePoint or Laserfiche.

For an in depth discussion of metadata please refer to Appendix A - Metadata.



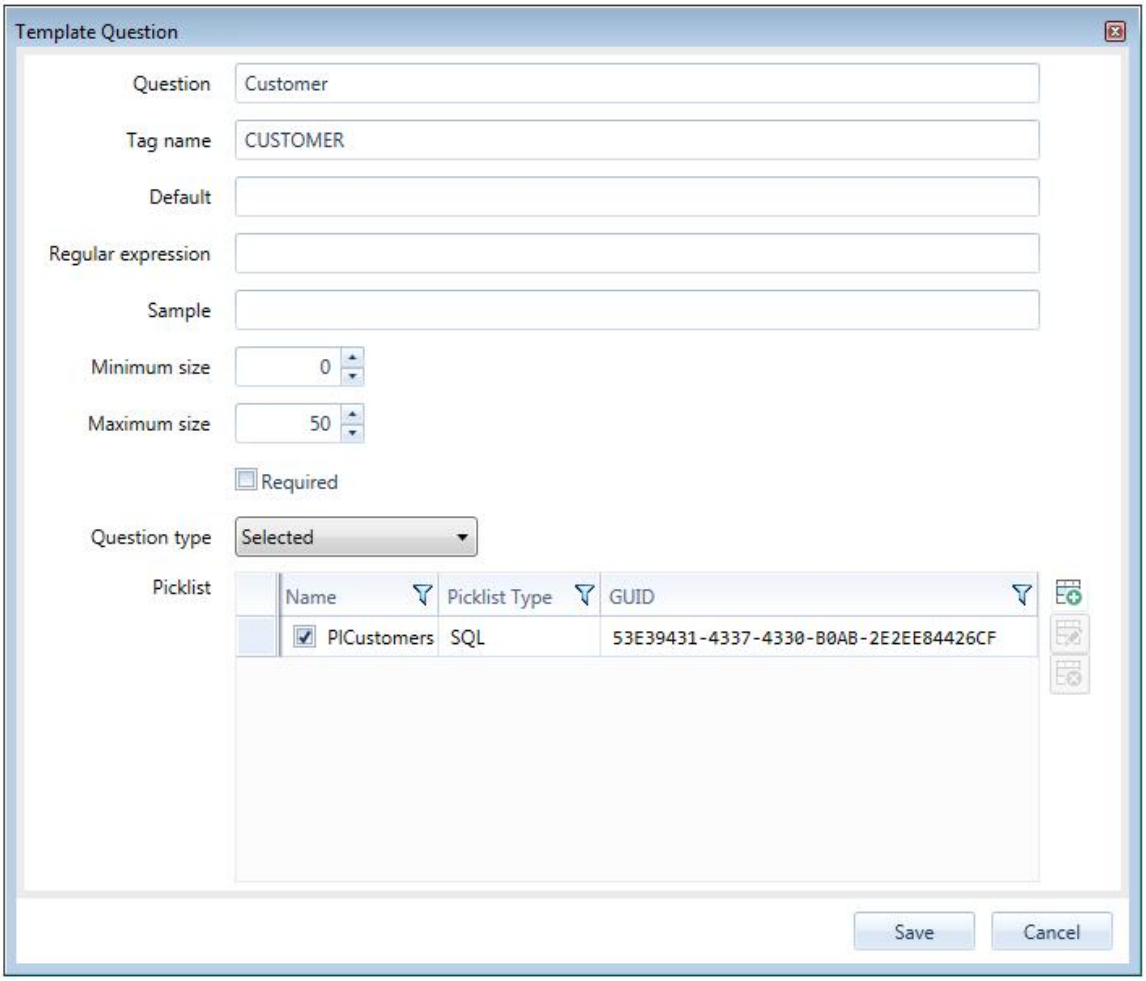
Context menu

When right clicking on the Questions grid the context menu appears with the following commands:

Add question	Adds a metadata question to a template.
Configure question	Configures an existing question.
Delete question	Delete a question.
Move down	Moves the selected question down.*
Move up	Moves the selected question up.*

*The order in which questions appear in the grid is the order in which they are displayed on the client.

The ScannerVision - Template Question dialog is shown:



Question (Required)

The prompt or question that is displayed to the user on the client.

Tag name (Required)

The metadata tag name by which this information is identified in the template. This value will appear in the ScannerVision Expression Editor list under the "Template tags" heading.

Default (Optional)

The default answer to the question/prompt if the user does not provide one. The user will see this

value in the respective edit field on the client.

Regular expression (Optional)

The data that is entered by the user can be validated by specifying a Regular Expression. The client application will try to find a match for the provided regular expression in the data that is entered by the user. If no match is found the user is presented with a message that states that data that was entered is incorrect. A sample of the correct data is also shown if you provide such in the Sample edit box.

By providing a regular expression, the quality and/or correctness of the captured information can be improved which if left unverified could potentially lead to problems later during the processing of the document.

Sample (Optional)

If a regular expression is provided and the user enters information that does not conform to the regular expression the user is prompted with a message to state that the entered data is not valid. The sample value you specify here is included in the message to show the user what the expected format of the data is.

Minimum size (Required)

The minimum number of characters that must be entered by the user.

Maximum size (Required)

The maximum number of characters that can be entered by the user.

Required (Optional)

Indicates that the question is required to be filled in by the user.

Question type (Required)

The selection you make here determines what options the user has for entering information on the client. The different options are:

Typed	The user is required to manually type in a value.
Typed & Selected	The user has the option to either type in a value or to select a value from a drop down list.
Selected	The user is required to select a value from a drop down list.
Browse	No further information is required. The particular folder which content is served is determined by the "Browse folder" value of the

	Add/Edit User window (see Users). Alternatively, if authentication is not enabled for a client or client group the "Default" value of the question is used.
Browse Picklist	This functionality is implemented by certain connectors such as SharePoint and configuration is specific to each implementation.

When the "Typed" or "Browse" option above is selected the "Picklists" section is not visible.

Picklists (Required)

The "Picklists" section appears when the "Question type" is either "Typed & Selected" or "Selected". Please refer to the Creating Picklists section for details on how to create picklists.

8.4.4.1.1 Creating Picklists

When you want a user to select a value from a drop down list box on the client instead of typing in a value you have to configure a picklist. The values of the picklist can be obtained in the following ways:

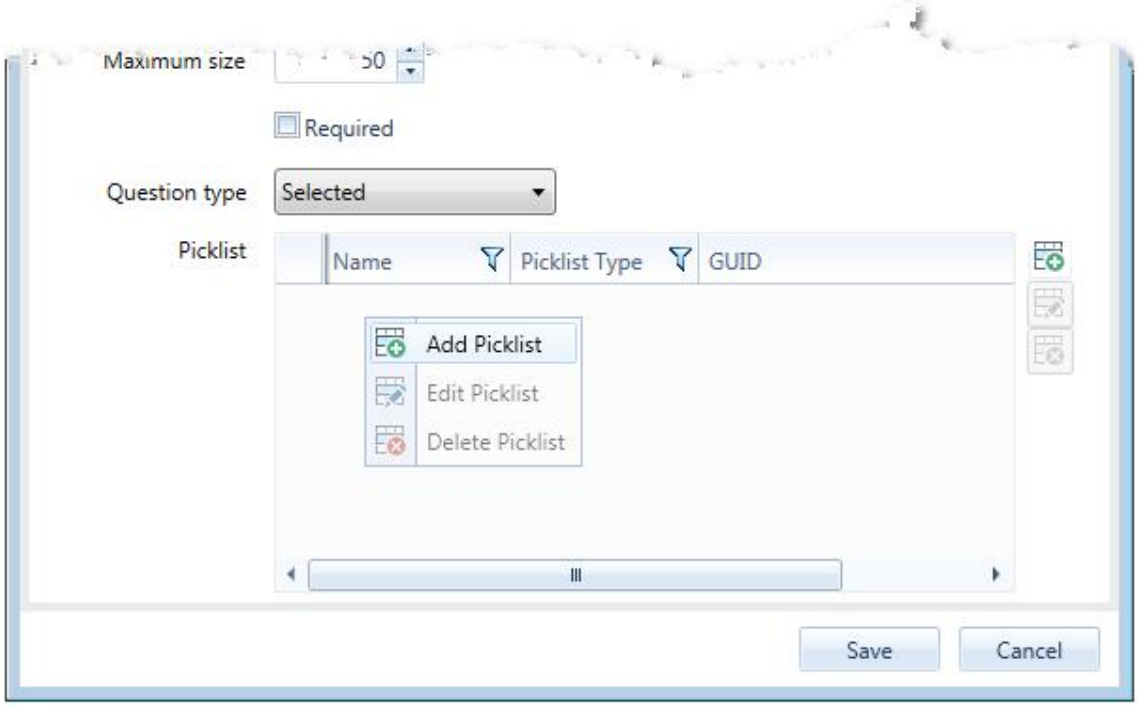
1. Static list
2. Database query
3. Visual Basic script
4. XML file

Certain connectors such as SharePoint publishes picklists when installed. So, depending on which connectors you have installed, you may have additional choices when configuring picklists.

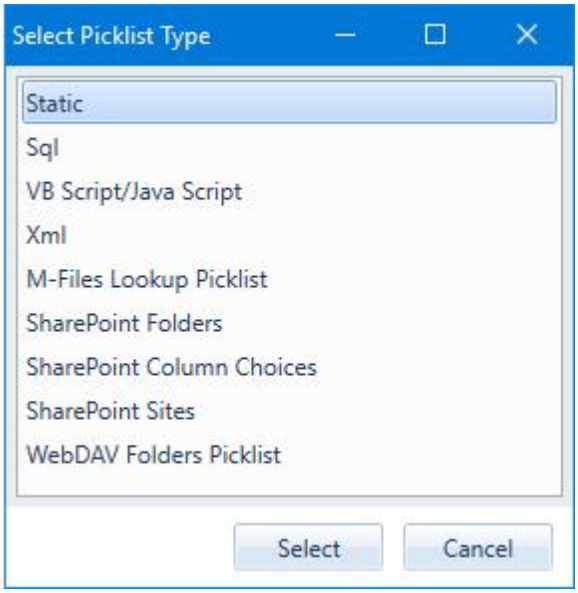
Picklists don't belong to a specific template or global metadata question even though they could be created from within either context. Once a picklist has been created it will appear in the list of

available picklists of all templates and global metadata questions.

To add a picklist to a template or global metadata question, right click in the "Picklist" grid and select the "Add Picklist" option or press the button to the right of the picklist grid:

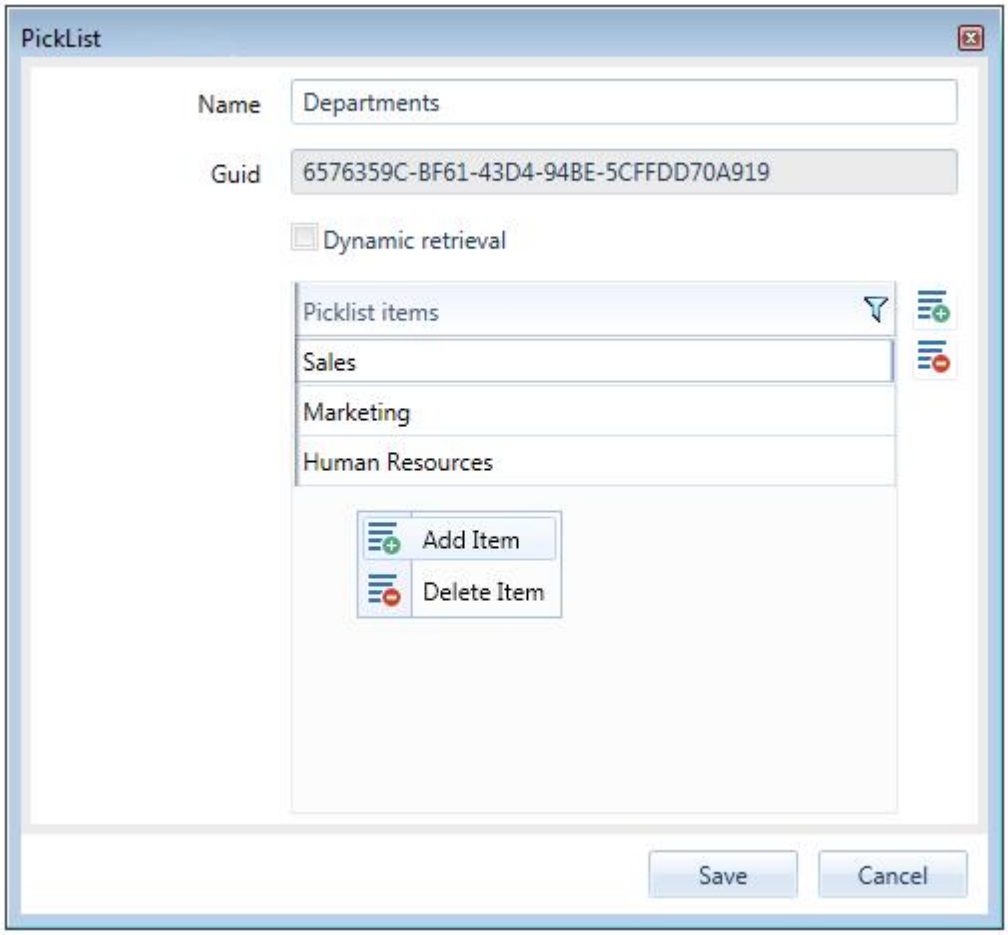


A window will appear with a list of all the picklist types you can configure. The list of available picklists you see may differ from the list shown below depending on which connectors you have installed. Make a selection here and click the "Select" button.



To edit or delete an existing picklist right click on the row and choose the desired option:

The Static picklist editor is shown below:



The upper part of the editor is the same for all picklist types. When the picklist type is changed the lower part of the editor changes according to the type of picklist.

Name (Required)

The name by which the picklist is identified.

Guid

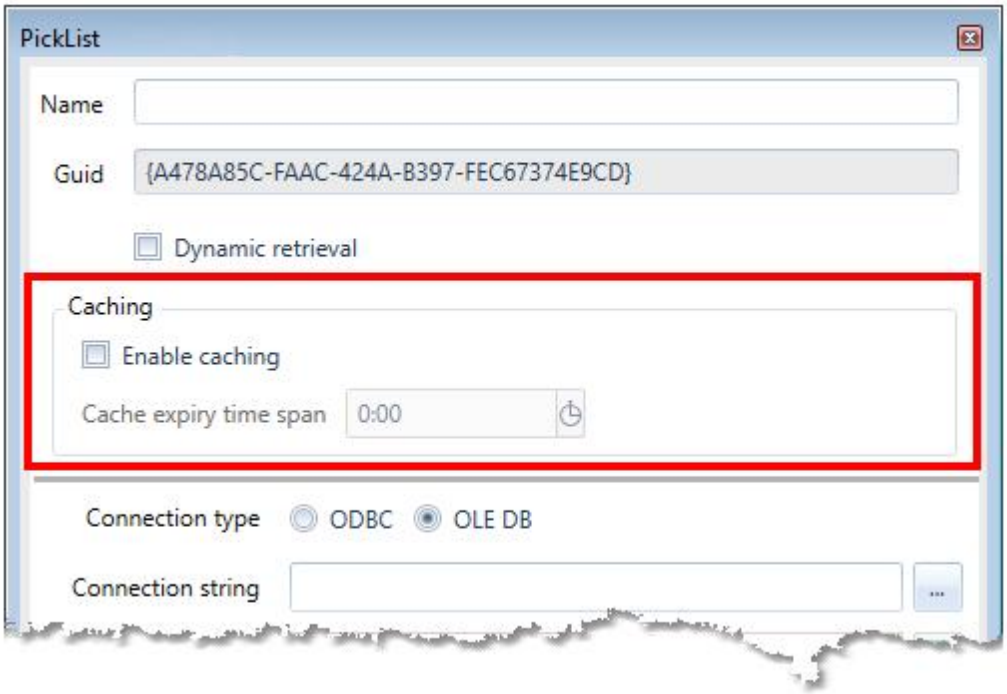
The auto generated, read-only ID of the picklist.

Dynamic retrieval (Optional)

The Dynamic retrieval check box indicates that the picklist makes use of metadata that has been captured by previous questions. Please refer to the Dynamic Picklists section for further details.

Caching*

Picklist values can be cached by the ScannerVision Processing Server High Performance Networking Server to improve performance. Caching is only available for non-static picklists.



Enable caching

Enables/disables caching for a picklist.

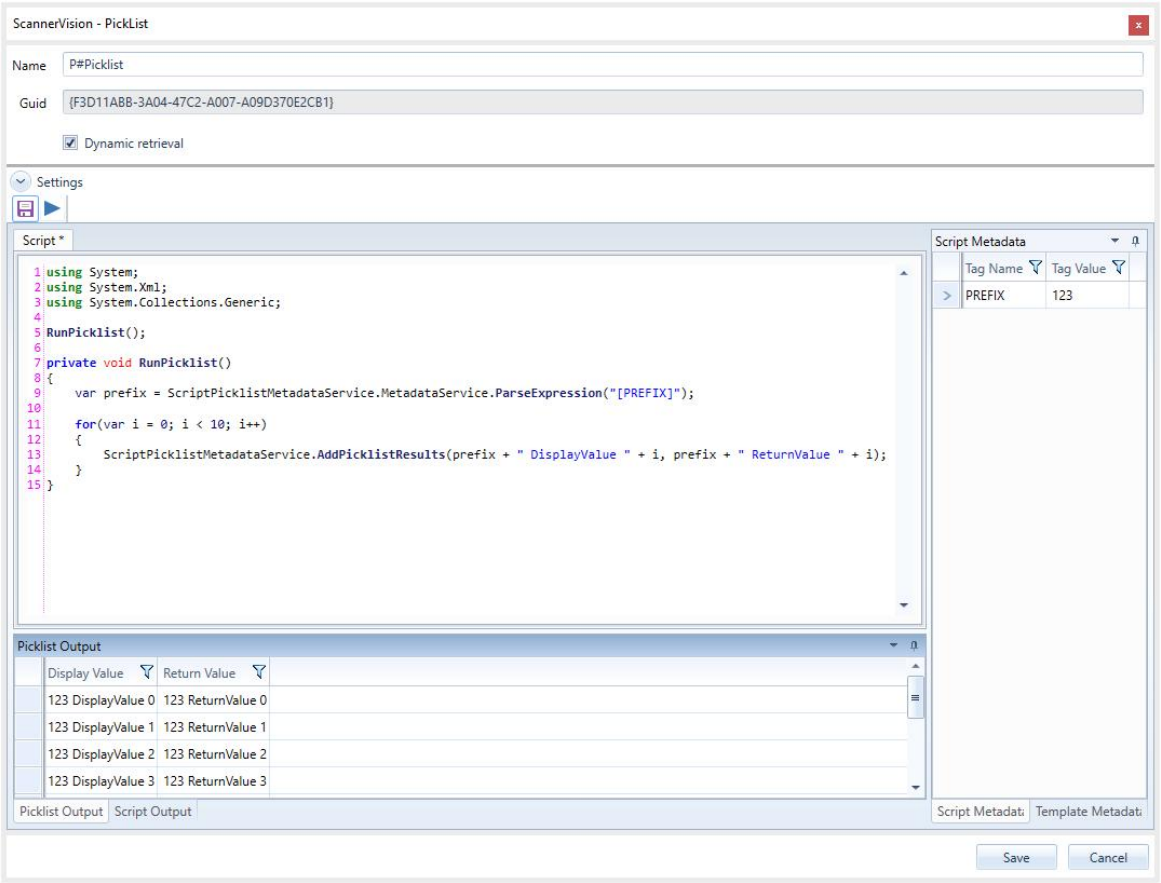
Cache expiry time span

The duration in hours and seconds that cached values should be kept. After the specified time span has lapsed the values in the cache are cleared and new ones fetched on the next call by a client to the picklist.

*ScannerVision Processing Server High Performance Networking Server is a stand-alone installation that replaces the built-in ScannerVision Networking Server. With HPN installed ScannerVision Processing Server can theoretically service thousands of clients depending on network throughput.

8.4.4.1.1.1 PowerSharp

A PowerSharp picklist obtains its values by executing a C# script. The editor window is shown below:



Please refer to the PowerSharp Connector for instructions to construct and execute a C# script.

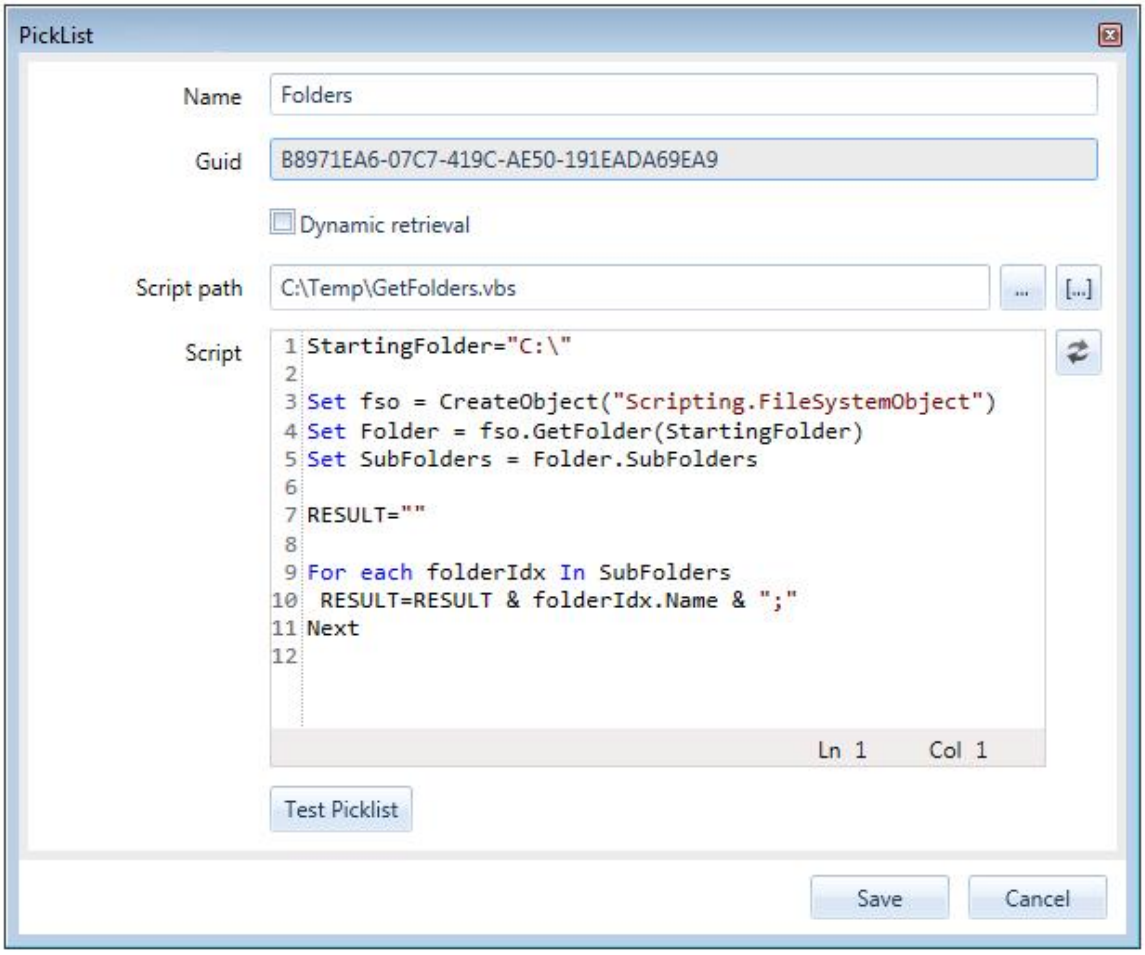
The context in which a picklist script executes is a bit different from that of a connector in the following ways:

1. Only metadata tags of previously answered template questions can be referenced.
2. No metadata tags can be published.
3. A picklist script is supposed to return a list of values by adding display/return value pairs with a call to the ScriptPicklistMetadataService.AddPicklistResults function.

In the example script above the [PREFIX] tag is populated by a previously answered template question. The value of this tag is parsed and then appended to the front of the list of values that is displayed to the user.

8.4.4.1.1.2 Script

A Script picklist obtains its values by executing a Visual Basic or JScript script. The editor window is shown below:



Script path (Required)

The path to the script file to execute. You can select a script file by pressing the "..." button to the right of the Script path edit box. The path may contain ScannerVision metadata tags. To edit the path in the ScannerVision Expression Editor press the "[...]" to the right of the Script path edit box.

Script

If you have selected a Script file it is loaded automatically into the script editor when the picklist editor is opened. If however the Script path contains ScannerVision metadata tags the script cannot be loaded automatically. To load the script press the  button to the right of the script editor. You will be presented with the "Metadata Tag Values" dialog discussed in the "Test Picklist" section below. If you provide values which, when parsed, resolves to an existing file on the system the file is

loaded into the script editor.

If the Script path does not contain any ScannerVision metadata tags, any changes you make to the script in the script editor is saved to the file when the "Save" button is pressed.

The values you want displayed on the client should be returned as a semicolon delimited list. For example, if you want to return a list of folders on the server you could write the following script:

```
StartingFolder="C:\Storage\"

Set fso = CreateObject("Scripting.FileSystemObject")

Set Folder = fso.GetFolder(StartingFolder)

Set SubFolders = Folder.SubFolders
RESULT = ""
For each folderIdx In SubFolders

  RESULT = RESULT & folderIdx.Name & ";"

Next
```

Metadata

ScannerVision metadata can be referenced in your script through the "Values" method of the "Metadata" object. The Values method takes a string parameter that represents the metadata tag you want to reference. To reference the [DATETIME] tag you would write the following:

```
Metadata.Values("DATETIME")
```

Value displayed & value returned

The Script picklist editor does not offer the ability to select a value to display and a value to return as SQL and XML picklists do but you can still achieve this result. To do this you have to return the value to display and the value to return as name-value pairs separated by ASCII character 30. The value to display is first then ASCII character 30 and then the value to return. Name-value pairs are separated by a semicolon e.g.

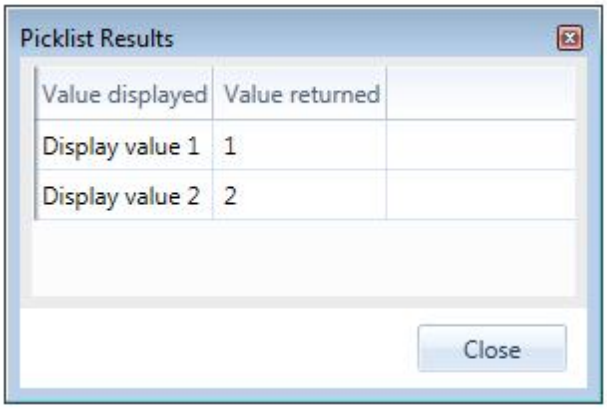
```
RESULT = "Display value 1" & Chr(30) & "1;" & "Display value 2" & Chr(30) & "2"
```

With this script the values "Display value 1" and "Display value 2" are show to the user, but the values "1" or "2" respectively are returned to ScannerVision.

Please refer to the Value displayed and value returned section for an explanation of the significance of the "Column displayed" and the "Column returned".

Test Picklist

To test that your picklist is configured correctly press the "Test Picklist" button. You will be presented with the "Picklist Results" dialog shown below:



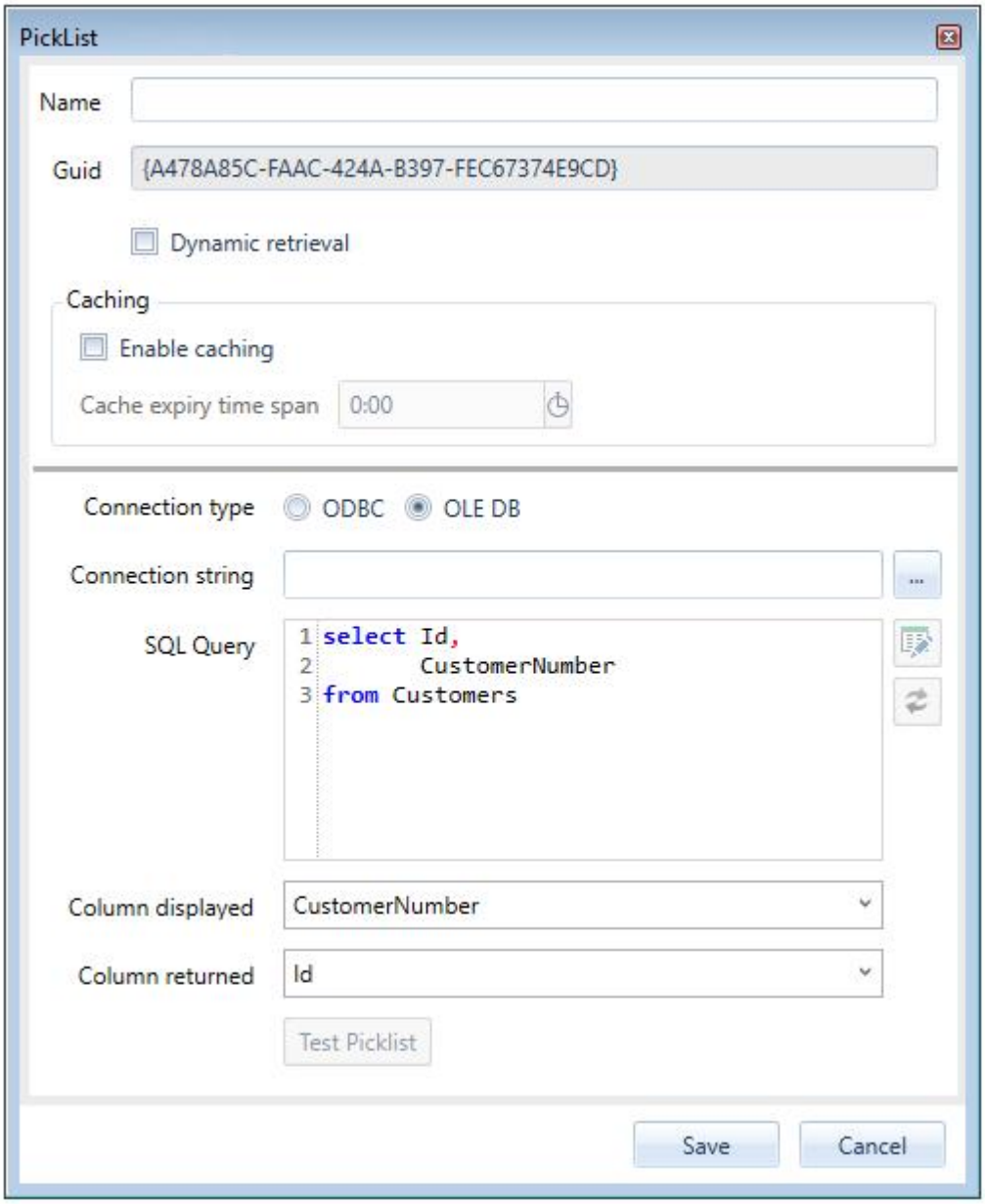
The dialog will show a maximum of 3 rows. Verify that the "Value displayed" and the "Value returned" values are what you expected.

Note

When testing your picklist only Standard ScannerVision metadata tags will be parsed when the script executes. If you reference a metadata tag in your script that does not exist in the context of the picklist editor, the script will not return any results. For example if you are configuring a picklist from within the context of a Global Metadata question, no metadata tags defined in templates will be parsed.

8.4.4.1.1.3 SQL

A SQL picklist obtains its values from a database. The editor window is shown below:

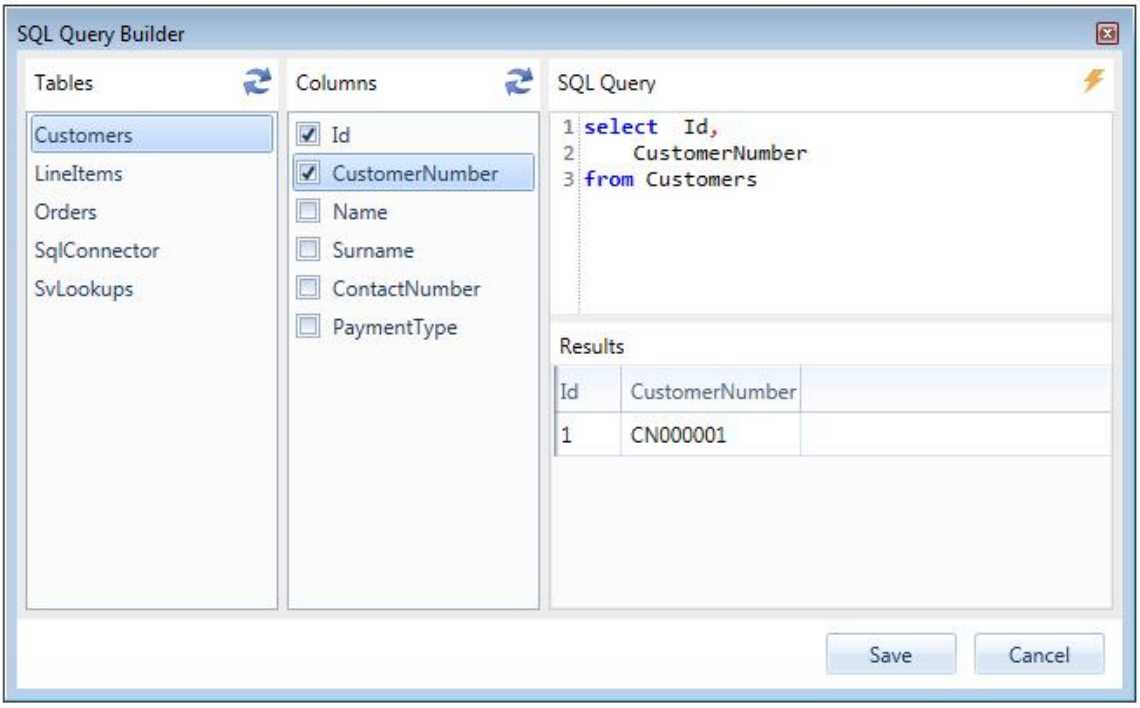


Connection string (Required)

The connection string contains the information necessary to connect to a database engine. Refer to Appendix C - Database Connection Strings for more information.

SQL Query (Required)

The SQL query is what gets executed by ScannerVision to obtain picklist values. You can be as specific or elaborate as you need to be with the query you specify. Any legal SQL query is allowed but try to observe common SQL query best practice. For example, don't use "select *" when you only want 1 or 2 columns. It puts unnecessary load on ScannerVision and the network as result sets get large. To this end we encourage you to make use of the SQL Query Builder which you can launch by pressing the  button to the right of the SQL editor. The SQL Query Builder window is shown below:



Select the table from which you want do the look-up in the "Tables" column. Once you've selected the table, the "Columns" list is updated with the columns in the table. To refresh the list of tables and columns press the respective button at the top of the list.

In the columns list select the columns you want to return from the table. The SQL query editor updates automatically with the selected table and columns.

To test the query press the button in the top right hand corner of the window. If the query execution is successful the first 3 records of the result set are shown in the "Results" grid.

Note

The SQL Query Builder is a one-way tool. If you launch the SQL Query Builder with an existing SQL query, the table and columns referenced in the query will not be selected automatically when the SQL Query Builder window opens. You will have to make your selections again to modify your query.

Column displayed/returned (Required)

After you have entered or changed a query you have to update the list of columns returned by the

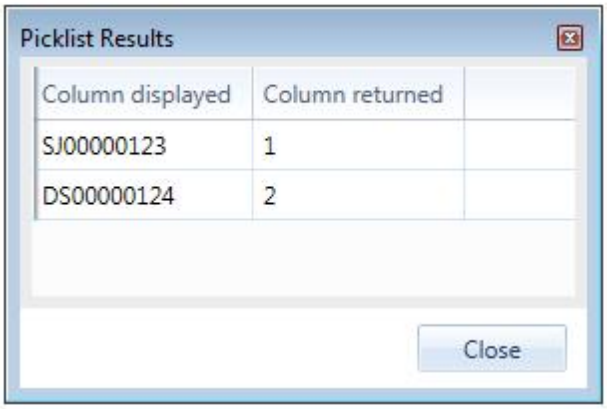
query. Do this by clicking the  button to the right of the SQL editor.

Select the column which you want displayed to the user and which column's value must be returned to ScannerVision as the selected value. You can enter the column names manually or you can select the column from the respective drop down list box. The list box is populated by executing the query against the database and extracting the column names from the result set.

Please refer to the Value displayed and value returned section for an explanation of the significance of the "Column displayed" and the "Column returned".

Test Picklist

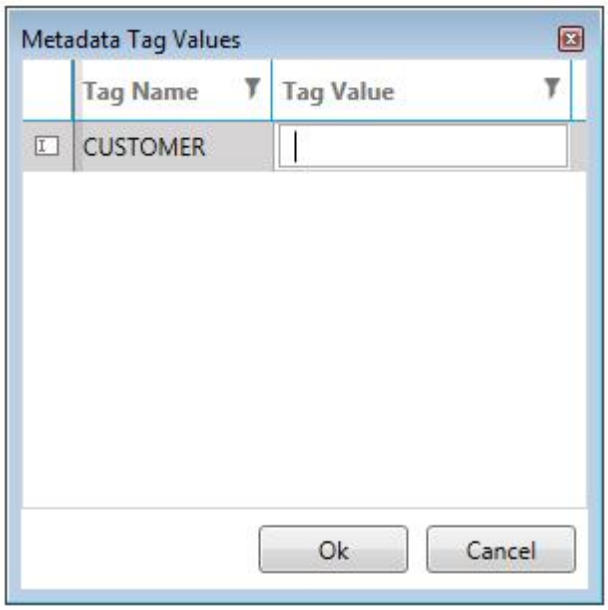
To test that your picklist is configured correctly press the "Test Picklist" button. You will be presented with the "Picklist Results" dialog shown below:



The dialog will show a maximum of 3 rows. Verify that the "Column displayed" and the "Column returned" values are what you expected.

Providing temporary metadata tags

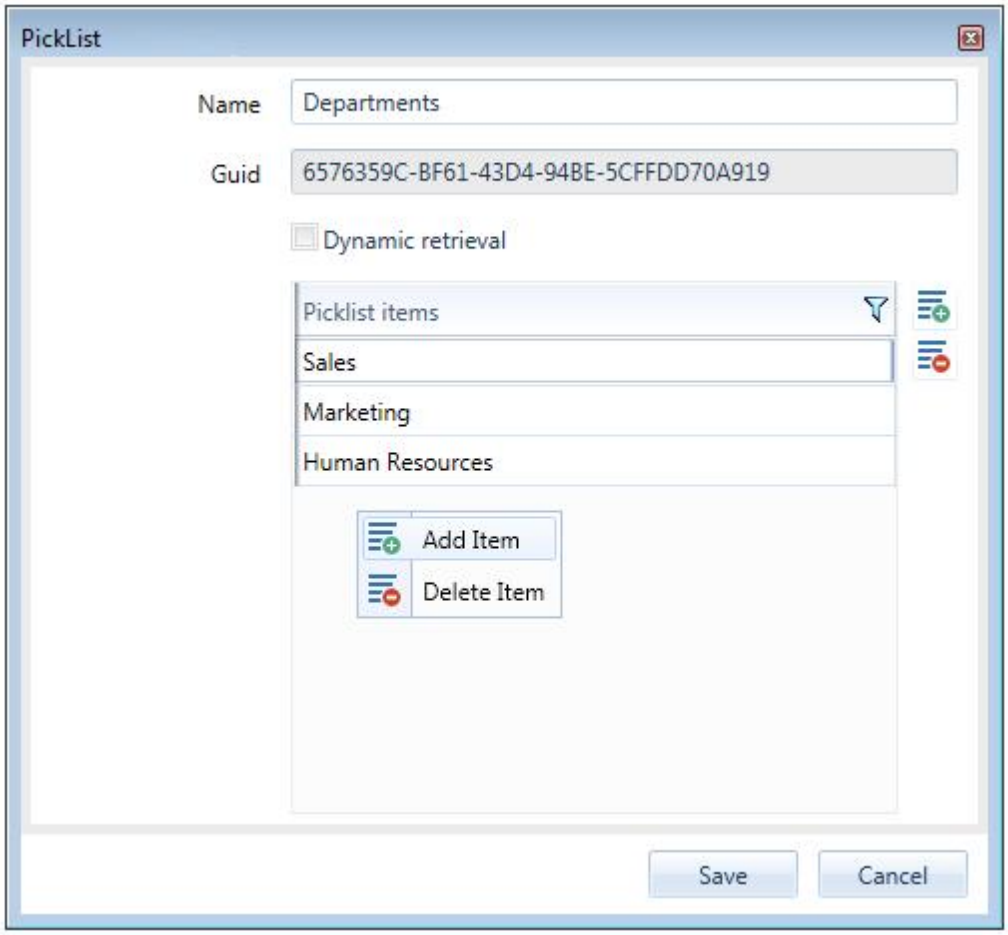
If you have referenced ScannerVision metadata tags in your query you will be prompted to provide temporary values for the metadata tags you've used when updating the columns or testing the picklist:



The values you specify here are used to by ScannerVision to substitute for the used metadata tags. If you enter legitimate values then the query should execute successfully.

8.4.4.1.1.4 Static List

A static picklist is one that has a fixed list of values that don't change often or that are not dependent on the answers of previous questions. This is not to say that the list can never change but that a change to the list requires the manual addition to or removal of items from the list. The static picklist editor window is shown below:



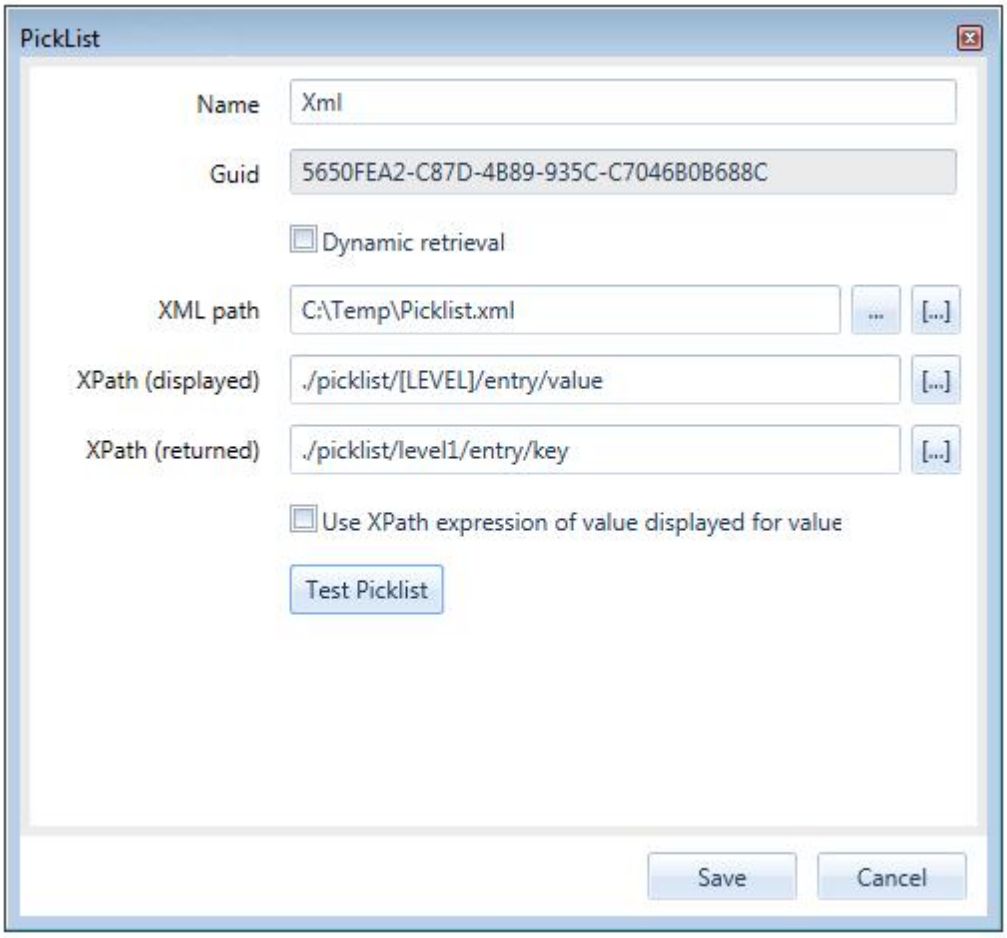
To add a list item right click in the "Picklist items" grid and select the "Add Item" menu option or press the button to the right of the "Picklist items" grid. A new item with the value "New Item" is added to the list.

To edit the value of a list item, select the item in the grid and press the "F2" key or click on the selected item again.

To delete an item right click on it and select the "Delete Item" menu option or press on the button.

8.4.4.1.1.5 XML

An XML picklist obtains its values from an Xml file. The editor window is shown below:



XML path (Required)

The path to the XML file. You can select a file by pressing the "..." button to the right of the XML path edit box. The path may contain ScannerVision metadata tags. To edit the path in the ScannerVision Expression Editor press the "[...]" to the right of the XML path edit box.

XPath displayed/returned (Required)

XPath is a query language for selecting nodes from an Xml document. An in depth discussion of the XPath language is beyond the scope of this document. Xml picklists allow you specify separate XPath expressions for the value displayed and the value returned. The XPath expressions may contain ScannerVision metadata tags. To edit the XPath expressions in the ScannerVision Expression Editor press the "[...]" to the right of the respective XPath edit boxes.

Please refer to the Value displayed and value returned section for an explanation of the significance

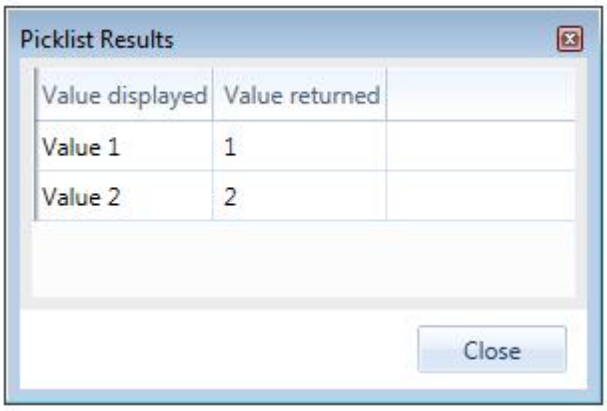
of the "Column displayed" and the "Column returned".

Use XPath expression of value displayed for value returned (Optional)

To use the same XPath expression for both the value displayed and the value returned, select the "Use XPath expression of value displayed for value returned" check box.

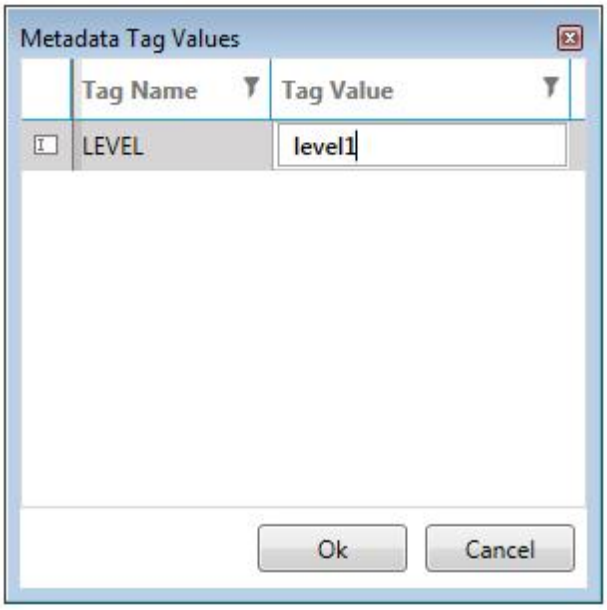
Test Picklist

To test that your picklist is configured correctly press the "Test Picklist" button. You will be presented with the "Picklist Results" dialog shown below:



The dialog will show a maximum of 3 rows. Verify that the "Value displayed" and the "Value returned" values are what you expected.

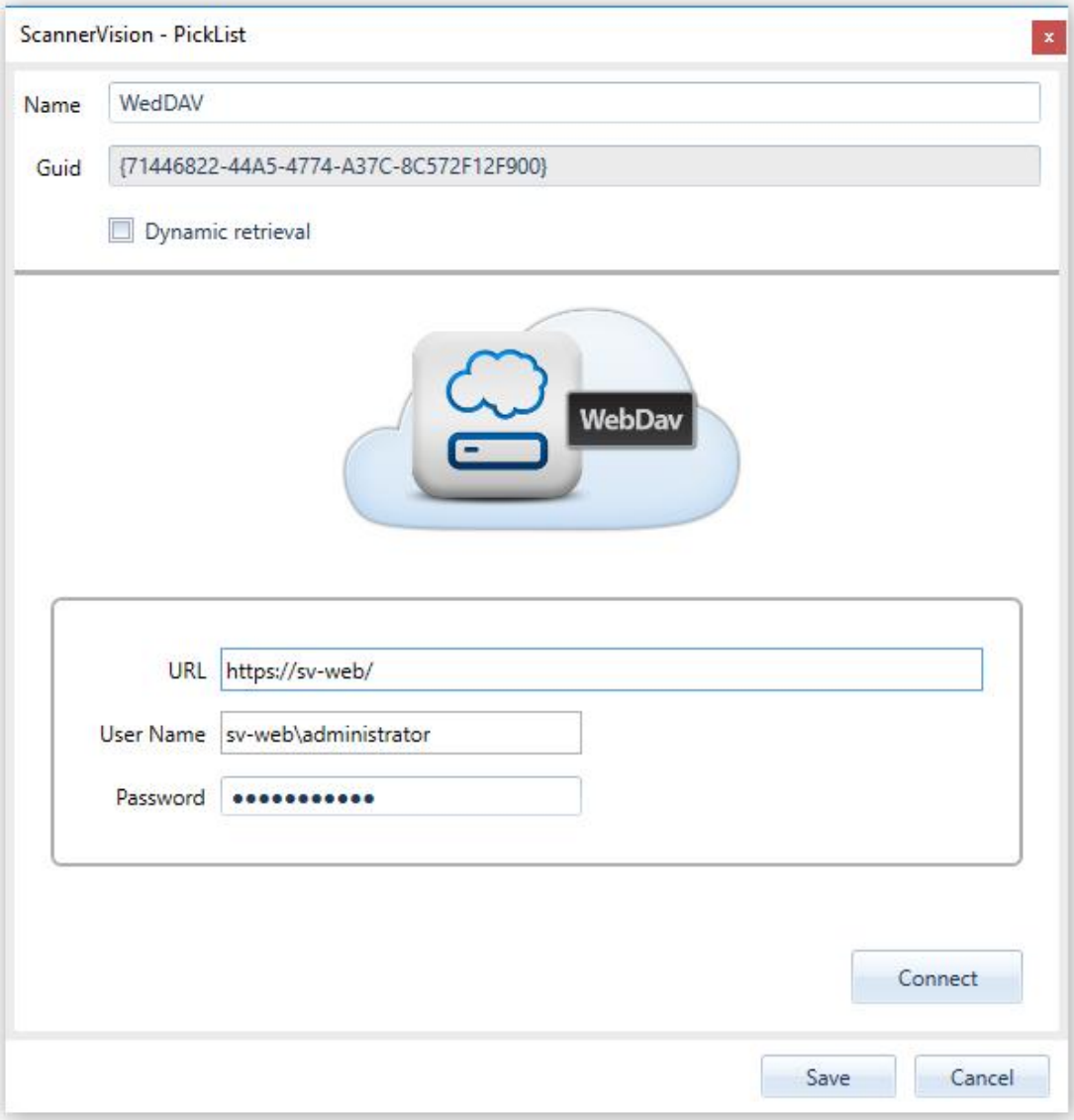
If you have referenced ScannerVision metadata tags in your query you will be prompted to provide temporary values for the metadata tags you've used:



The values you specify here are used to by ScannerVision to substitute for the used metadata tags. If you enter legitimate values then the XPath expression should execute successfully.

8.4.4.1.1.6 WebDAV

A WebDAV picklist presents the folders of the server connect to. This picklist requires no configuration apart from the server URL and credentials you want to connect to. The login screen is shown below:



URL

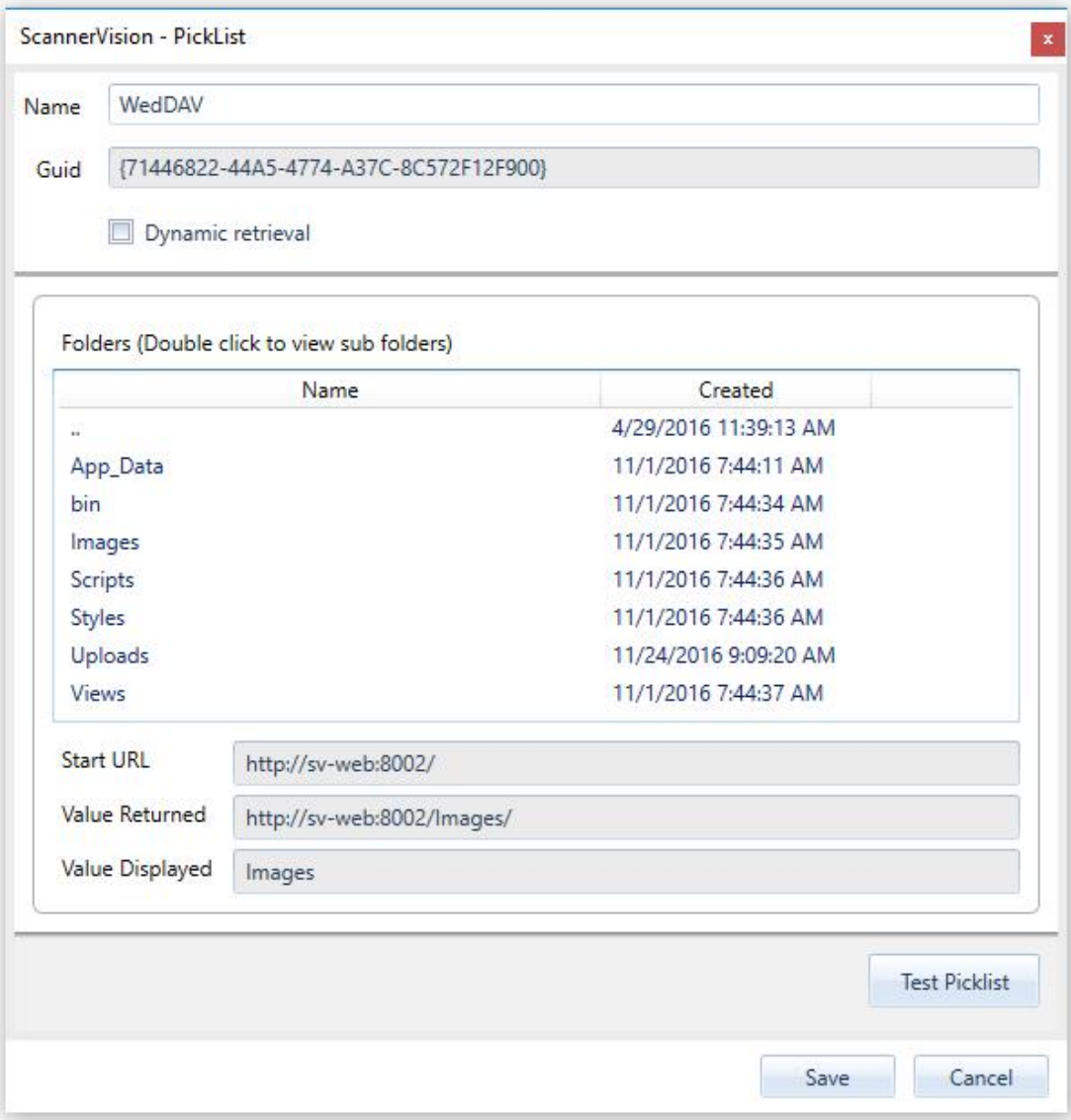
The URL of the web server. To secure the communication with the WebDAV server, specify the **https** protocol.

Please refer to Appendix I - SSL Certificates if you are presented with a SSL certificate warning dialog.

User Name & Password

The user name and password of the user with permissions to connect to and upload documents to the web server.

Press the "Connect" button to see the list of folders that will be presented to the user:



You can press the "Test Picklist" button to verify that the picklist works.

8.4.4.1.2 Dynamic Picklists

Dynamic picklists make use of metadata captured by previous template and/or global metadata questions to filter the values returned by the picklist.

We will work through a fictitious scenario to explain how you could make use of this functionality.

We are required to create a ScannerVision template for the storing of order payment receipts in a

SharePoint library. The library has fields for storing the customer number and order number. The template must require the user to select the customer and order numbers from drop down lists with values obtained from a SQL Server database.

To achieve this we will configure 2 template questions (or prompts to be more accurate) and 2 picklists. The prompts will be of type "Selected" and the picklists will be of type "SQL" (refer to the SQL query section for details on how to configure a SQL picklist).

The details of the prompts and picklists are tabulated below:

Question/Prompt	Tag Name	Picklist Name	Picklist Type
Customer number:	CUSTOMER	PlCustomers	SQL
Order number:	ORDERNUMBER	PlOrders	SQL

The PlCustomers picklist will be a simple look-up from the Customers table:

```
select    Id,
          CustomerNumber
from Customers
```

Every time the "Customer number:" drop down list box is opened on the client a request is sent to the ScannerVision server to obtain a list of customers. The server will execute the query above to get the list.

Now, once the user has selected the customer number, it would be great if we could limit the list of order numbers in the "Order number:" drop down list box to only those belonging to the selected customer. In order to do this several things have to be in place:

1. The "Customer number:" prompt must be the first one in the list of template questions. It is intuitive for users to start at the top of the list of questions and to work their way down. By putting the "Customer number:" first in the list we can be reasonably sure that the user will have made a selection of the customer number by the time the list of order numbers is fetched. The order of the questions in the "Questions" area of the "Client Capture" screen can be changed by right clicking on an item and selecting the relevant menu option from the context menu that appears.

2. The "Dynamic retrieval" option must be selected when we create the PlOrders pick list. This tells the ScannerVision client application to send all previously captured metadata (the customer number in this case) to the server when it asks for a list of order numbers.

3. The SQL query of the PlOrders picklist must filter the result set of the orders query on the customer Id to ensure that only order numbers for the selected customer are returned to the ScannerVision client.

The SQL query that we will use for the PlOrders picklist looks like this:

```
select    Id,
          OrderNumber
from Order
where CustomerId = "[CUSTOMER]"
```

In this query we are filtering the result set on the CustomerId column. Only records with a value in the CustomerId column equal to the [CUSTOMER] ScannerVision metadata tag will be returned. As stated above we have ensured that the [CUSTOMER] tag will contain a value by putting the "Customer number:" prompt at the top of the questions list.

Before the ScannerVision server executes the query it will replace all references to ScannerVision metadata tags with actual values.

If for example the user has selected the customer number "CUST00012345" the query that the ScannerVision server executes will look like this:

```
select    Id,
          OrderNumber
from Order
where CustomerId = "12345"
```

Why "12345" and not "CUST00012345" you may ask? Please refer to the Value displayed and value returned section for an explanation.

Now, when the user selects the "Order number:" drop down list box on the client, it will only contain the order numbers for the selected customer.

Please refer to the Character Escaping for information on character escaping in SQL queries.

8.4.4.1.3 Value displayed and value returned

For the purposes of the discussion to follow let us assume you have the following "Customers" database table:

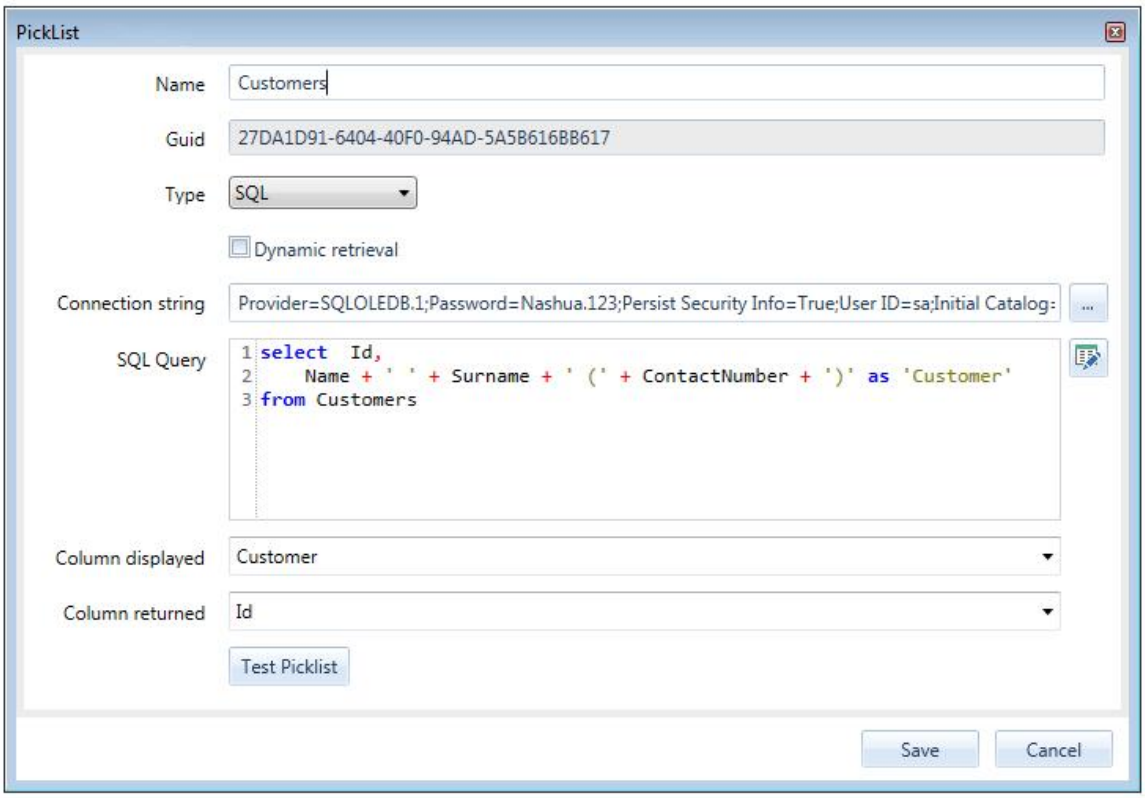
Id	Name	Surname	CustomerNumber	ContactNumber
1	Sally	Jones	SJ00000123	123-456-7890
2	Donald	Smith	DS00000124	321-654-0987

Relational database tables records usually have a unique number that identifies each record in the table. This number is called the primary key. In the table above the primary key is called "Id" and as you can see it is just an incrementing number. When a record is added to the Customer table, the database engine will automatically insert a new unique Id for the record (assuming that the table was designed to do this). Whenever a reference needs to be made to a customer in another table, say Orders, only the Id of the customer would be needed. Knowing the Id of the customer record is enough to uniquely identify the customer. We could also use the CustomerNumber field as it would also be unique but it is a bigger value and therefore inefficient to deal with in other tables. There are other reasons why you would not typically use the CustomerNumber field as a reference in other tables but those are beyond the scope of our discussion here.

Now, if a ScannerVision template requires the user to select a customer from a list we want to present the user with a list that contains the customer name, surname and possibly the telephone number or even the CustomerNumber. We need to be sure that the user is given enough information to be able to uniquely identify the customer. If we have two customers called "Donald Smith" we need to present the user with more information in order to distinguish between them. The important thing to understand is that whatever combination of fields from the Customers table we present to the user, that is not what we want to return to ScannerVision when the user submits the document. We want to return the customer's Id. So there is a distinction between what information we want to so to the user and what information we want to return to ScannerVision.

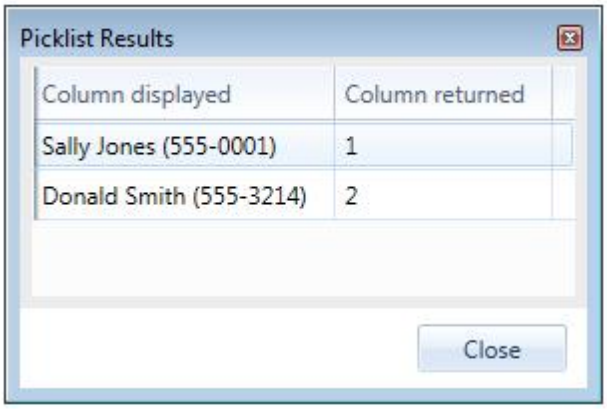
ScannerVision gives us the ability to make this distinction with SQL, VBScript/JScript and XML picklists. In the screen shot below you can see an example of how we could construct the solution we discussed above. We know we need to select enough information from the Customers table to present to the user as well as the Id of the record. So we are going to need the "Id", "Name", "Surname" and "CustomerNumber" columns but we need to return only two columns in the result set of the query we execute. The one column must be the Id and the other a combination of the "Name", "Surname" and "CustomerNumber" columns. To do this we are going to concatenate the "Name", "Surname" and "CustomerNumber" fields and call it "Customer". We will also add some formatting to the data by putting in spaces and brackets where appropriate to improve readability.

Here is a screen shot of the SQL picklist editor and the SQL query:



Notice that we have selected the "Id" column in the "Column returned" drop down list and the "Customer" column in the "Column displayed" list.

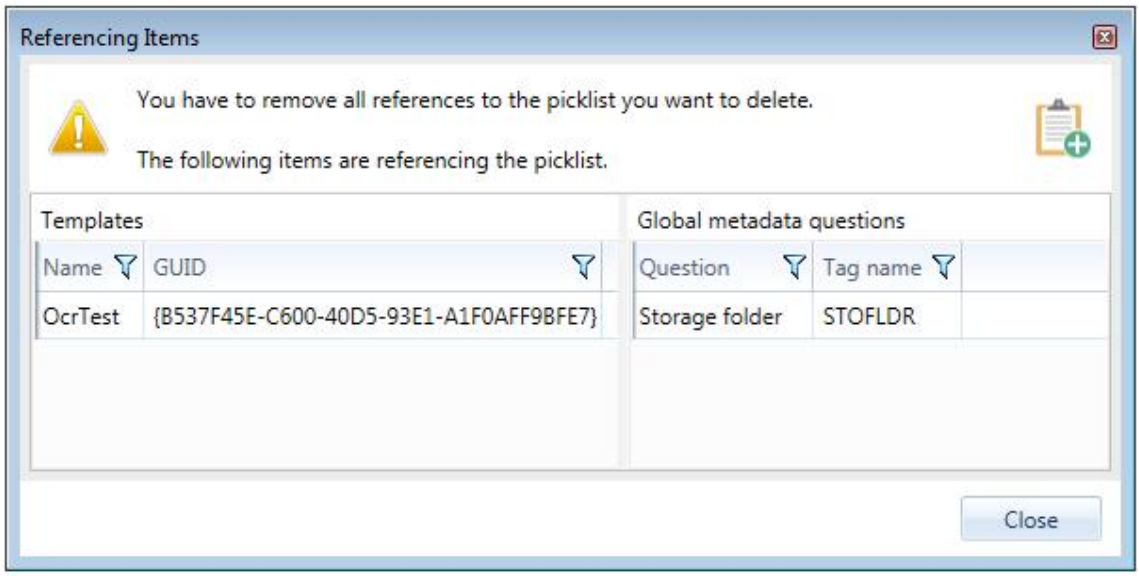
When we execute this query by pressing the "Test Picklist" button we get the following results:



Make sure that the column displayed and column returned show the correct data.

8.4.4.1.4 Deleting a Picklist

ScannerVision will not allow you to delete a picklist that is being referenced by a template or a global metadata question. If you tried to do so you will be greeted by the following dialog:



This dialog will list all the templates (left) and the global metadata questions (right) that are referencing the picklist. You will have to remove all the references to the picklist from the listed entities before you will be able to delete the picklist. To help you with this you can press the button in the top right hand corner. This will copy the contents of the dialog to the Windows clipboard. You can then paste this information into a text editor and print it out if needed.

8.5 Process

ScannerVision offers powerful image processing capabilities which fall into three categories:

Manipulating Documents

When manipulating a document the aim is usually to improve its quality by deskewing, despeckling, dot removal etc. or to improve it useability by rotation, changing color, contrast and brightness.

Reading Document Content

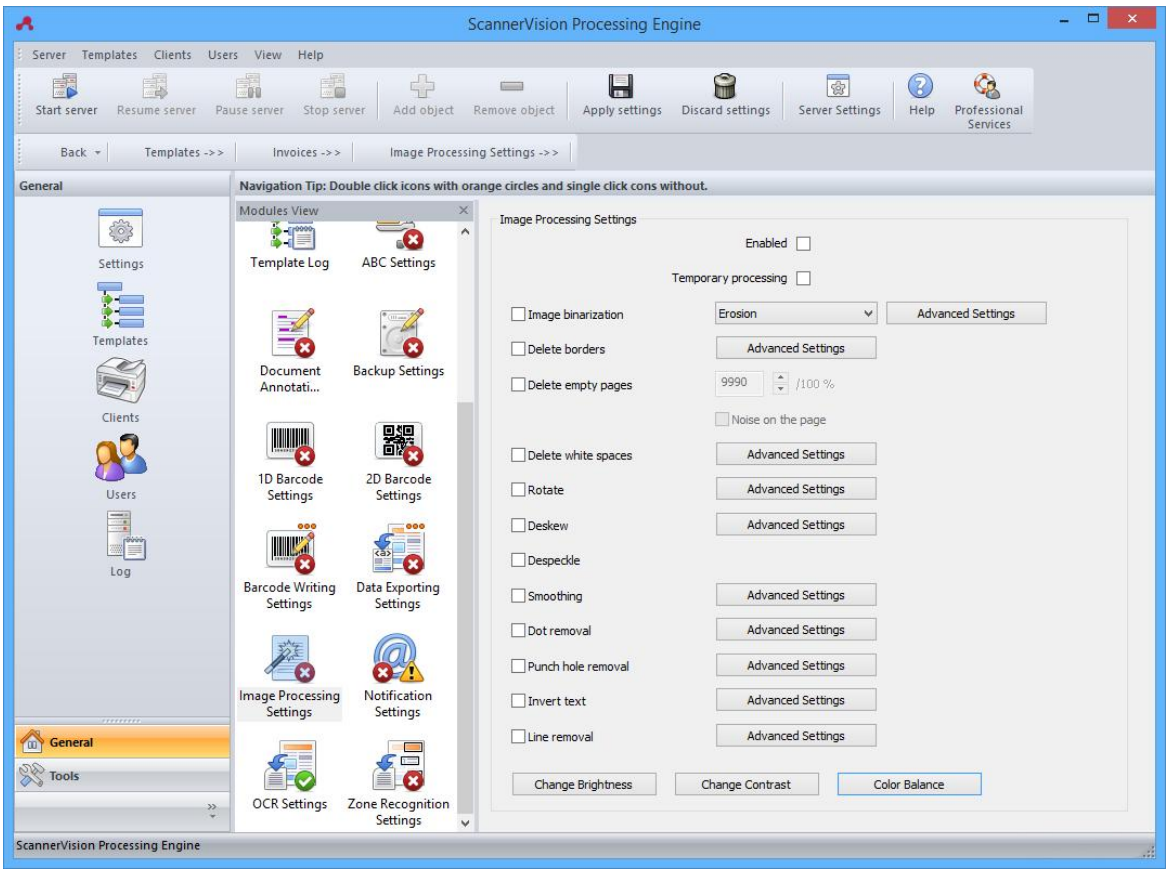
Reading document content involves the reading of barcodes or OCRing using any of the 5 OCR functions offered by ScannerVision.

Adding Content

Adding of content to a document could take the form of annotations (shapes, text, stamps etc.) or the writing of barcodes.

8.5.1 Manipulating Documents

All the image manipulation tools are found on the Image Processing Settings screen show below.



Enabled

Enables/disables all image processing functions. If enabled, individual function can still be enabled or disabled.

Temporary processing

When selected, the modifications to the original document will only apply to the Reading Document Content phase of the processing with the aim of improving the accuracy of the document reading functions such as OCRing and barcode reading. The document that is passed on to the connectors is the unmodified original. If "Temporary Processing" is not selected then the changes made by the

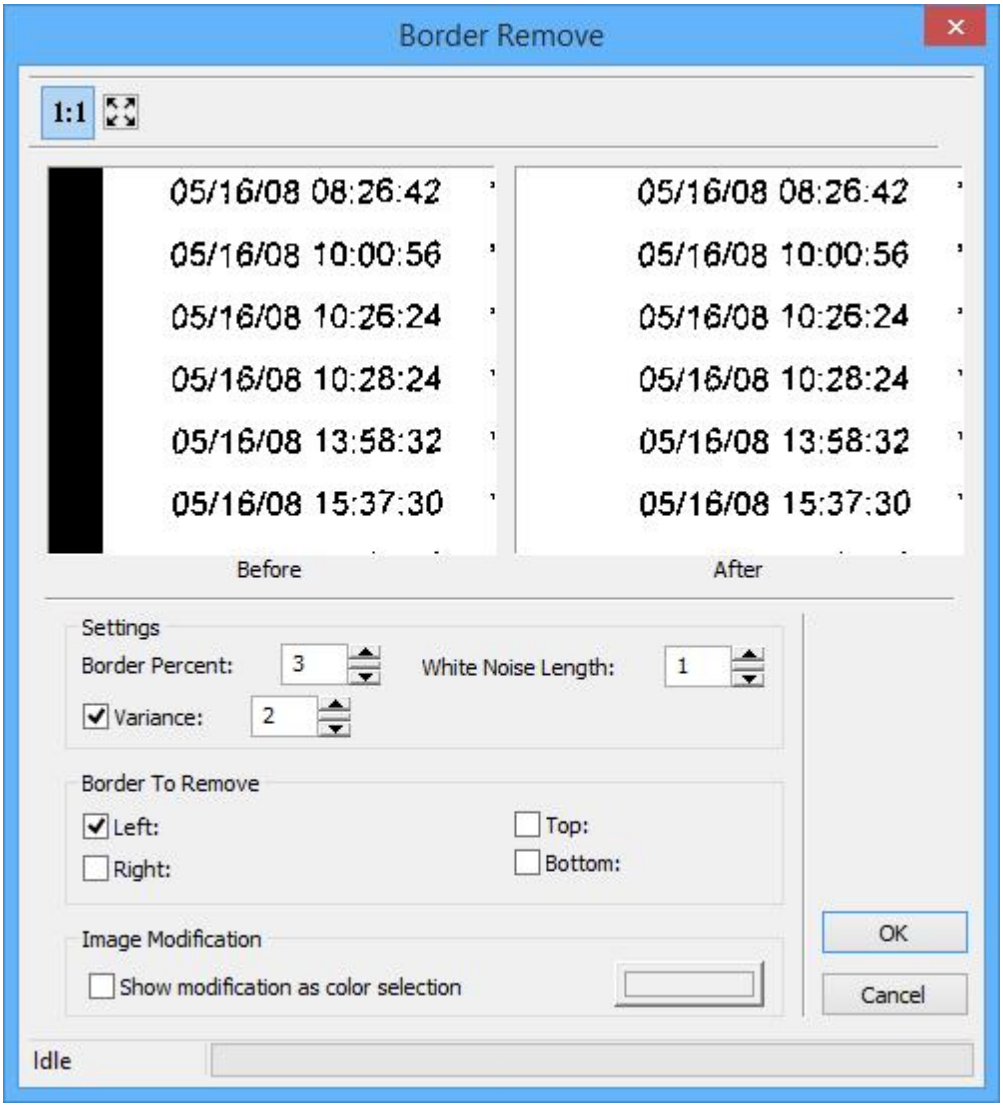
image processing functions are permanent.

Note

The "Image binarization" function is also available on the 1D/2D Barcode Settings screens.

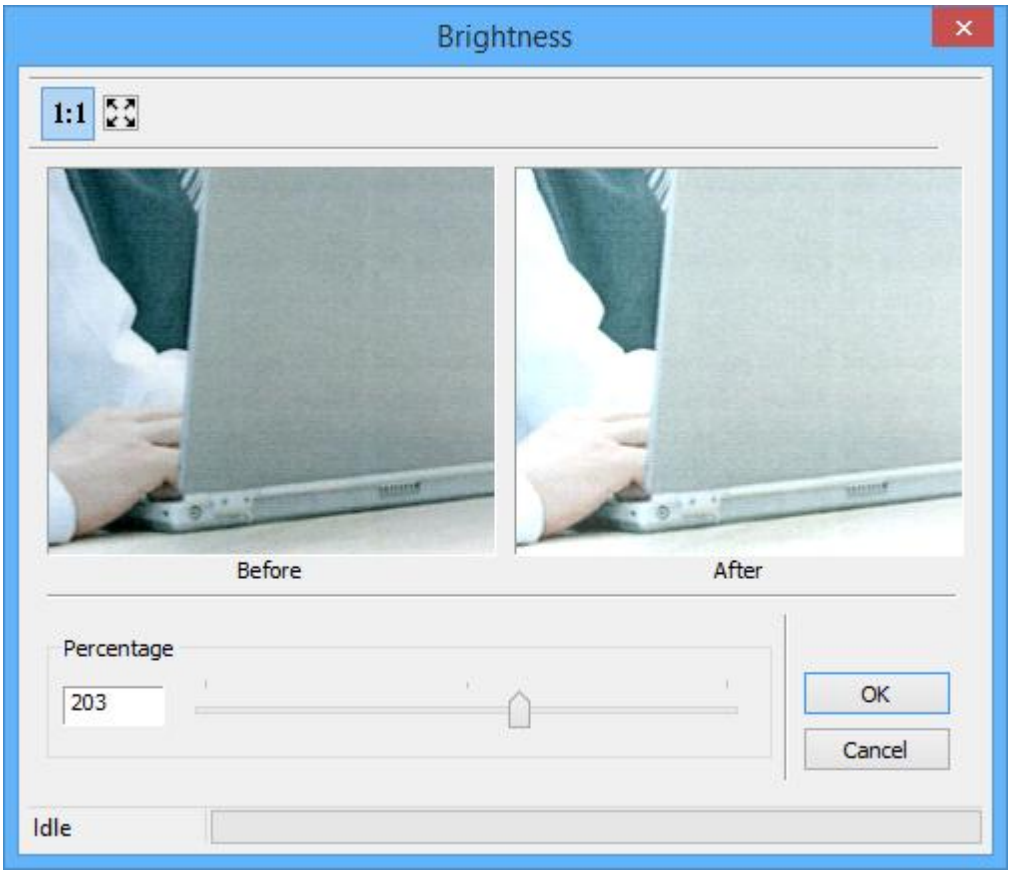
8.5.1.1 Border Removal

Use this function to remove borders from 1 bit black and white documents.



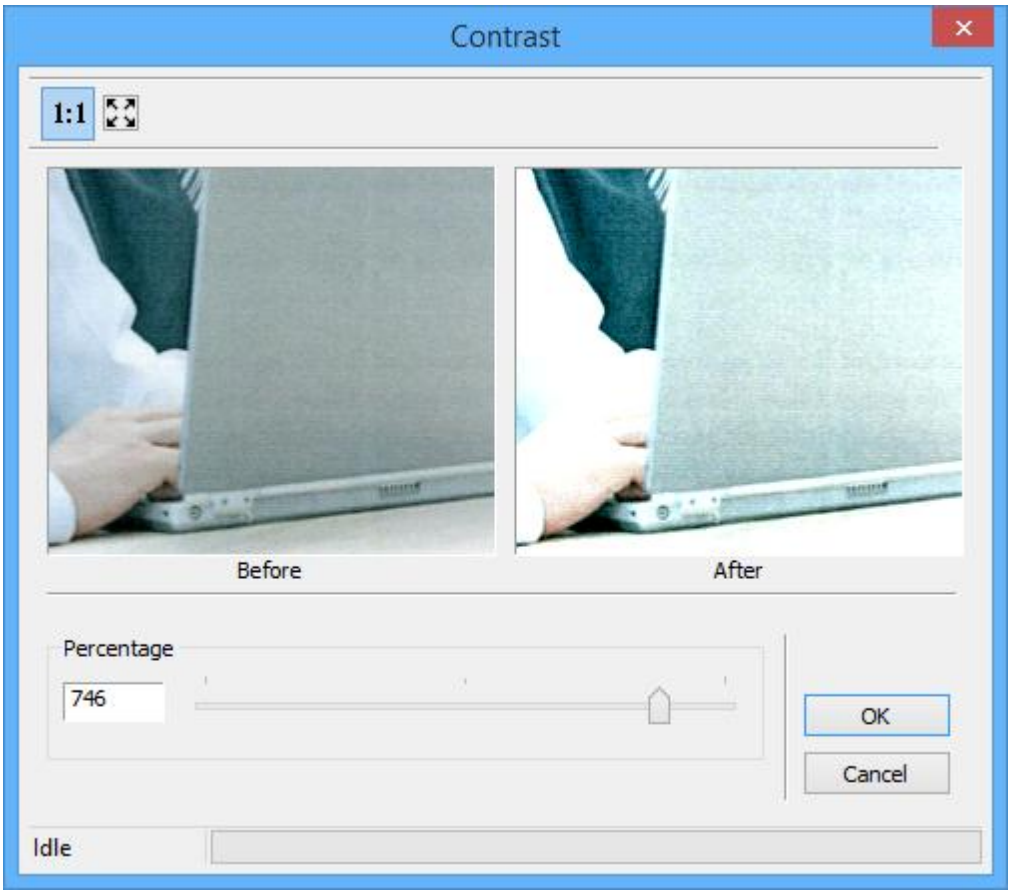
8.5.1.2 Change Brightness

Use this command to the change the brightness of a document.



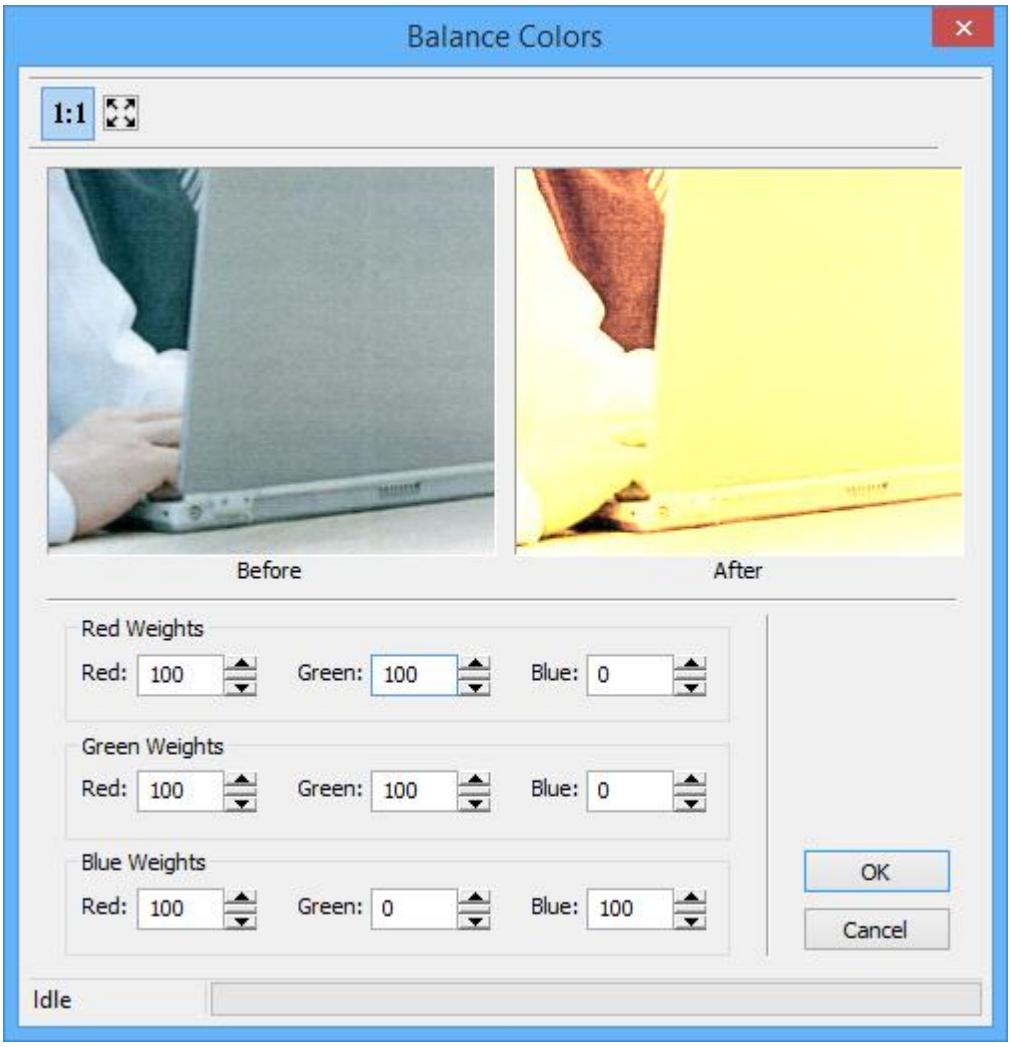
8.5.1.3 Change Contrast

Use this command to the change the contrast of a document.



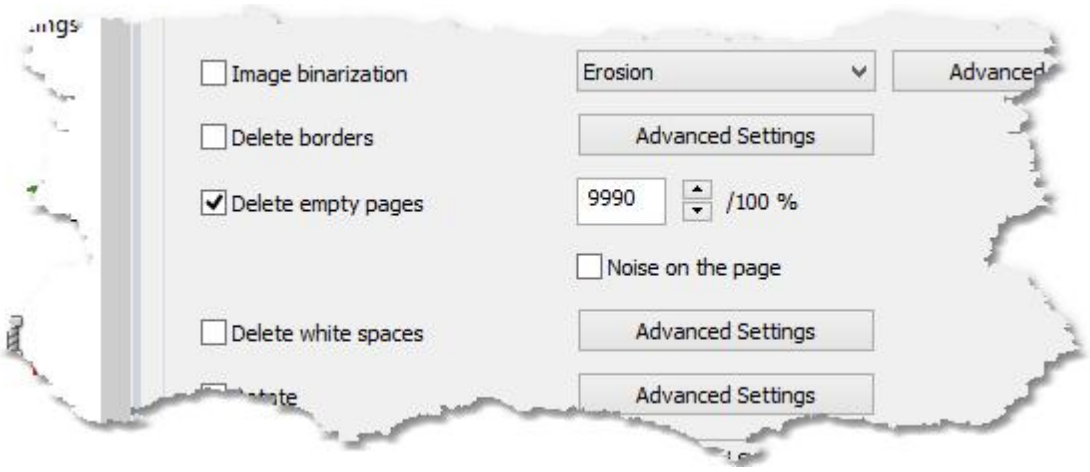
8.5.1.4 Colors Balancing

Use this function to adjust the color balance in documents.



8.5.1.5 Delete Empty Pages

The Delete Empty Pages" feature allows you to remove pages which do not contain any data. When ScannerVision detects a blank page it makes an estimation of the certainty that the page is blank in the form of a number with a value between 0 and 10000 with a higher the value indicating a higher certainty that the page is blank. The "Delete empty pages" edit box allows you to specify the threshold of when a page should be considered blank. If for example you want ScannerVision to remove a page only if it is 98% or more certain that it is blank, you would put in a value of 9800 = 98 * 100. The default value is 9990 which is equivalent to 99.90% certainty.



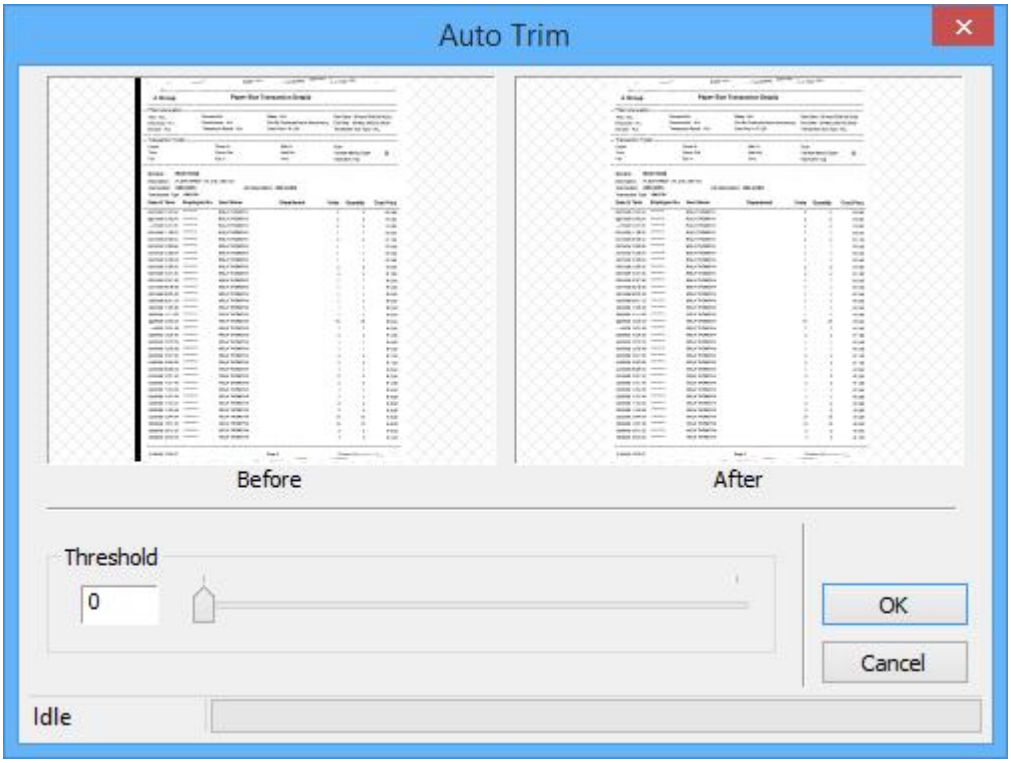
Be careful with the accuracy though. You may lose pages if you set the accuracy too low. ScannerVision will warn you if you choose a value lower than 9800. Rather err on the side of caution and choose a high level of certainty. If you are too aggressive you could lose pages that were not supposed to be removed.

Noise on the page

If pages contain noise such as dots and speckles that are introduced by poor scanning you can select the "Noise on the page" check box. This will instruct ScannerVision to take this into consideration when determining if a page is blank.

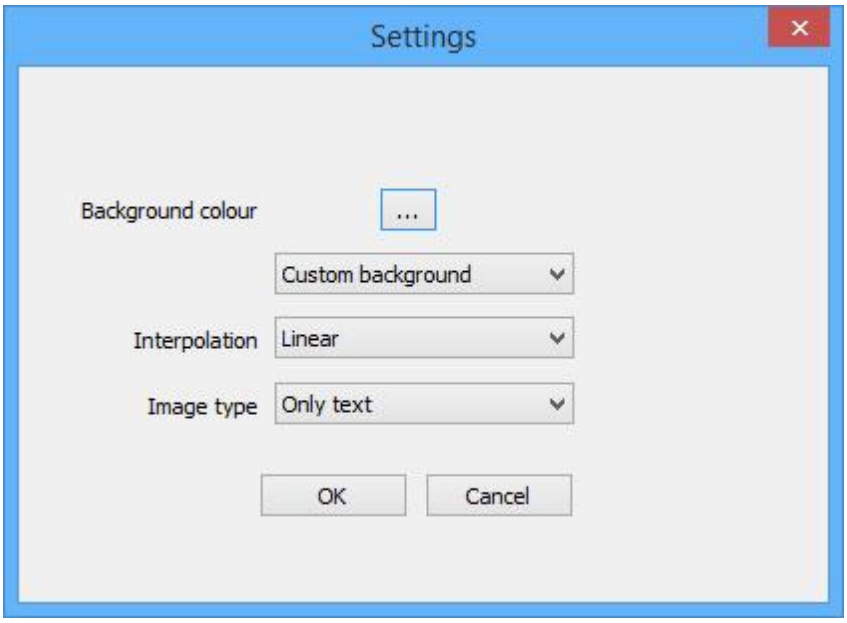
8.5.1.6 Delete White Spaces

Use this function to remove white space that is introduced with the scanning of non-standard paper sizes. The threshold bar determines how much space is removed.



8.5.1.7 Deskew

Use this function to automatically correct slight skewing of scanned documents. The calculated rotation is limited to 20 degrees in either direction and performs best when documents contains mainly horizontal lines of text.



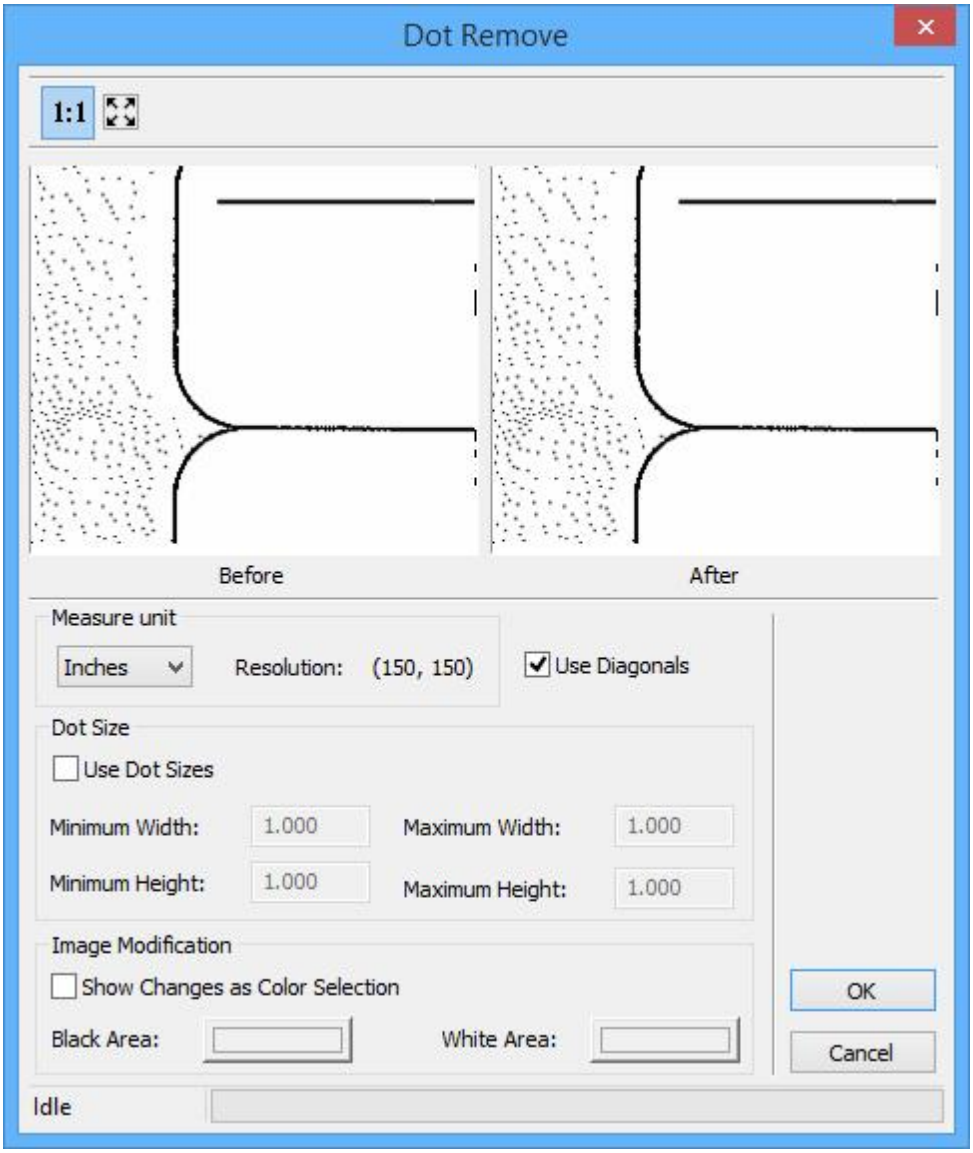
8.5.1.8 Despeckle

This function removes noise from documents such as faxes or images scanned on a scanner with a dirty exposure glass.

This function is automatic and requires no user configuration.

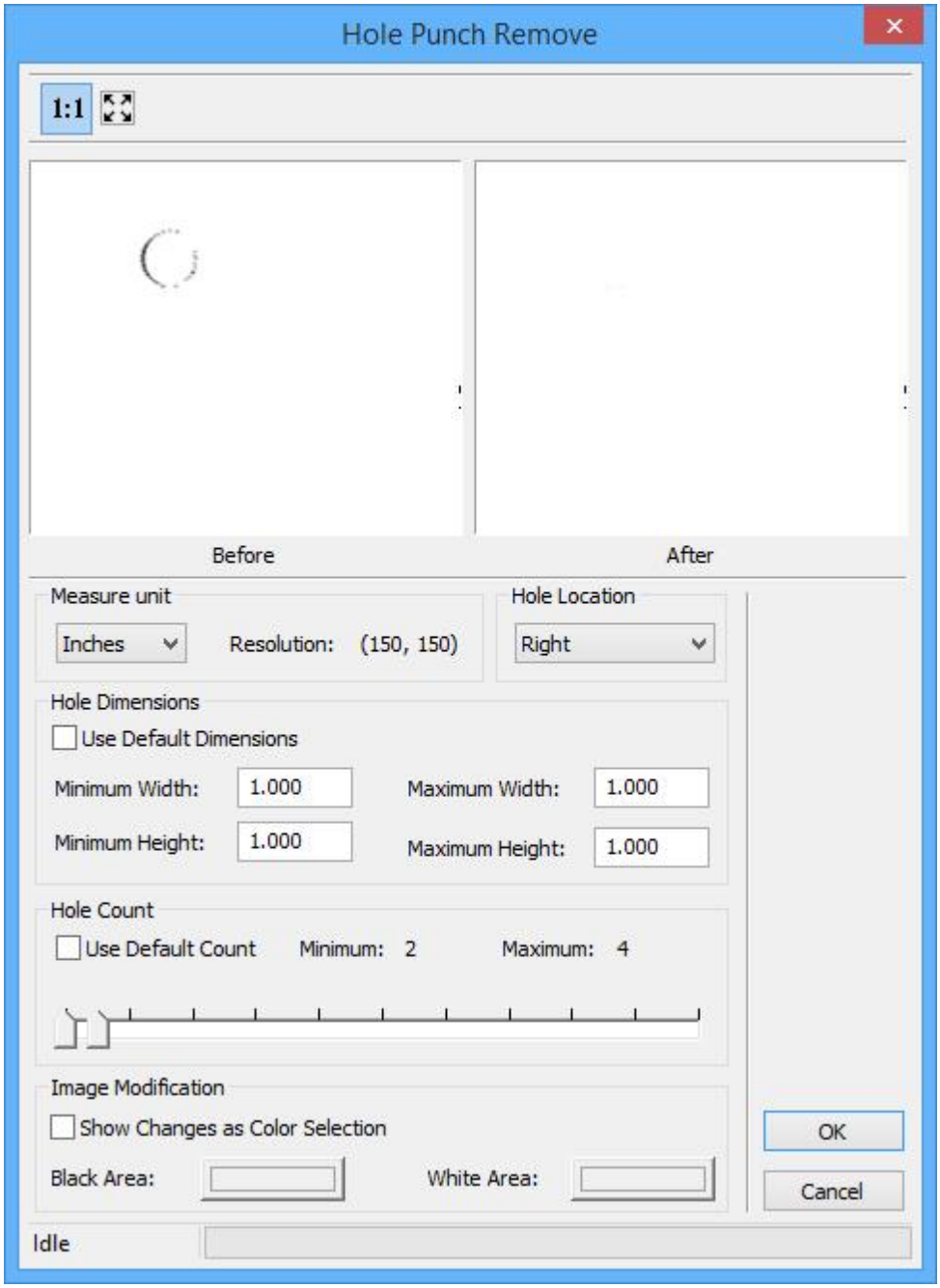
8.5.1.9 Dot Removal

Use this function to remove dots from 1 bit black and white documents. You can specify the size range of dots that must be removed.



8.5.1.10 Hole Punch Removal

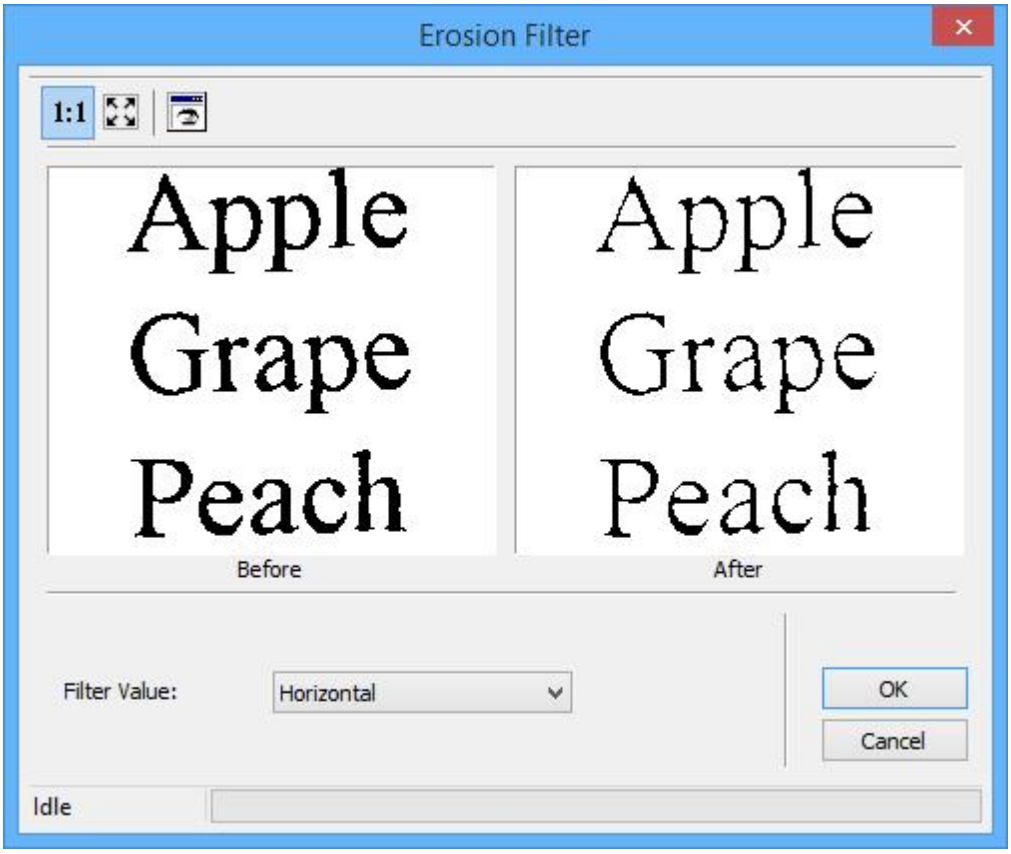
Use this function to remove punch hole marks from 1 bit black and white documents. The location, dimensions and number of punch holes to be removed can be specified.



8.5.1.11 Image Binarization

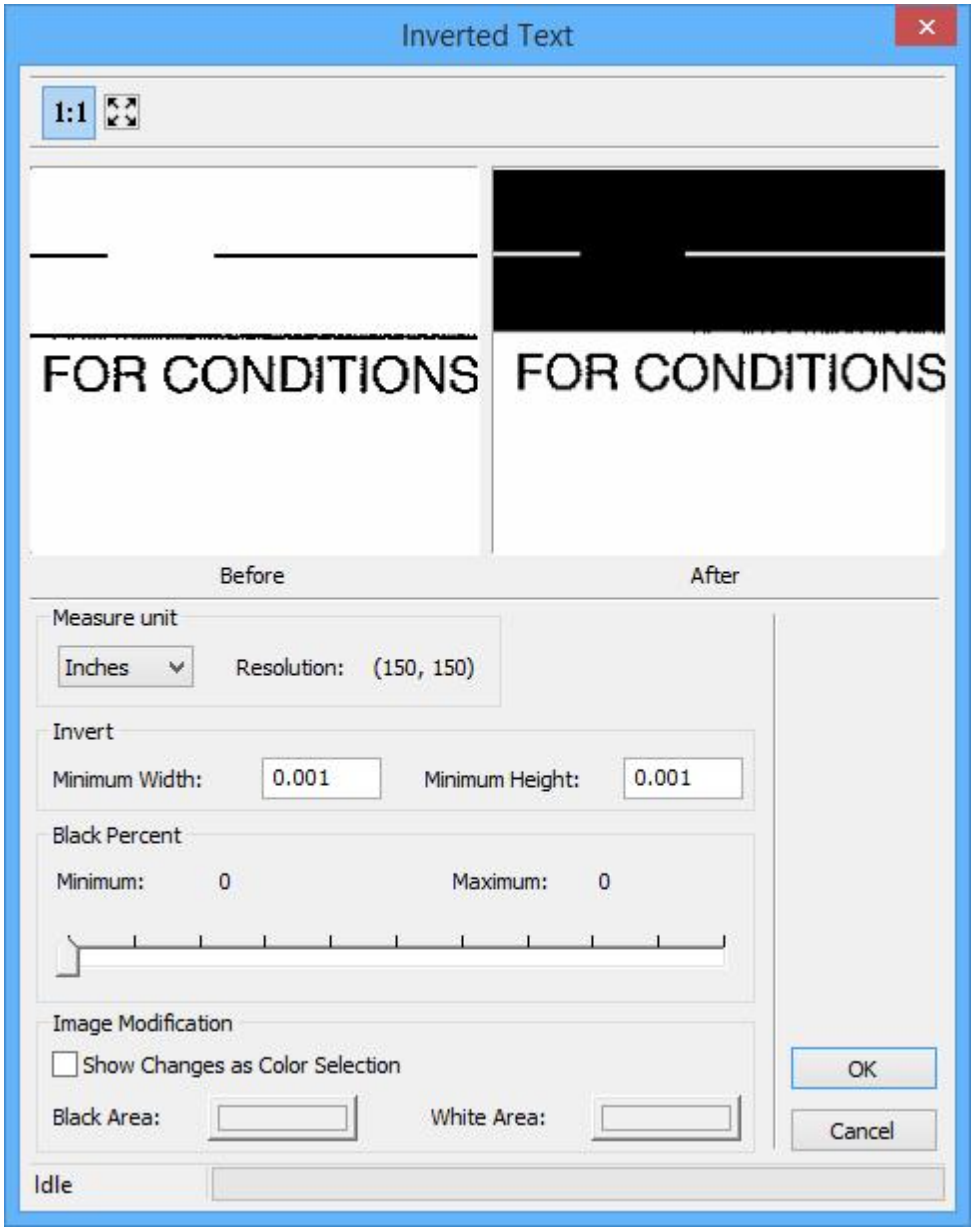
The Image binarization function is used to convert 256 gray-levels images into monochrome levels. This means that the threshold value for the binary image is not fixed, but rather dynamic. This is calculated irrespective of changes both in foreground darkness and in background lightness and overall illumination levels. By adjusting contrast levels image binarization increases the accuracy of low-contrast documents containing hand-written or typed characters with variable lines thickness and darkness automatically.

Using the Advanced Settings, binarization can be applied Horizontally, Vertically, Omnidirectional or diagonally



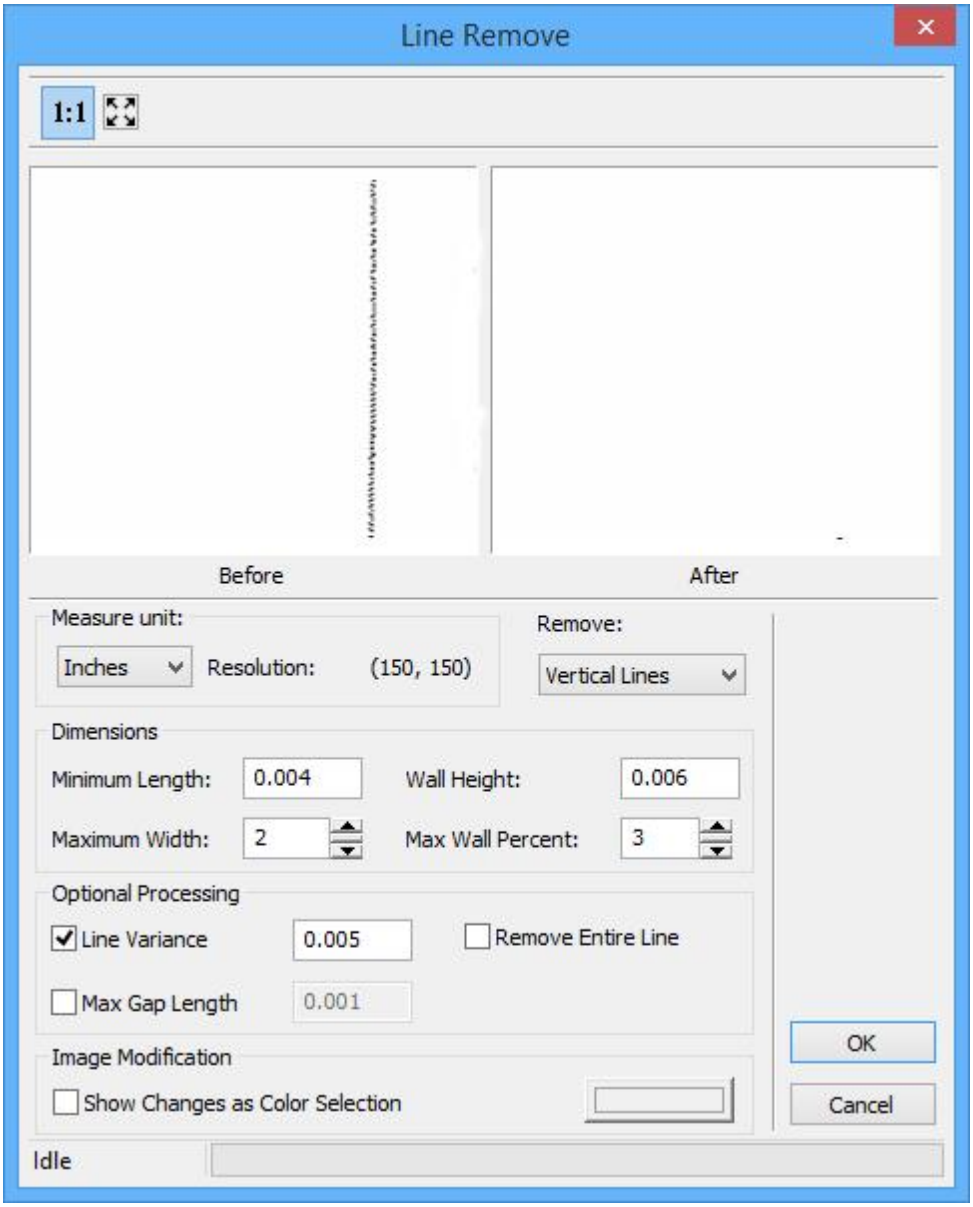
8.5.1.12 Invert Text

Use this function to invert the color of pixels in 1 bit black and white document. This is used to invert white-on-black text to black-on-white and visa versa.



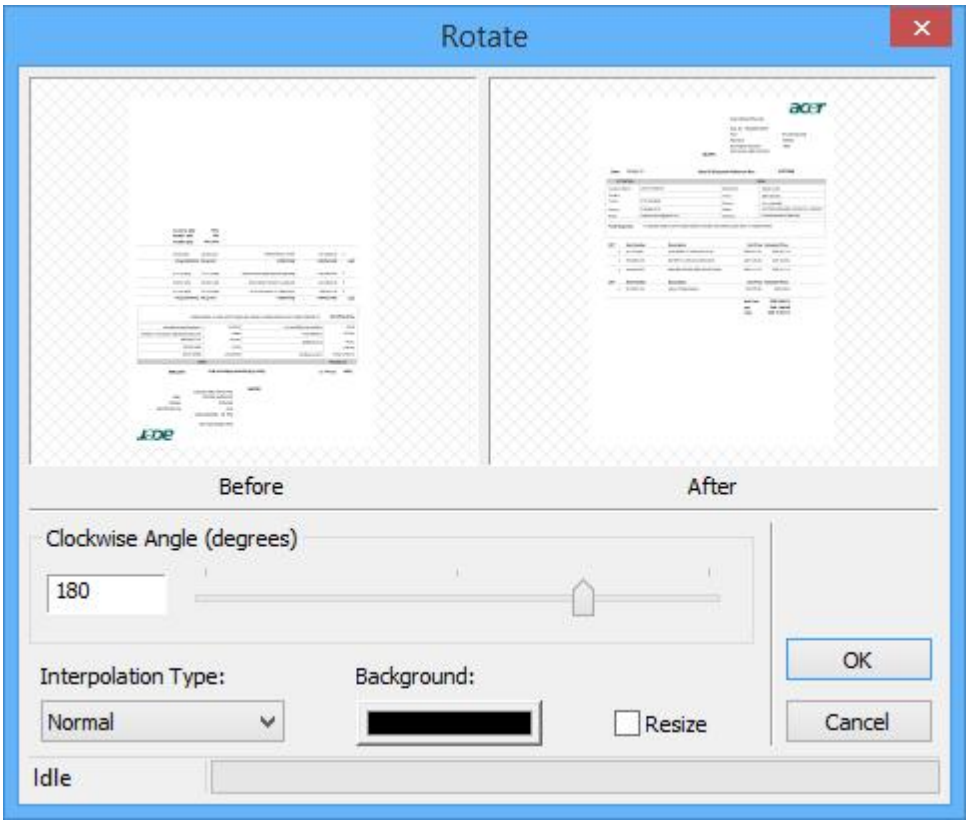
8.5.1.13 Line Removal

Use this function to remove horizontal and vertical lines from scanned, 1 bit black and white documents.



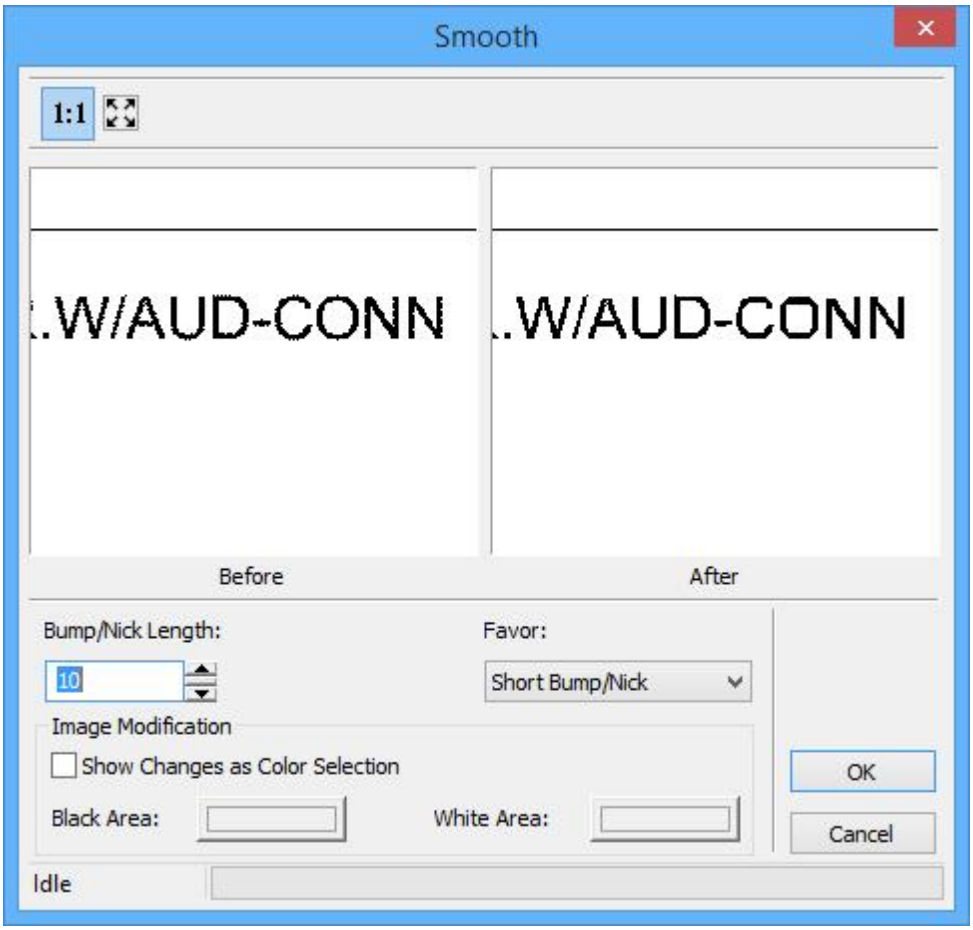
8.5.1.14 Rotate

The rotate function rotates pages between 0 and 360 degrees.



8.5.1.15 Smoothing

Use this function to smooth the bumps and fill the nicks of 1 bit black and white documents.



8.5.2 Reading Document Content

The reading of document content involves the reading of 1D and 2D barcodes as well as optical character recognition (OCR). OCR fall into 2 categories namely full document OCR and zonal OCR. Full document OCR does exactly that, it OCRs the whole document. Zone OCR on the other hands constrains OCRing to a specific region on the page. There are four types of zone OCR namely:

OCR Optical Character Recognition - the recognition of regular text.

ICR Intelligent Character Recognition - the recognition of hand written text.

OMR Optical Mark Recognition - the recognition of tick marks, X's, lines, check marks, and scribbles commonly found on surveys, polls, academic exams and official applications.

MICR Magnetic Ink Character Recognition(MICR) - the recognition of special numbers and symbols typically found on checks.

The content that is read from a document is available as Metadata. With all content reading functions except full document OCR can you specify what the name of the metadata is, e.g. when reading a barcode you can specify the name of the metadata tag in which the barcode data must be stored. In the case of full document OCR the tag names in which OCRed text is stored are fixed and are called OCRTEXT, OCRTEXTXML, OCRTEXTFILE and OCRTEXTXMLFILE. Since documents can be very large the size of the OCRTEXT and OCRTEXTXML tags are limited to 64MB*. If you need to get access to the full body of OCRed text you can reference the OCRTEXTFILE or OCRTEXTXMLFILE tags which hold the fully qualified path to text files containing all the OCRed text.

* There is one exception to the 64MB limit. When you configure a data export rule in the Data Export Settings tab and you reference ONLY the OCRTEXT or OCRTEXTXML tags (no metadata functions, custom text etc.) in the "Custom output" field e.g. [OCRTEXT] or [OCRTEXTXML], all OCRed text will be exported to the specified file.

For details on the format of the Xml output of OCR text please refer to the Tags section.

Achieving good results

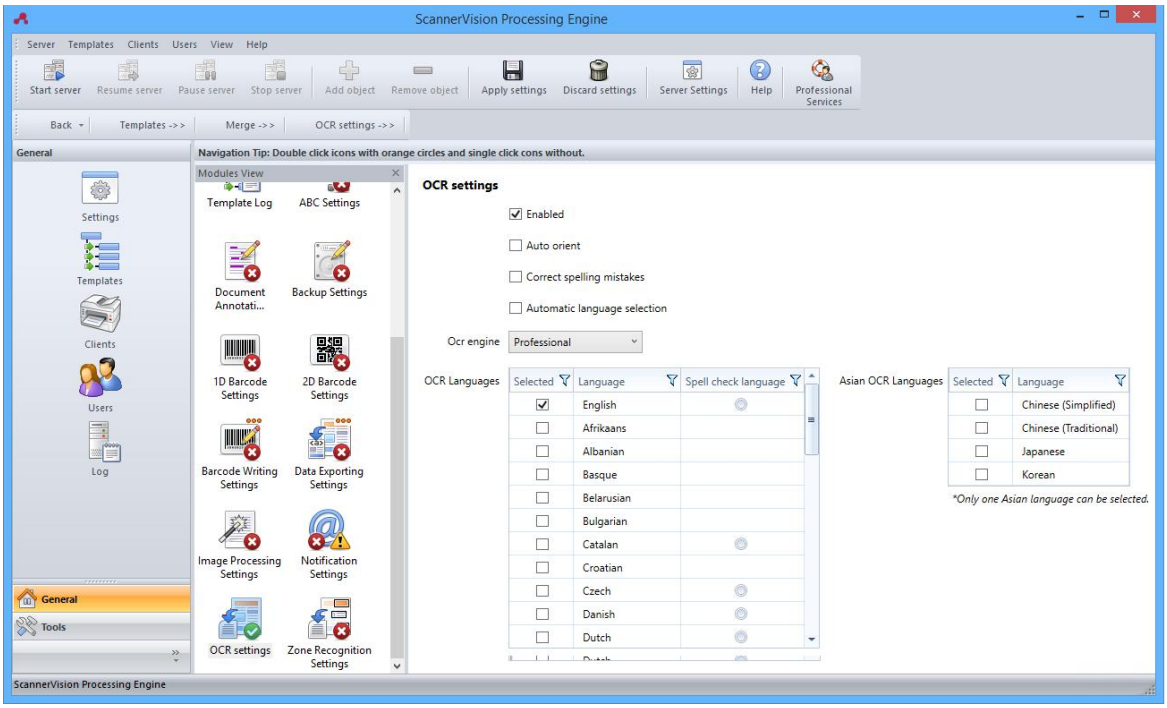
The ability for humans to read a document or recognize shapes, marks and lines on a page comes naturally. For a computer this is very difficult and therefore we have to give it the best possible chance to do so. The single most important factor that determines the accuracy of the computer's reading ability is the quality of the document it is reading.

As far as possible try to:

1. Ensure that documents are scanned at 300dpi or higher.
2. Use Tiff documents instead of PDF documents as input to ScannerVision. See Appendix D - PDF Input Documents for more information on this.
3. Ensure that scanners are clean so that noise is not introduced in the scan.

8.5.2.1 Full Document OCR

ScannerVision offers two OCR engines namely Nuance Omnipage and Leadtools Advantage with the former being the default. You may find that the accuracy and speed of the two engines may differ in your environment so choose the one that best suits your needs.



Enabled

Enables/disable the OCR engine. OCR has to be enabled for certain output document types such as Searchable PDF and PDF/A.

Auto orient

Automatically rotates the page being OCRed to the upright position if it is rotated.

Correct spelling mistakes

Automatically correct spelling errors based on the language you have selected in the "OCR Languages" grid's "Spell check language" column. There can be only one spell check language even if multiple OCR languages are selected. This will be explained in more detail below.

Automatic language selection

Automatically detects the language of the text being OCRed. When you select this option all the languages in the "OCR Languages" grid are selected and the ability to select/deselect languages becomes disabled. When the option is deselected the language selection of the saved template is loaded.

If you want an Asian language to be included in the language selection, you have to select it explicitly since only one Asian language can be used at a time.

OCR Engine

Selects the OCR engine to use. The choice you make here will depend on your particular requirements. The engines do not support the same languages and you may find that the one performs better in your environment than the other. The default engine is "Professional". You will notice that as you change the OCR engine the list of OCR languages also changes.

OCR Languages

The languages that are supported by the OCR engine. These are different for the different OCR engines. The languages that you select helps the OCR engine to recognize characters. For example, if you choose English only then the engine knows that if it encounters a character that could be either of 2 or more possibilities but only one exists in the standard English character set, it will choose that character. In addition, if you also select English to be the spell check language, the OCR engine can go one step further and verify that the words it has recognized form part of the English language. If for example the engine recognizes a character that could either be a zero or and the letter "o" it will look at the context of the character. If the character forms part of a word that exists in the language, the engine will choose the character "o" instead of the number "0" (zero).

Language	Advantage Engine		Professional Engine	
Arabic	**Character Set**	**Spell Check**	**Character Set**	**Spell Check**
English	X	X	X	X
Afrikaans			X	
Albanian			X	
Basque			X	
Belarusian			X	
Bulgarian	X		X	
Catalan	X		X	X
Chinese (Simplified)	X		X	
Chinese (Traditional)	X		X	
Croatian			X	
Czech	X		X	X
Danish	X		X	X
Dutch	X		X	X

Estonian			X	
Faroese			X	
Finnish	X		X	X
French	X	X	X	X
Galician			X	
German	X	X	X	X
Greek	X		X	X
Hungarian	X		X	X
Icelandic			X	
Indonesian	X		X	
Italian	X		X	X
Japanese	X		X	
Korean	X		X	
Latvian	X		X	
Lithuanian	X		X	
Macedonian			X	
Norwegian	X		X	X
Polish	X		X	X
Portuguese	X		X	X
Brazilian Portuguese			X	X
Romanian	X		X	
Russian	X		X	X
Serbian	X			
Serbian - Latin			X	
Serbian - Cyrillic			X	
Slovak	X		X	
Slovenian	X		X	X
Spanish	X	X	X	X

Swedish	X		X	X
Telugu	X			
Thai	X			
Turkish	X		X	X
Ukranian	X		X	
Vietnamese	X			

Asian OCR Languages

Asian languages are supported by the "Advantage" and "Professional" engines. Only one Asian language can be selected at any given time - including for "Automatic language selection" - and no spell check support is available.

8.5.2.2 Reading Barcodes

ScannerVision supports the reading of the barcode symbologies listed below:

1D

- EAN 8
- EAN 13, EXT 2, EXT 5
- UCC/EAN 128
- UPCE
- CODE 39 (CODE 3 OF 9)
- CODE 93
- CODE 128
- CODABAR
- Interleaved 2 of 5
- Standard 2 of 5
- RSS14, Limited, Expanded, Stacked
- Code 11 C, K
- MSI MOD10, 2 MOD 10, MOD 11, MOD11 MOD 10
- Patch Code
- PostNet

- Planet
- Australian Post C, N, Bar State
- Royal Mail
- USPS OneCode

2D

- PDF417 Mode 0, Mode 1, Mode 2, Mode 3 Ext
- Micro PDF417
- DataMatrix Square, Rectangle, Small
- QR

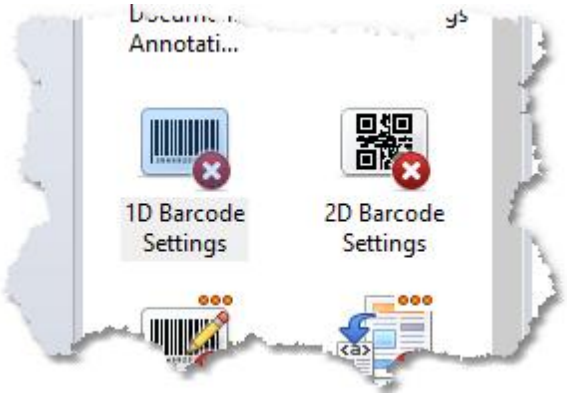
Barcodes in ScannerVision can be used as the source of metadata or as document splitters or both. As the source of metadata the data contained in the barcode is assigned to a user defined metadata tag for use elsewhere in the template. As a document splitter the existence of the barcode on a page triggers the separation of the original document into separate documents on the particular page that the barcode was found. The page on which the barcode was found can optionally be omitted from the resulting separated document. The contents of the barcode that triggered the split is written to the DOCUMENTSPLITTER metadata tag. Multiple barcode symbologies can be defined as document splitters in which case they are processed in the order they appear in the list. This will be explained in detail later.

When multiple instances of a particular barcode symbology can appear in a document or on a single page a Regular Expression (Regex) can be used to identify which barcode was found.

Zones can be defined in which to read barcodes. Refer to the Zone OCR section for details.

8.5.2.2.1 Configuring Barcodes

Reading of 1D and 2D barcodes are supported by ScannerVision and each type is configured separately. To access the configuration screens click the relevant button in the Modules View as shown below:



Although 1D and 2D barcodes are configured separately, their settings are exactly the same in all respects except for the particular barcode symbology that is configured. As such, the discussion below focuses on 1D barcodes only but everything applies to 2D barcodes as well.

You can read barcodes in one of two modes. You can let ScannerVision automatically read all the barcodes it can find by selecting the "Automatic barcode recognition" option. The values of the barcodes that are found are written to metadata tags with the following format:

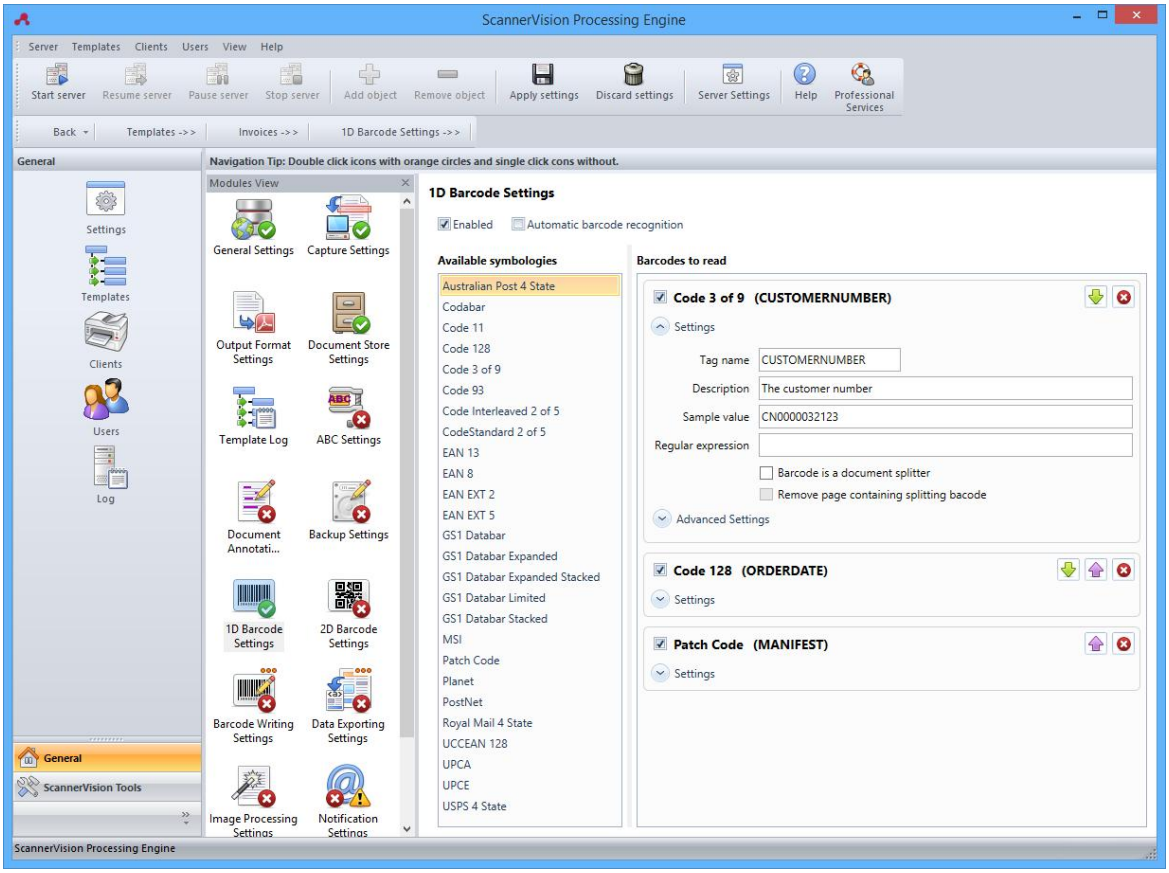
BC + 1D/2D + VALUE + Page number

e.g. BC1DVALUE0, BC1DVALUE2, BC2DVALUE0 etc.

With automatic barcode recognition the barcode engine uses default settings for all barcode symbologies which cannot be modified.

If you configure individual barcode symbologies on the other hand you have control over every setting on a per barcode level. To configure individual barcodes drag the required barcode symbology from the "Available symbologies" list onto an empty spot on the "Barcodes to read" area. If you drag a symbology over an existing one the symbology of the barcode is changed to the new one without affecting any of the other properties and settings.

Since there are many settings you can configure, you can show as much or as little information as you need by collapsing or expanding the "Settings" and "Advanced Settings" groups.



Enable/Disable

To enable or disable a barcode symbology, select the check box to the left of the symbology name as show below:

Tag name

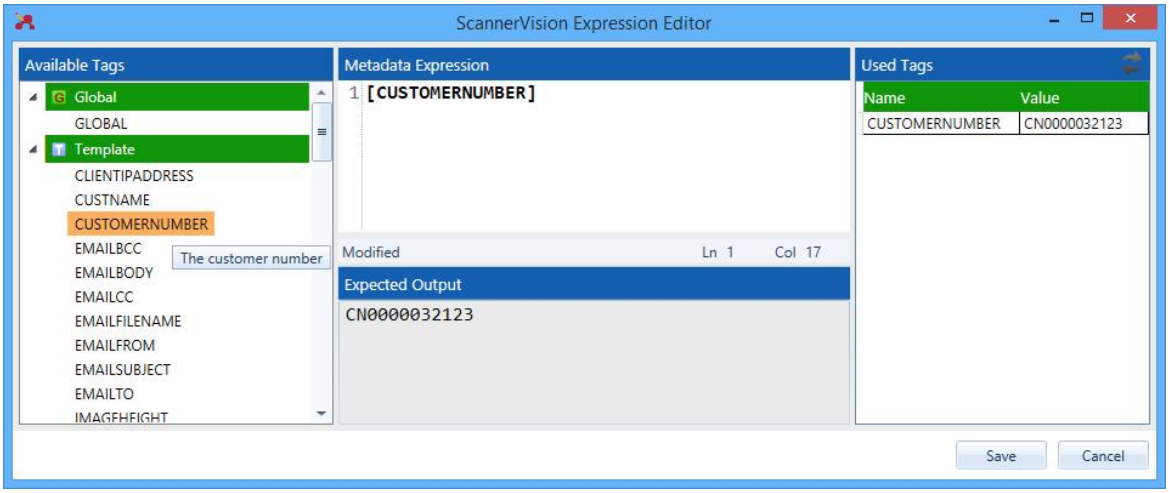
This is the name of the metadata tag that will hold the content of the barcode.

Description

A description of the data that the barcode represents. This description is shown as a tool tip in the ScannerVision Expression Editor when the mouse is hovered over the metadata tag name. You can see this in the screen shot of the ScannerVision Expression Editor below. The tool tip of the CUSTOMERCUMBER tag is "The customer number." which is what was entered in the description field of the barcode.

Sample value

A value that is typical of the data contained in the barcode. This value is assigned as the metadata tag value in the ScannerVision Expression Editor. In the screen shot below you can see that the CUSTOMERCUMBER tag contains the value "CN0000032123". This value was assigned automatically to the sample value.



Regular Expression

A regular expression can be specified to ensure that the correct barcode value is read. This is useful in situations where more than one barcode of a specific type exists on a page. An example would be where two Code 3 of 9 barcodes are found on a page, one containing an invoice number and the other an order number.

If for example the invoice number starts with "INV" and the order number with "ORD" you can add two Code 3 of 9 symbologies. For the invoice number symbology you would add a regular expression that checks if the data in the barcode starts with the characters "INV" and if it does the value is assigned to tag. You would do the same for the order number but with the regular expression checking for "ORD" instead.

Barcode is a document splitter

When this option is enabled ScannerVision will split the original document on the page on which the barcode is found. All remaining pages will form part of the new document unless another document splitting barcode is found. By default the page on which the barcode was found is included in the new document as the first page. If you want the page to be excluded select the "Remove page containing

splitting barcode" option.

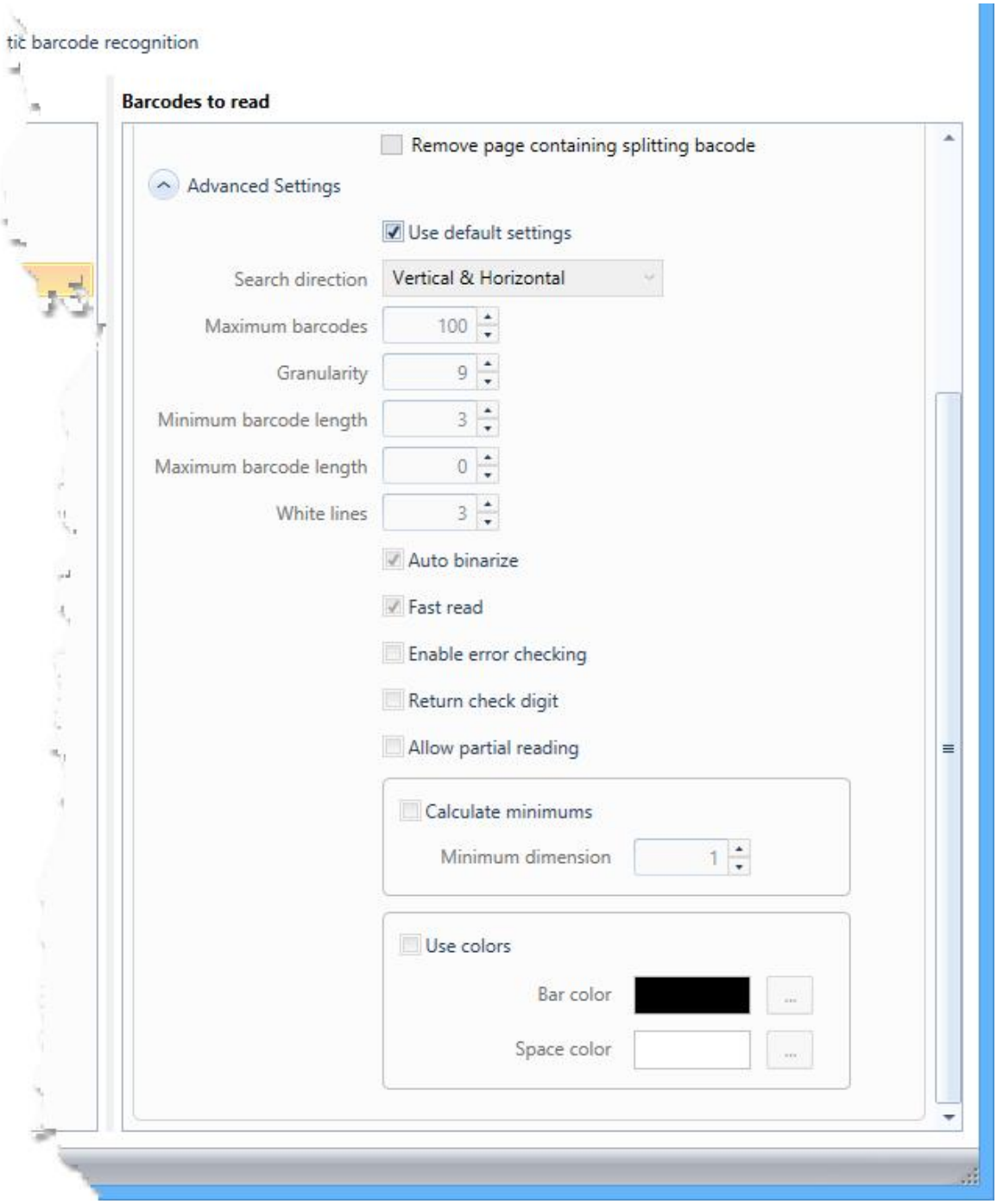
Once the original document has been separated into multiple documents, each new document is passed through the same template again in the order in which they were created. Document splitting will however not be performed on the second pass.

Remove page containing splitting barcode

This option is only available when the "Barcode is a document splitter" check box is selected. When this check box is selected the page on which the splitting barcode was found will not be included in the new document. This option is typically used in situations where barcode separator pages are used to indicate the start of a new document.

Advanced Settings

The Advanced Settings section shown below offers more control over how the barcode is read and will largely determine how successful the barcode engine is with the recognizing of barcodes. In most situations the default settings will yield good results but there are situations where adjusting certain settings may improve the results.



Use default settings

Select this option if you want the barcode engine to use its defaults settings for this particular symbology. It is recommended that you always try this option first and if you find that the barcode recognition is not working as expected you can change some of the advanced settings.

Search direction

Indicates in which direction the barcode engine searches for barcodes. To cater for pages that are rotated through 90°, 180° or 270°, it is recommended that you choose the "Horizontal & Vertical" option. This will force the barcode engine to search for barcodes in all possible directions. There is a slight performance penalty with this configuration but it will yield more reliable results.

Max Barcodes

Specifies the maximum number of barcodes that will be read per page**.** A value of 0 means all barcodes are read.

Granularity

Number of scanned lines per column to skip when reading a barcode. With a value of 9 for example each tenth line will be scanned when looking for a barcode. Scanning every line will slow down the search process while skipping too many lines may skip over the barcode. A value of 9 is typical.

Minimum barcode length

The minimum length of a barcode string when searching for a non fixed length bar code.

Maximum barcode length

The maximum length of a barcode string. The maximum value is 63.

White Lines

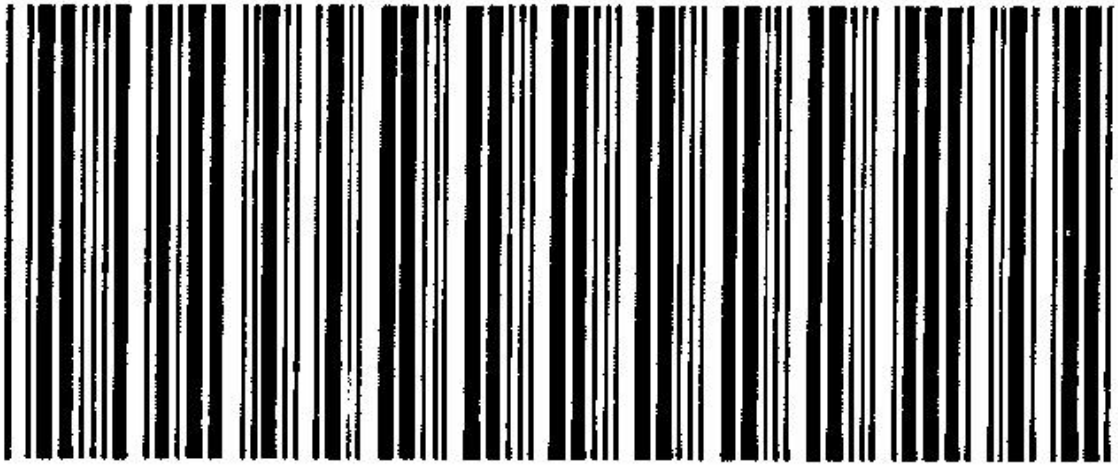
The number of lines of white space above and below the barcode symbol.

Auto binarize

Apply automatic binary segmentation to improve barcode recognition results. The effect of this setting is shown below:

Before binarization:

After binarization:



This command does not support the following image formats:

1. 12, 16 or 32 bit grayscale images

2. 48 or 64 bit color images

3. Signed data images

Fast read

Employs a fast barcode recognition algorithm which is suited for most barcode. This option is selected by default.

Use error check digit

Enables the use of the optional check word for validity checking when reading a barcode (if supported). If the validity check fails no barcode is returned.

This option will only have effect when reading barcode symbologies that support the optional check word which include:

- Code 3 of 9 (Code 39)
- Standard 2 of 5 (Industrial 2 of 5)
- Interleaved 2 of 5
- Code 11 (USD-8)
- CODABAR

Return Check Digit

When selected the error check digit should be returned as part of the barcode data.

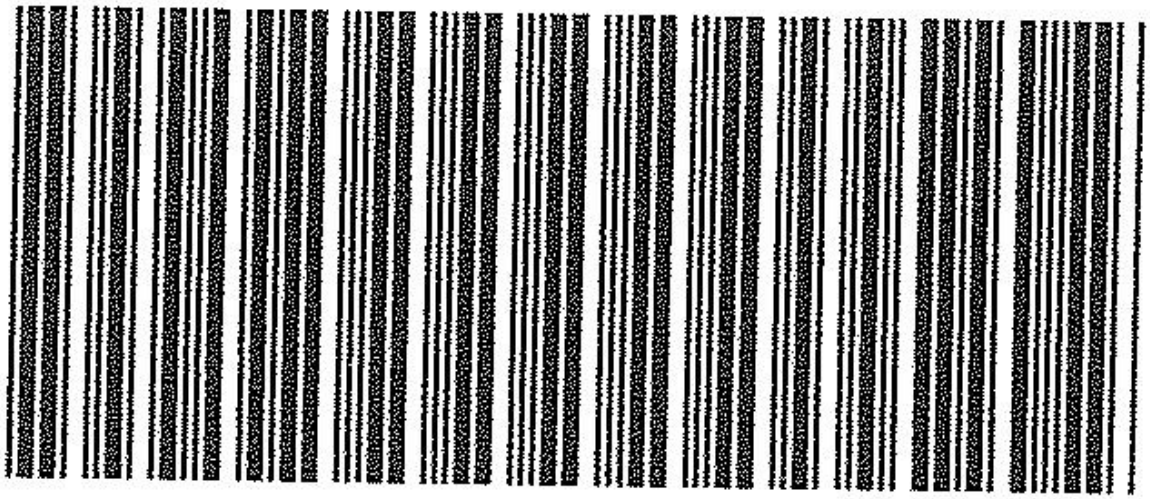
Allow partial reading

If this option is selected then barcodes that could not be read completely such as when error checking fails or when a certain area of the barcode is corrupted, the engine will return the barcode data that it was able to read. There is no guarantee what part of the barcode is read correctly and hence it is not recommended that you enable this option.

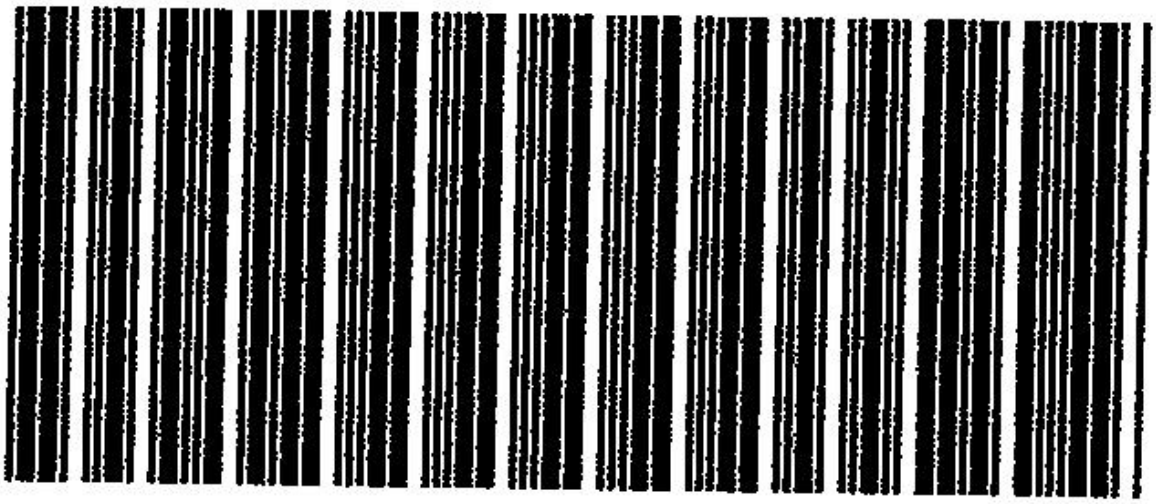
Calculate minimums

Dilates dark objects by the specified amount. The effect can be controlled by specifying the size of the neighborhood that is used for calculating each pixel value. For example for 2 x 2, set the "Minimum dimension" property to 2. Each pixel is replaced with the minimum value of its neighborhood. A black pixel has a numerical value of 0 while a white pixel has a value of 255. The calculate minimums function will therefore have the effect of filling in blank areas as shown below:

Before minimums calculation:



After minimums calculation:



This command supports 12 and 16 bit grayscale and 48 and 64 bit color images. 32 bit grayscale images are not supported.

Use Colors

When barcodes are printed in color the bar color and space color can be specified to improve recognition accuracy.

Bar Color	The color of the barcode lines.
Space Color	The color of the spaces between the barcode lines.

8.5.2.2.2 Barcode Ordering

The barcode engine sorts the barcodes that it has read from top to bottom then left to right - **on a per symbology basis**. So, when you configure barcodes make sure that the order in which the symbologies appear on the ScannerVision Bacode Settings screen is the same as what the barcode engine would return.

If you have barcodes as is shown below:

the barcode engine would number them as indicated. You can configure the EAN-13 symbologies first and the EAN-8 symbologies after that or even in between the EAN-8 symbologies as long as EAN-8 and QR symbologies respectively are in the correct order.

8.5.2.2.3 Troubleshooting

When ScannerVision does not recognize barcodes you can look at:

Source document quality

The quality of the document that is processed has a big impact on the accuracy of the barcode recognition. It is recommended that documents have a minimum resolution of 300dpi.

Auto binarize

When documents are not black and white the auto binarize function could improve accuracy.

Calculate minimums

When barcodes are dithered, i.e. lots of white areas within the bars of the barcode the calculate minimums functions can be applied to fill in the gaps.

PDF Load Resolution

When your source document is a PDF try to increase the PDF load resolution in the Template General Settings section.

Enable Partial Read

Sometimes barcodes cannot be read completely for example when the error check fails or when certain areas of the barcode has been corrupted. By enabling the partial read option, you will get back the data that could be read.

Duplicate Barcodes

The image above is of a poor quality barcode. The barcode engine erroneously thinks that there are four different barcodes here. You could tweak the advanced barcode settings to try and achieve

better results and you may in fact be successful with a given sample document but there can be no guarantee that these settings will yield good results with all documents.

The only real solution to this problem is to ensure that your source documents are of higher quality.

8.5.2.3 Zone OCR

On the Zone Recognition Settings screen you define regions known as zones in which OCRing or barcode reading has to be done. Only the area of the page that falls within the boundary of the zone is processed. ScannerVision offers four different types of OCRing that can be performed as well as barcode reading. The OCR options are:

OCR Optical Character Recognition - the recognition of regular text.

ICR Intelligent Character Recognition - the recognition of hand written text.

OMR Optical Mark Recognition - the recognition of tick marks, X's, lines, check marks, and scribbles commonly found on surveys, polls, academic exams and official applications.

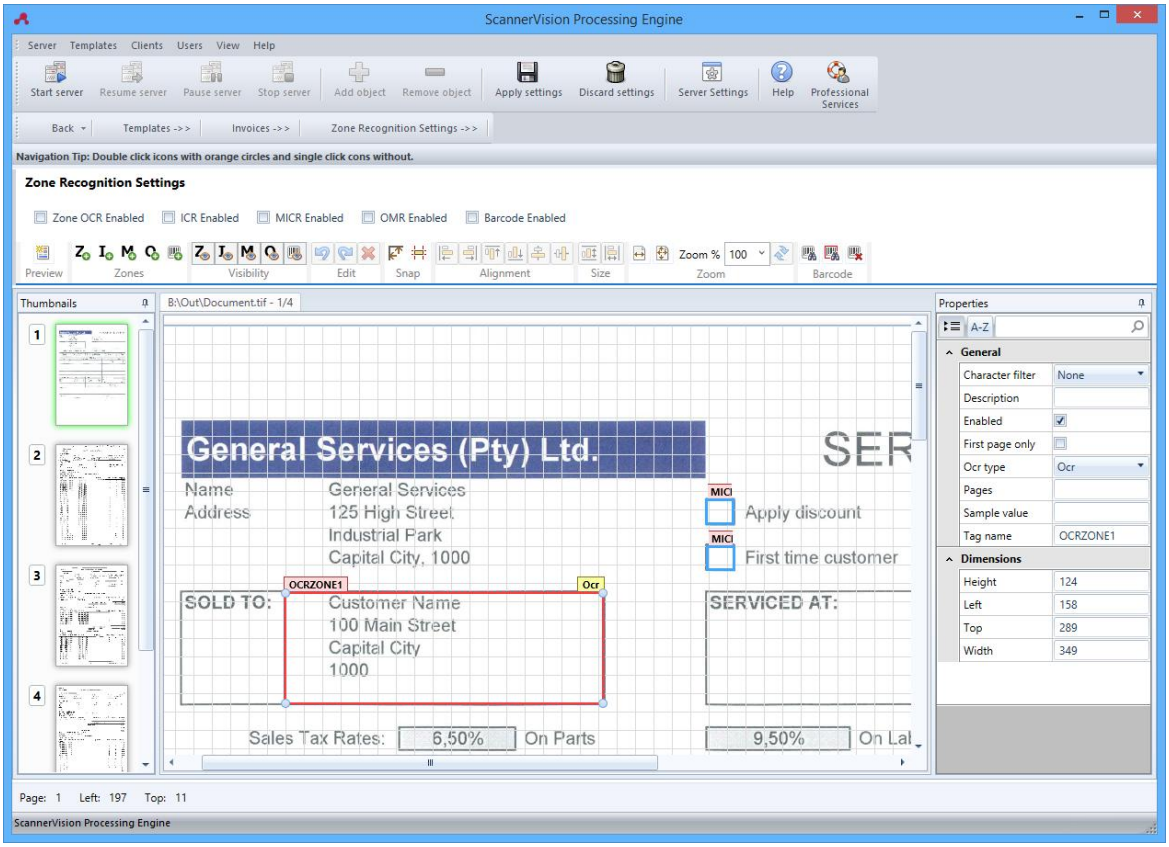
MICR Magnetic Ink Character Recognition(MICR) - the recognition of special numbers and symbols typically found on checks.

Metadata tag names

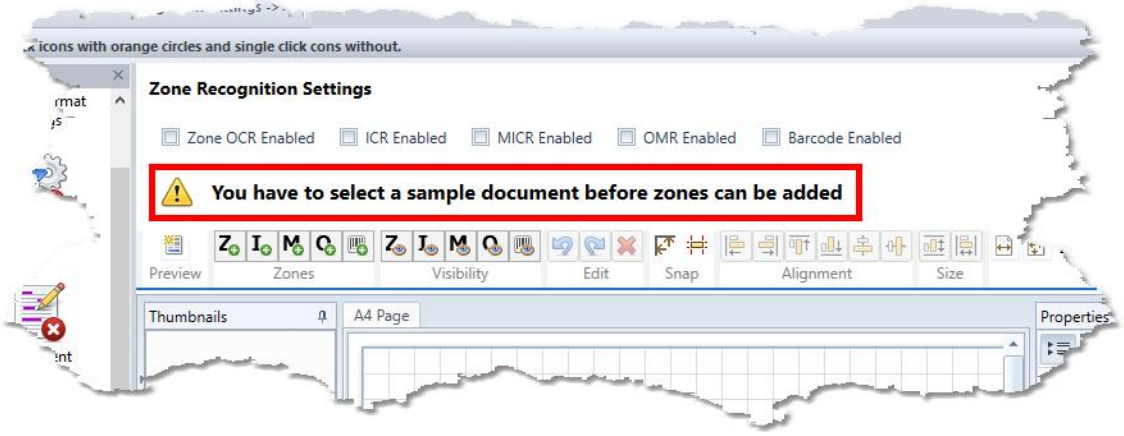
All five types of zones can be defined on the same page and can be enabled and/or disabled individually. Every zone that you configure must be given a unique name which is the name by which you would refer to the data that was retrieved from that zone. Zones can be configured to be processed on multiple pages in which case ScannerVision appends the page on which the zone was processed to the name you specify in the form of NAME + Page. If for example your zone name is CUSTOMERADDRESS and you have configured the zone to be processed on page 1 and 4, the tag names that you would reference would be CUSTOMERADDRESS1 and CUSTOMERADDRESS4. There will also be a tag name without a number which will hold the value of the zone that was read on the last page so using the example above CUSTOMERADDRESS would hold the same value as CUSTOMERADDRESS4. You can configure a zone to be processed on any number of pages e.g. "1, 2, 5, 10-12". The pages on which the zone would be processed would be pages 1, 2, 5, 10, 11 and 12.

A zone allows you to provide a description and a sample value for use in the ScannerVision Expression Editor. The description you enter will appear as a tool tip when the mouse is hovered over the tag name and the sample value is used as the tag's value.

The Zone Editor is shown below.



Zones cannot be configured without a sample document being attached. If you don't have a sample document attached, you will see a warning at the top of the zone editor window as shown below:



Enabling and disabling Zone OCR

You can enable or disable the 5 types of zone processing individually by selecting the respective check box namely "Zone OCR Enabled", "OMR Enabled", "MICR Enabled", "OMR Enabled" and "Barcode

Enabled".

Opening a sample document

Press the "Open sample document" button in the "Preview" group shown above to open a sample document. It stands to reason that the sample document you open be representative of the documents that would be processed by the template.

Adding zones

In order to add a zone to the page select the zone type by pressing the respective button in the "Zones" group shown above. The letters represent the type of zone in the following way:

Z Zone OCR

I ICR

M MICR

O OMR

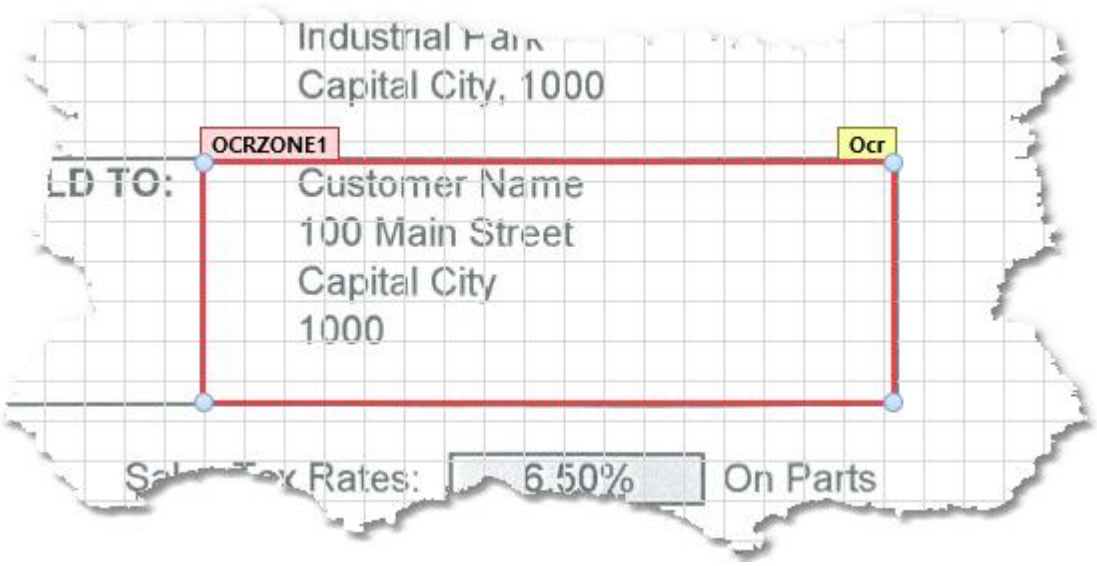
 Barcode

If you want to place a zone in which you want regular text to be OCRed, click the "Z" button. The toolbar will now look like this:

You can now draw a zone by left clicking in the zone editor window to define the top left corner and dragging the zone's width and height while the left mouse button is pressed. If you just click and release the left mouse button in the zone editor, a zone is drawn that is 50 pixels high and 150 pixels wide. After you have placed the zone, you will notice that the selected zone type is reset. If you want to place multiple zones of the same type, press the Control key (Ctrl) on the keyboard while you select the zone type. To stop placing zones click the selected zone type button again.

Re-sizing and repositioning zones

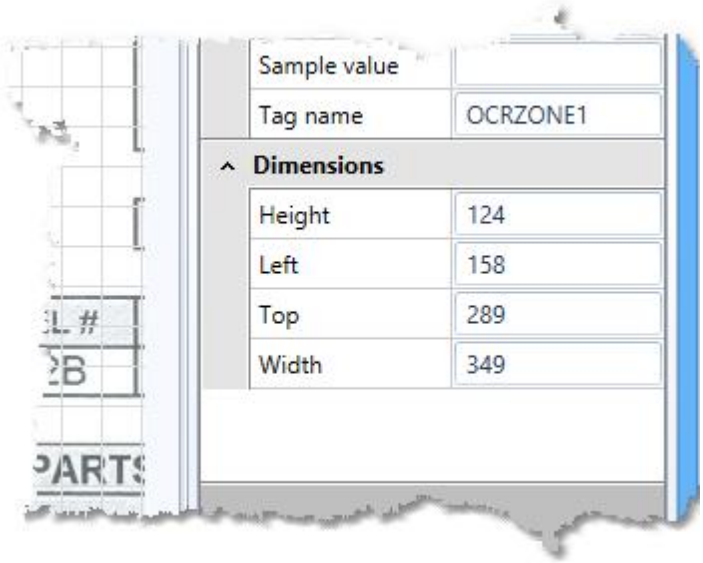
Zones can be re-sized and repositioned using the mouse or by modifying the coordinates in the Property Editor. To re-size or reposition a zone select it. You will notice that the zone changes color from blue to red and that sizing handles (blue circles) appear in the corners:



To re-size the zone drag the blue circle which you want to move to the desired location.

To reposition the entire zone, left click anywhere in the zone with the mouse and while the left mouse button is pressed drag the zone to the desired position.

You could also type in new values for the Height, Width, Top and Left positions in the property editor.



Barcode zones are allowed to have a negative "Top" and "Left" coordinate. You would need this in

situation where barcodes are placed close to the top and/or left edge of the page. See Zonal Barcodes for more details.

Changing the visibility of zones

You can show or hide zone types. Changing the visibility of zones has no influence on the processing of the zone by the ScannerVision Processing Server so hidden zones will still be processed if they are enabled.

The visibility buttons in the "Visibility" group toggles the visibility of zone types. At least one zone of a particular type has to be placed on any of the pages for its respective visibility button to become available. In the screen shot above you can see that the "Z" and "M" buttons are enabled while the "I" and "O" buttons are not. To disable for example all the Zone OCR zones, click the "Z" button. The open eye icon overlay changes to an eye with a red line through it as shown below:

Snapping

The zone editor supports two snapping modes namely "Snap to grid" and "Snap to zone".

"Snap to grid" (left icon above) forces the boundaries of the shape to fall on the grid lines that are visible over the sample document.

"Snap to zone" (right icon above) locks on to or sticks to the boundary of an existing zone that is in close proximity of the zone being placed.

Undo, Redo and Delete

Some actions of the zone editor can be undone and actions that have been undone can be redone. Whenever an action is performed that can be undone the "Undo" button (left facing arrow button on the left shown above) is enabled. If an action has been undone the "Redo" button (right facing arrow button in the middle) is enabled. Actions that can be undone and/or redone are deletion of a zone

and all the alignment actions. The make-same-size actions in the "Size" group on the toolbar cannot be undone but manual sizing actions performed on individual zones can.

Alignment

When multiple zones are selected they can be aligned in one of six ways:

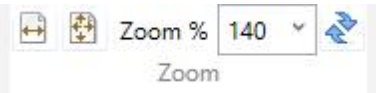
1. Left border
2. Right border
3. Top border
4. Bottom border
5. Vertical centers
6. Horizontal centers

When multiple zones are selected alignment is done relative to the first zone that was selected which will be shown in red while the remaining selected zones are green.

Sizing Zones

When multiple zones are selected their width and height can be made the same as the zone that was selected first - which will be shown in red while the remaining selected zones are green.

Zooming



The zone editor working area can be zoomed in one of four ways:

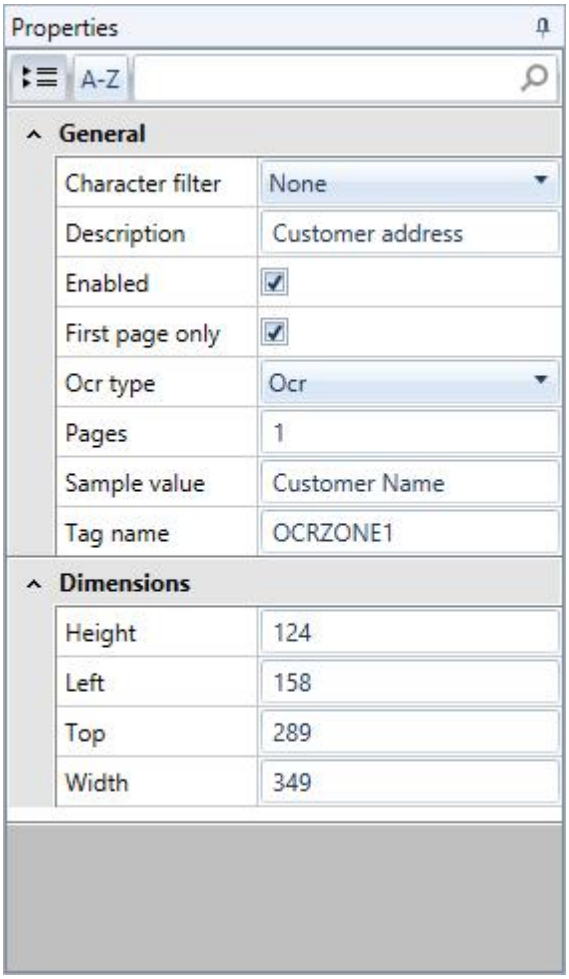
1. Fit the width of the page to the visible editor area.
2. Fit the longest edge (typically height) to the visible editor area. This is equivalent to fitting the whole page.
3. Select a custom zoom factor from the drop down list or by typing in a zoom factor.

4. While pressing the control key on the keyboard zoom in by scrolling the mouse wheel upwards and zoom out by scrolling down.

To reset the zoom factor press the "Reset zoom" button to the right of the drop down list box.

Zone Properties

Zones have properties that can be configured in the property editor shown below:



Character filter

Specifies the type of character filter to apply. A character filter allows only the selected type of character to be recognized. For example, if you set the filter to "Digit" only numeric characters will be recognized.

Character filter options are:

Alpha Upper and lowercase letters only. This is a combination of "Uppercase" and

	"Lowercase".
Digit	Recognition of numerals only. For example: "3" (Digit Three).
Lowercase	Recognition of lowercase letters only including accented ones. For example: "a" (Lowercase a).
Miscellaneous	Recognition of miscellaneous characters only. For example: "+" (Plus sign).
None	All characters are recognized.
Punctuation	Recognition of punctuation signs only. For example: "!" (Exclamation Mark).
Uppercase	Recognition of uppercase letters only, including accented ones. For example: "A" (Capital A).

Description

A description of the data that the zone represents. This description is shown as a tool tip in the ScannerVision Expression Editor when the mouse is hovered over the metadata tag name.

Enabled

Enables/disables the zone. If a zone is disabled it will be OCRed by the ScannerVision Processing Server and is drawn in a light gray color in the zone editor.

First page only

When this option is selected the zone is OCRed on the first page of the document only, regardless of what the Pages property specifies.

Ocr type

The type of OCRing that would be performed by the zone. You can change the OCR type regardless of what its original type was. If you select the Zone OCR ("Z" button) and place a zone, you can change it to ICR by selecting Icr from the list without needing to delete the Zone OCR zone and re-adding an ICR zone.

Pages

Specifies the pages on which the zone must be OCRed. Distinct pages as well as page ranges are supported e.g. "1, 2, 5-10". You can also add round brackets to ranges if it makes it easier for you to read e.g. "1, 2, (5-10)".

Note

If any form of document splitting is enabled the pages you specify here are relative to the split document and not the original document unless no split occurred.

Sample value

A value that is typical of the data contained in the zone. This value is assigned as the metadata tag value in the ScannerVision Expression Editor.

Tag name

The name of the metadata tag to which the OCRed data must be assigned.

Height

The height of the zone in pixels.

Left

The left position of the zone in pixels.

Top

The top position of the zone in pixels.

Width

The width of the zone in pixels.

Barcode Zones

Please refer to the Zonal Barcodes section for barcode specific properties.

8.5.2.3.1 Zonal Barcodes

Zonal barcode reading offers an alternative way of reading barcodes to the full page barcode reading function described in Reading Barcodes. Zonal barcode reading could in some situations yield a higher success rate. Since a zone defines a region on a page where you expect to find a barcode, the barcode reading engine uses additional techniques to try and recognize and read a barcode if it was unsuccessful after the first pass. Applying these techniques to a whole page is not practical in general because of the associated performance overhead.

After you have loaded a sample document that contains the barcodes you want to read you can perform 2 auto detect functions. You will find the button for these to the right of the toolbar, shown below.

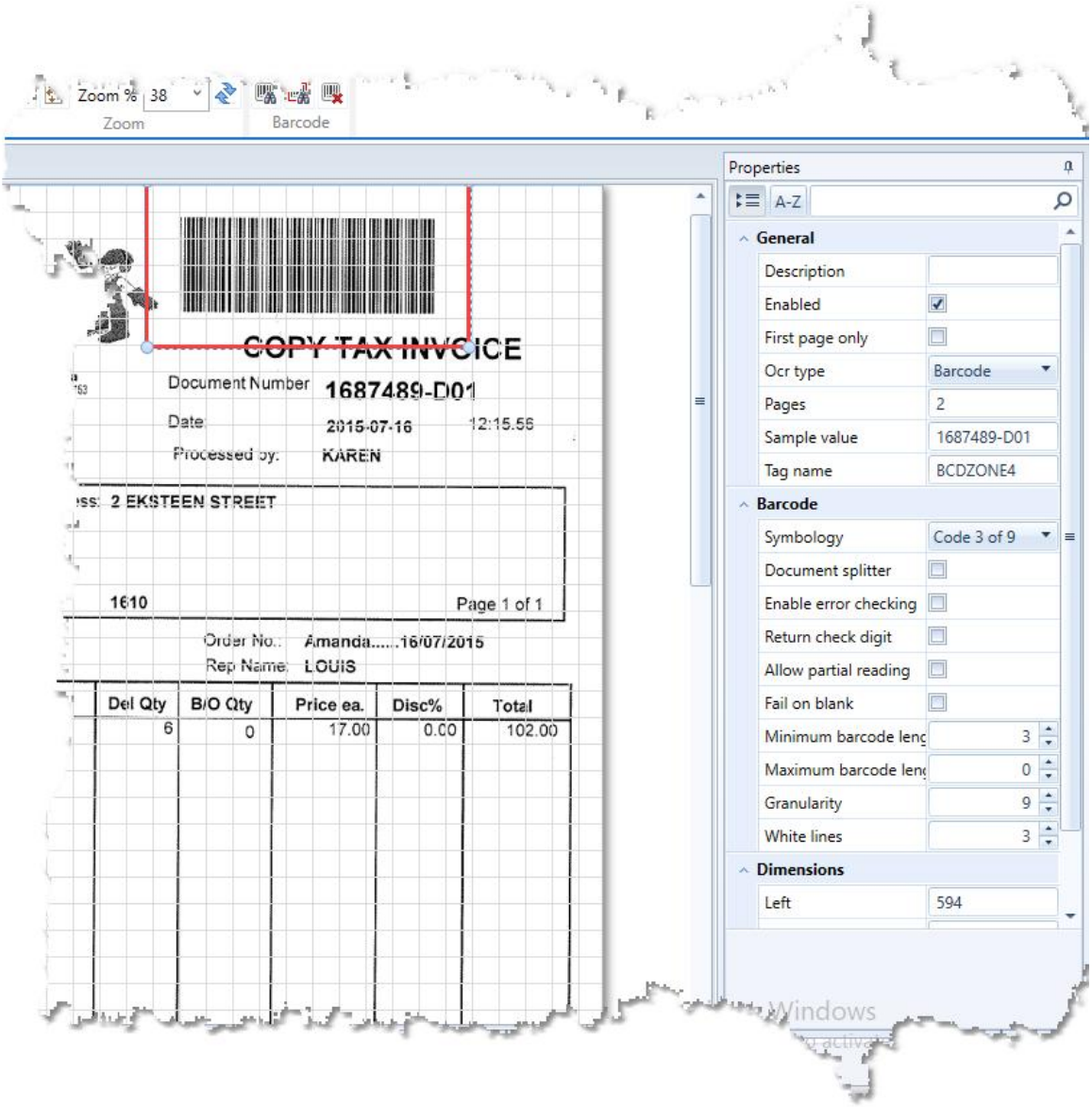
The left button performs auto detection across the entire page that is displayed. A zone is drawn around each barcode that is detected. When a zone is selected, the properties of the detected barcode is updated in the property editor.

The right button performs auto detection within a pre-existing zone. With this function you draw a zone around the barcode that you see on the page and then click this button. ScannerVision will then proceed to detect the barcode in the zone. If a barcode is detected it's properties are updated in the property editor.

Zone Size

You will notice that the zone that is drawn around an auto detected barcode appears to be much bigger than what you would have expected. An example of this is shown below. The reason for this is to make sure that the barcode engine covers a large enough area within which to perform detection. As mentioned above, the properties of the detected barcode are completed automatically. This includes the coordinates of the zone. By making the zone a bit larger you prevent the scenario where barcodes are missed during processing because of drift. In other words, if a barcode does not always appear in exactly the same place it would still be detected. This is also the reason why the zone shown below extends beyond the top of the screen with a negative top coordinate of -4 (not visible in the screen shot). This is not an error and processing of the zone will succeed.

It is recommended that when you draw zones manually you do the same.



The section which follows highlights properties in the property editor that are not present for Zone OCR zones. In addition, properties that overlap with the **Configuring Barcodes** section are not repeated. Please refer to the **Configuring Barcodes** section for details.

Pages

Specifies the pages on which the zone must be processed. Distinct pages as well as page ranges are supported e.g. "1, 2, 5-10". You can also add round brackets to ranges if it makes it easier for you to read e.g. "1, 2, (5-10)". If no page number is provided the zone is processed on every page.

If a zone is marked to be a document splitter you cannot specify page numbers.

Note

If any form of document splitting is enabled the pages you specify here are relative to the split document and not the original document unless no split occurred.

Document splitter

This allows you to perform document splitting on the page that the barcode is detected which will be the first page of the newly created split document. If you select the "Remove page" option, the page on which the barcode is detected is omitted from the split document.

Remove page

Removes the page on which the barcode is detected from the split document. This option is only available when the "Document splitter" option is selected.

Fail on blank

When selected the template fails if a barcode cannot be detected in the specified zone. This option is only available if you have specified page ranges.

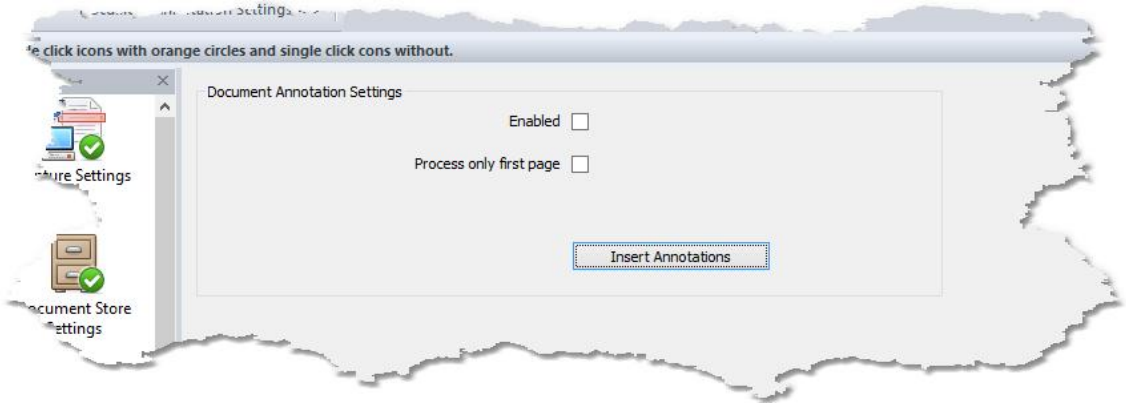
8.5.3 Adding Content

ScannerVision offers two mechanisms with which to add content to a document namely:

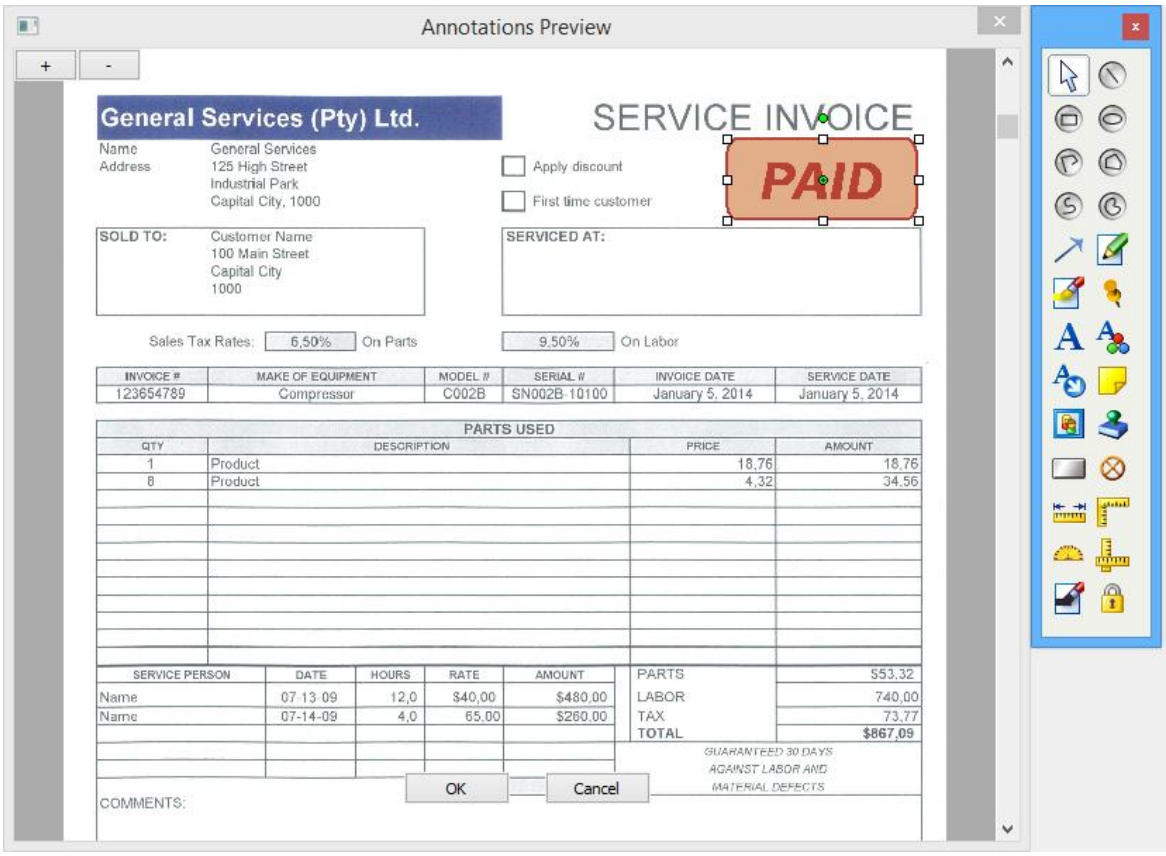
1. Annotation
2. Writing Barcodes

8.5.3.1 Annotation

Annotations can be added to a document's first page only or to all pages. The "Document Annotation Settings" screen is shown below:



To add annotations click the "Insert Annotations" button and select a sample document. The document is shown in the Annotations Preview window shown below. A floating toolbar is also shown from which you can select the type of annotation to add.



To add an annotation select the type from the floating toolbar and drag a region on the document

that should be filled by the annotation. You can re-size and reposition an annotation by selecting it to show the sizing handles (white squares). Click and drag the relevant sizing handles to the desired position. To reposition an annotation, click and drag it to the desired location.

Options that are specific to the type of annotation are accessed by right clicking on the annotation and selecting the relevant option from the context menu.

The following table lists the annotations and functions that are available on the annotation toolbar:

Icon	Description
	Selection tool. After placing an annotation this option is selected automatically.
	Line
	Rectangle
	Ellipse
	Poly line
	Polygon
	Curve
	Closed curve
	Pointer
	Freehand
	Highlight
	Pushpin
	Text

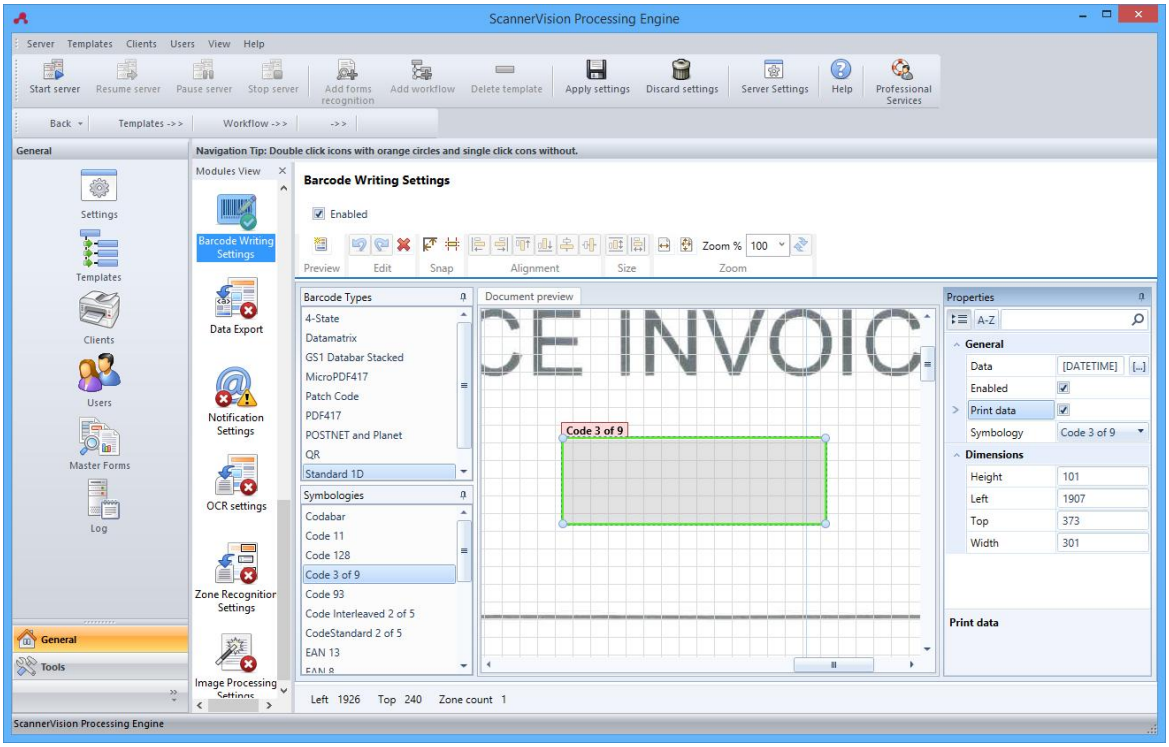
Icon	Description
	Rich Text
	Note
	Stamp
	Rubber Stamp
	Button
	Point
	Ruler
	Poly Ruler
	Protractor
	Cross Product
	Redact
	Encrypt

Note

Documents in which the "Encrypt" annotation has been applied the ScannerVision Processing Server will output an additional file with a .ann extension. This file will be uploaded by all connectors configured in the template. In order to decrypt the file with the Image Decryptor utility, you would need this file.

8.5.3.2 Writing Barcodes

ScannerVision offers the ability to write 1D and 2D barcodes on documents. To write a barcode, select the symbology from the "Modules View" and drag it onto the white area to the right. When the symbology has been dropped its settings screen appears as shown below. To configure an existing barcode, right click on its entry in the list and select the "Configure Barcode Rule" option from the context menu.



Barcode Type

The type of the barcode to write. The position and size of the barcode is modified by by pressing the "..." button to the right of the LTRB edit boxes. This allows you to select a sample document to use as reference when positioning and sizing the barcode.

Barcode Value

This is the data which will be written to the barcode. You can select metadata by pressing the "..." button to the right of the edit box. This will open the ScannerVision Expression Editor.

Print barcode value

A human readable annotation of the barcode value will be placed below the barcode if this option is selected.

Alignment

The barcode can be aligned within the region specified above by selecting Left, Center or Right.

Use Colors

Barcodes can be printed in color. Bar Color and Space Color must be specified.

Bar Color	The color to be used when solid bar colors are written (Normally the BLACK areas of a barcode)
Space Color	The color to be used when white spaces should be written (Normally the WHITE areas of a barcode)

8.5.4 Exporting Data

All metadata that is accompanies an incoming document or that is created by a template can be exported to any of 3 formats namely Text, XML and HTML. To export metadata, select the output type from the Modules View and drag it onto the white area to the right. The export settings screen will appear. To edit previously configured export rules, right click on the respective rule and select "Configure rule" from the context menu.

Below are excerpts of the output produced by the respective formats:

TXT

UTC=2014-06-23T04:48:52.4461958Z

UCT=2014-06-23T04:48:52Z

TZD=+02:00

DATETIME=2014-06-23 06:48:52

DATE=2014-06-23

TIME=06:48 AM

...

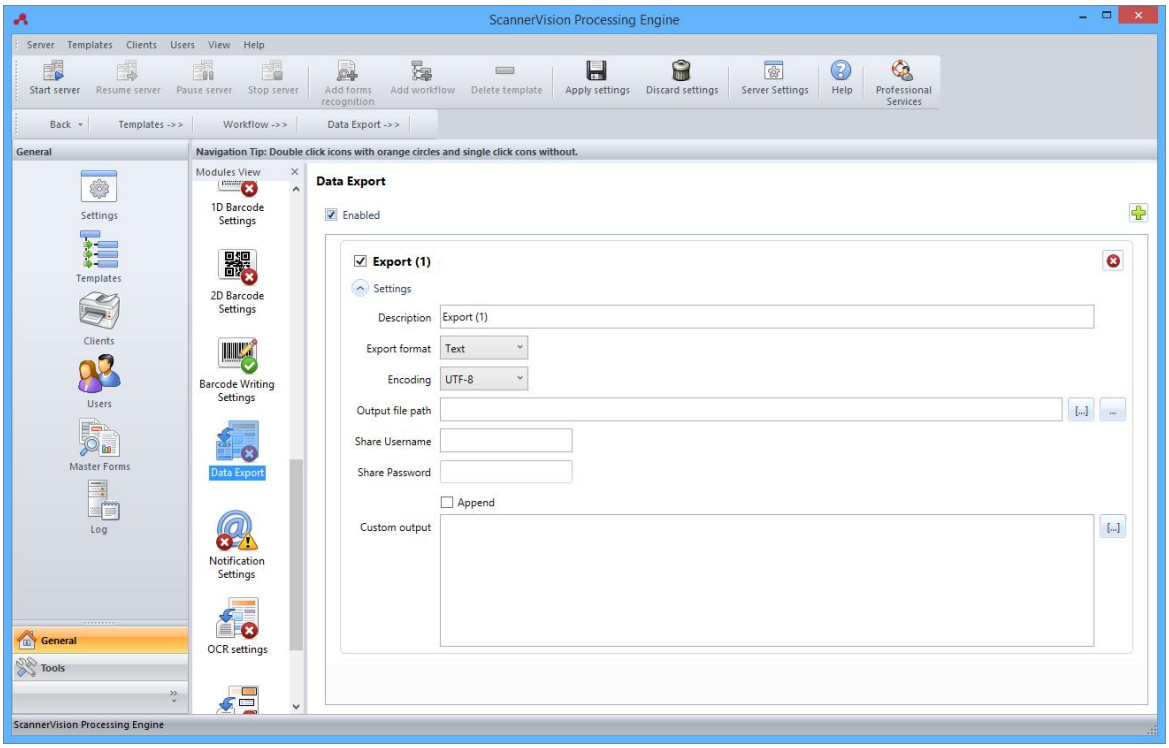
XML

```
<metadata>
	<UTC>
		<![CDATA[2014-06-23T04:48:52.4461958Z]]>
	</UTC>
	<UCT>
		<![CDATA[2014-06-23T04:48:52Z]]>
	</UCT>
	<TZD>
		<![CDATA[+02:00]]>
	</TZD>
	<DATETIME>
		<![CDATA[2014-06-23 06:48:52]]>
	</DATETIME>
	<DATE>
		<![CDATA[2014-06-23]]>
	</DATE>
	<TIME>
		<![CDATA[06:48 AM]]>
	</TIME>
	...
</metadata>
```

HTML

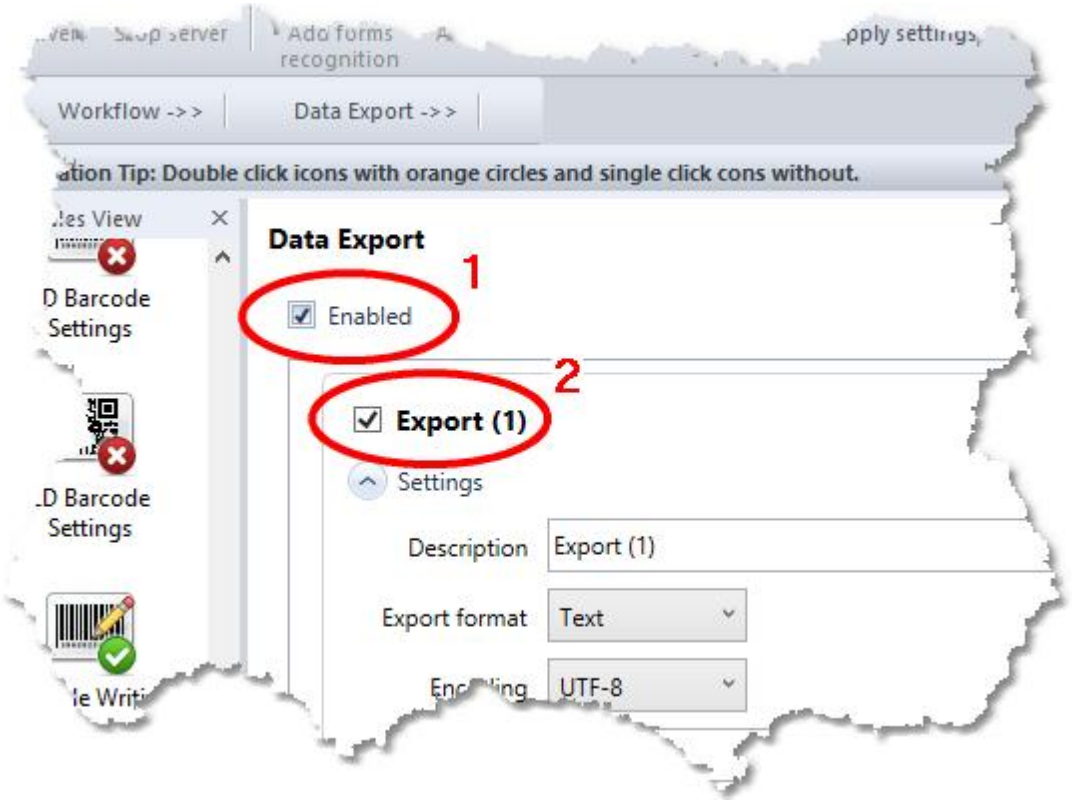
```
<html>
```

```
<head>
        <title>METADATA</title>
</head>
<body>UTC=2014-06-23T04:48:52.4461958Z<br />UCT=2014-06-23T04:48:52Z<br />TZD=+02:00<br />DATETIME=2014-06-23 06:48:52<br />DATE=2014-06-23<br />TIME=06:48 AM...
</body>
</html>
```



Enable/Disable exports

To disable all exports disable the "Enabled" check box circled and marked as 1 below. To disable an individual export disable the check box circled and marked as 2 below.



Description

A description of the export. The value you enter here is for your own use and has no effect on the functioning of the export.

Export format

The output format of the exported data. Possible values are:

TXT

Output specification:

Document processed: [ORIGINALFILENAME][#13]
Processed date & time: [DATETIME][#13]
Server version: [SVSERVERVERSION]

Result:

Document processed: Document.tif
Processed date & time: 2014-06-23 07:28:43

Server version: 6.0.0.200

XML

Output specification:

```
<metadata>[#13]
    <DocumentProcessed>[ORIGINALFILENAME]</DocumentProcessed>[#13]
    <ProcessedDateTime>[DATETIME]</ProcessedDateTime>[#13]
    <ServerVersion>[SVSERVERVERSION]</ServerVersion>[#13]
</metadata>
```

Result:

```
<metadata>
    <DocumentProcessed>Document.tif</DocumentProcessed>
    <ProcessedDateTime>2014-06-23 07:28:43</ProcessedDateTime>
    <ServerVersion>6.0.0.200</ServerVersion>
</metadata>
```

HTML

Output specification:

```
<html>[#13]
    <head>[#13]
            <title>METADATA</title>[#13]
    </head>[#13]
    <body>[#13]
            <h1>Document processed</h1>[#13]
            [ORIGINALFILENAME][#13]
            <h1>Processed date & time</h1>[#13]
            [DATETIME][#13]
            <h1>Server version</h1>[#13]
            [SVSERVERVERSION][#13]
    </body>[#13]
</html>
```

Result:

```
<html>
    <head>
            <title>METADATA</title>
```

```
    </head>
    <body>
        <h1>Document processed</h1>
        Document.tif
        <h1>Processed date & time</h1>
        2014-06-23 07:28:43
        <h1>Server version</h1>
        6.0.0.200
    </body>
</html>
```

Encoding

Select the character encoding to be used for the output file. Possible values are:

Text	Output is regular ASCII with no byte order mark.
UNICODE	Output is Unicode and a byte order mark is added.
UTF-8	Output is Utf-8 and a byte order mark is added.

Output file path

The fully qualified path and file name of the export file. To make use of metadata click the "[...]" button to the right of the edit box. The select a file click the "..." button to the right of the edit box.

Share Username & Password

Credentials for exporting to a protected network share.

Append

Appends metadata to the export file if it exists instead of overwriting it.

Custom Output

You can override the default output format shown above by specifying which metadata tags must be exported and optionally any static text the must be included. When you want to output custom Xml and Html, you have to define the complete Xml and Html structure i.e. you have to specify all the tags explicitly. Custom output can be authored in the ScannerVision Expression Editor by clicking the "[...]" button to the right of the "Custom output" edit box.

Below are examples of custom output.

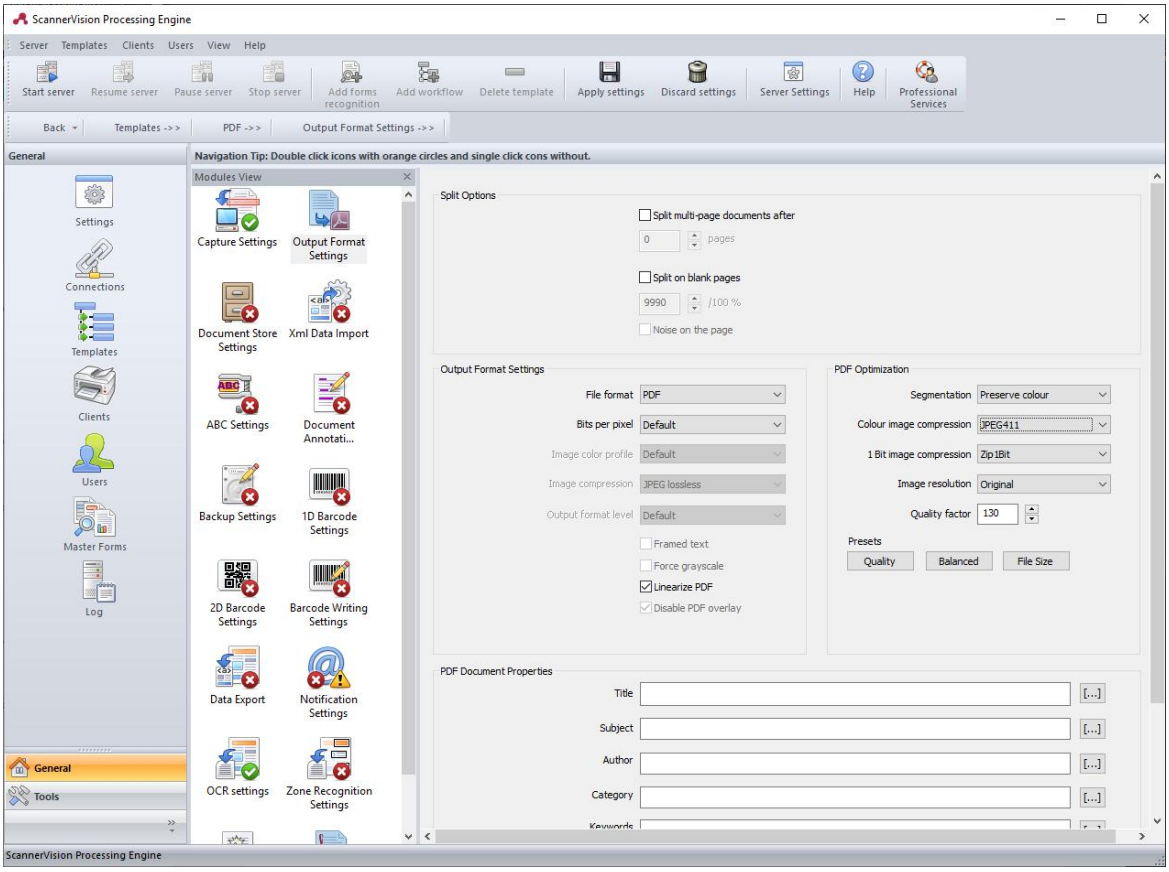
8.6 Store

The final stage in the ScannerVision document processing work flow is the storing of the processed document and optionally its metadata. You have control over which format the document is stored in and where. You can also configure backup locations of the original document and sending of notifications after processing has completed.

- Output Format Settings
- ABC Settings
- Backup Settings
- Notifications Settings
- Connector Settings

8.6.1 Output Format Settings

ScannerVision support document output to various file formats. A Tiff document scanned on a multi-functional printer could for example be converted to a searchable PDF or MS Word document. The output document properties are configured on the "Output Format Settings" screen shown below.



Split multi-page documents after

With this option enabled the output document is split into multiple documents with each containing up to the number of pages specified. The suffix "-split-00000000.tif" is added to the output file name with the zeros indicating the document number e.g. "-split-00000000.tif", "-split-00000001.tif", "-split-00000002.tif" etc.

Split on blank pages

With this option enabled the output document is split on every blank page that is detected. Please refer to the Delete Empty Pages section for more information on the accuracy and noise values. The suffix "-split-00000000.tif" is added to the output file name with the zeros indicating the document number e.g. "-split-00000000.tif", "-split-00000001.tif", "-split-00000002.tif" etc.

File format

The document output format. Document output formats fall into 3 categories namely Original, Raster and Vector formats. If you choose a vector output format you have to enable the OCR engine.

Original:

When the output format is set to original the original document is output untouched. Also, certain

controls are disabled since they have no effect in this context. The controls that remain enabled have an influence on the OCR engine but not on the output document. For example, when selecting a bits per pixel value of 1, the OCR engine will rasterize the document as 1 bit before the OCRing is performed. Similarly, when the "Force grayscale" option is enabled the document on which OCRing is performed is converted to grayscale prior to OCRing while the output document remains the original.

If the incoming document is a PDF and OCRing is enabled, the original document is overlaid over the OCRed text. With this option enabled the fidelity of the original document is retained and the size only grows with that of the OCRed text. This behavior can be disabled by selecting the "Disable PDF overlay" option.

Raster formats:

- BMP
- GIF
- JPEG
- JPEG 2000
- PCX
- PDF
- PNG
- PostScript
- TIFF
- WMF*
- XPS*

Vector formats:

- Microsoft Word 2000
- Microsoft Excel 2000
- Word ML
- RTF
- Searchable PDF
- PDF/A (Level B conformance)

* Although the WMF and XPS file formats support vector information ScannerVision only produces raster versions of these files.

Bits per pixel

The color or bit depth of the output document. If you convert to a lower bits-per-pixel format ScannerVision optimizes the colors automatically. For example, when converting a 24-bit file (16 million colors) to an 8-bit file (256 colors) ScannerVision selects the best 256 colors to represent the 24-bit image. This setting also determines the bit depth at which a page from the document is loaded into memory. This means that if you choose 1 bpp a colour document is loaded as black and white and all processing is performed on the black and white document - including output processing. The only exception to this is when PDF Optimization is enabled. This ensures that the segmentation engine has all the information in the original document to apply its algorithm to.

Image color profile

The following color profiles are available: 4:1:1:, 4:2:2:, RGB and CMYK. File type dependent.

Image compression

The following compression modes are available: No Compression, RLE, LZW, ZIPLib, CCITT, G4, JPEG 4:4:4, JPEG 4:2:2, JPEG 4:1:1, Flate and Jbig2. File type dependent.

Output format level

The output format level specifies how much formatting is retained in Microsoft Word and Excel documents. The options and their meaning are:

Default	ScannerVision automatically determines level of formatting to retain.
None	No format conversion. All formatting information is ignored and replaced by a default value. (One column, left aligned paragraphs, no font attributes, a default font, etc.)
Retain FP	Retain Font and Paragraphs. The formatting information of fonts and paragraphs is retained but layout related information is ignored. (This level has a special purpose when saving to Excel: each detected table or spreadsheet in a document is saved to a separate worksheet. Other content is placed on the last worksheet and functions as an index. The tables are replaced by hyperlinks to their own sheet.)
True Page	This keeps the look of the original layout of the pages. This is done by absolute positioning of the texts, pictures and tables on the page with boxes, frames or other target application specific methods. This level is only available for target applications capable of handling these.
Flowing Page	Preserves the original layout of the pages, including retaining columns. Boxes and frames are only used when necessary.
Spreadsheet	This level exports the results in tabular form, suitable for use in

spreadsheet applications. Each page is placed in a separate worksheet.

Framed text

This property applies to the WORD 2000 and RTF output formats. When enabled text blocks are put into frames which can be sized and moved individually.

Force grayscale

When the bits per pixel setting has a value of 8 or 16 bit and this option is selected the output document will be converted to grayscale before any operations are performed on it. If the "File format" option is set to "Original" this conversion to grayscale will not affect the output format.

Linearize PDF

A linearized PDF file is a special format of a PDF file that makes viewing faster over the Internet. Linearized PDF files contains information that allow a byte-streaming server to download the PDF file one page at a time. If byte-streaming is disabled on the server or if the PDF file is not linearized, the entire PDF file must be downloaded before it can be viewed. All supported versions of IDS produce linearized PDF files. (from What are linearized PDF files?).

This option is only available when the PDF, Searchable PDF or PDF/A file format is selected.

Disable PDF overlay

When this option is selected the original PDF document is not overlaid over OCRed text. If document OCR is not enabled this option has no effect.

PDF Document Properties

PDF documents support the inclusion of extra information in the document and include the fields listed below:

- Title
- Subject
- Author
- Category
- Keywords
- Comments

ScannerVision metadata can be inserted for these fields by clicking the "[...]".

PDF Optimization

Please refer to the PDF Optimization section for details.

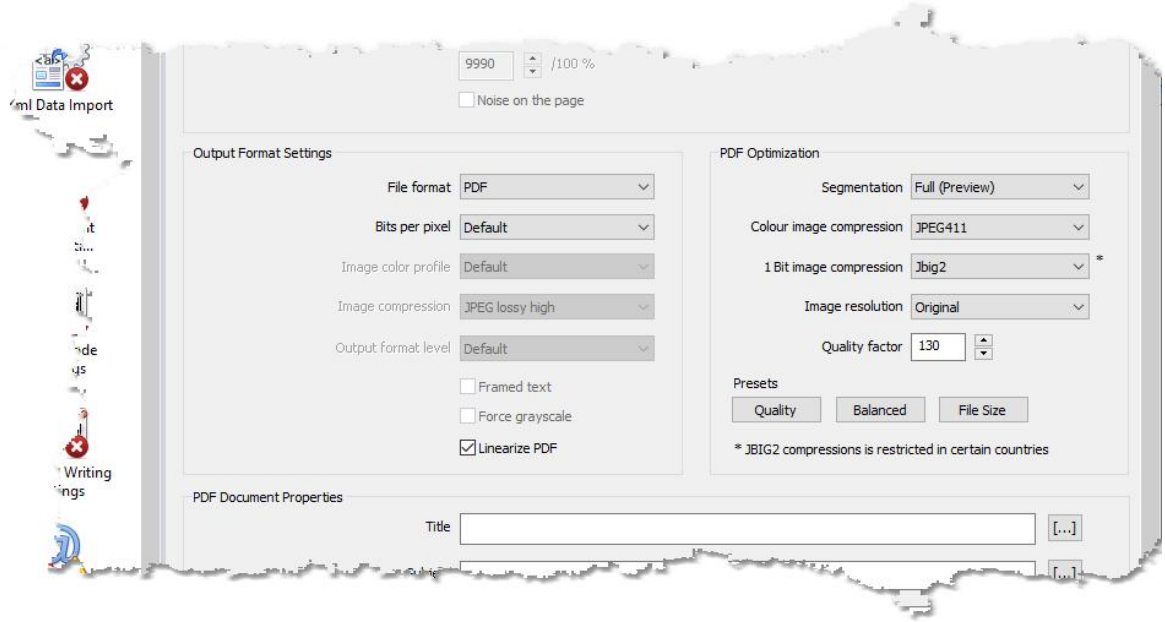
8.6.1.1 PDF Optimization

When PDF Optimization is enabled, ScannerVision applies a technique called Image Segmentation by which every page of a document is segmented into 1 bit and multi-bit (full colour) segment types. Segments are then compressed using a compression algorithm that is optimized for the type of segment. The size of the document resulting from this process could be significantly smaller than the original with no perceivable/visible quality degradation (depending on the options chosen as described below).

Image Segmentation is a computationally expensive operation that will add to the document processing time. In addition, the segmentation algorithm is driven by various input parameters - some of which are provided in the PDF Optimization section discussed below while others are pre-configured for optimal performance.

Depending on the choices you make with the optimization options below, you may encounter artifacts in the output document. For example, if you prioritize smaller file sizes you will typically choose the "Full" segmentation option. In this case you may find that the segmentation engine has segmented a region as 1 bit where it should have been colour. If, on the the other hand, document fidelity is more important to you, you would choose the "None" segmentation option but with the downside of a potentially larger output document.

Any reference to "image" in the text to follow refers to a single page in a document.



PDF Optimization settings are enabled when "PDF", "Searchable PDF" or "PDF/A" is selected as the "File format".

Segmentation

The level of segmentation that is to be applied.

None	No segmentation is performed. Each image is treated as a single segment and depending on the bit depth of the output image ("Bits per pixel" option), the appropriate compression is applied according to your "Colour image compression" and "1 Bit image compression" choices.
Preserve colour	This option forces all segments to be treated as full colour which will prevent the occurrence of color artifacts.
***Full**	Full segmentation is performed which means each image is segmented according to its content and the appropriate compression is applied according to your "Colour image compression" and "1 Bit image compression" choices.

This option is disabled when the "Bits per pixel" is set to 1. In this case, and in the case that the original image being processed is already 1 bit, there would be no benefit in segmenting the image since it is known to be 1 bit. The segmentation engine will treat the whole image as a single 1 bit segment and apply the "1 Bit image compression" algorithm that you have selected.

* This is a preview feature.

Colour image compression

The compression algorithm to apply to colour segments. Options are JPEG, LZW, ZIP, JPEG422 and JPEG411.

1 Bit image compression

The compression algorithm to apply to 1 bit segments. Options are FaxG31D, FaxG32D, FaxG4, Jbig2, Lzw1Bit, Zip1Bit.

NOTE

The use of JBIG2 compression may be prohibited in certain countries for certain types of documents.

Image resolution

This determines the output resolution of segments. If the document being processed has a higher resolution than what you specify here the document's resolution is reduced to the value specified. If it has a lower resolution however the document's original resolution is retained. The lower the resolution of the output document the smaller the file size, but the lower the quality of the document.

Quality factor

This determines how much information is thrown away by the JPEG, JPEG422 and JPEG411 compression algorithms. Jpeg is a "lossy" compression algorithm but you have control over how lossy it is. The lower the number the less information is thrown away resulting is a higher quality but larger

document.

Presets

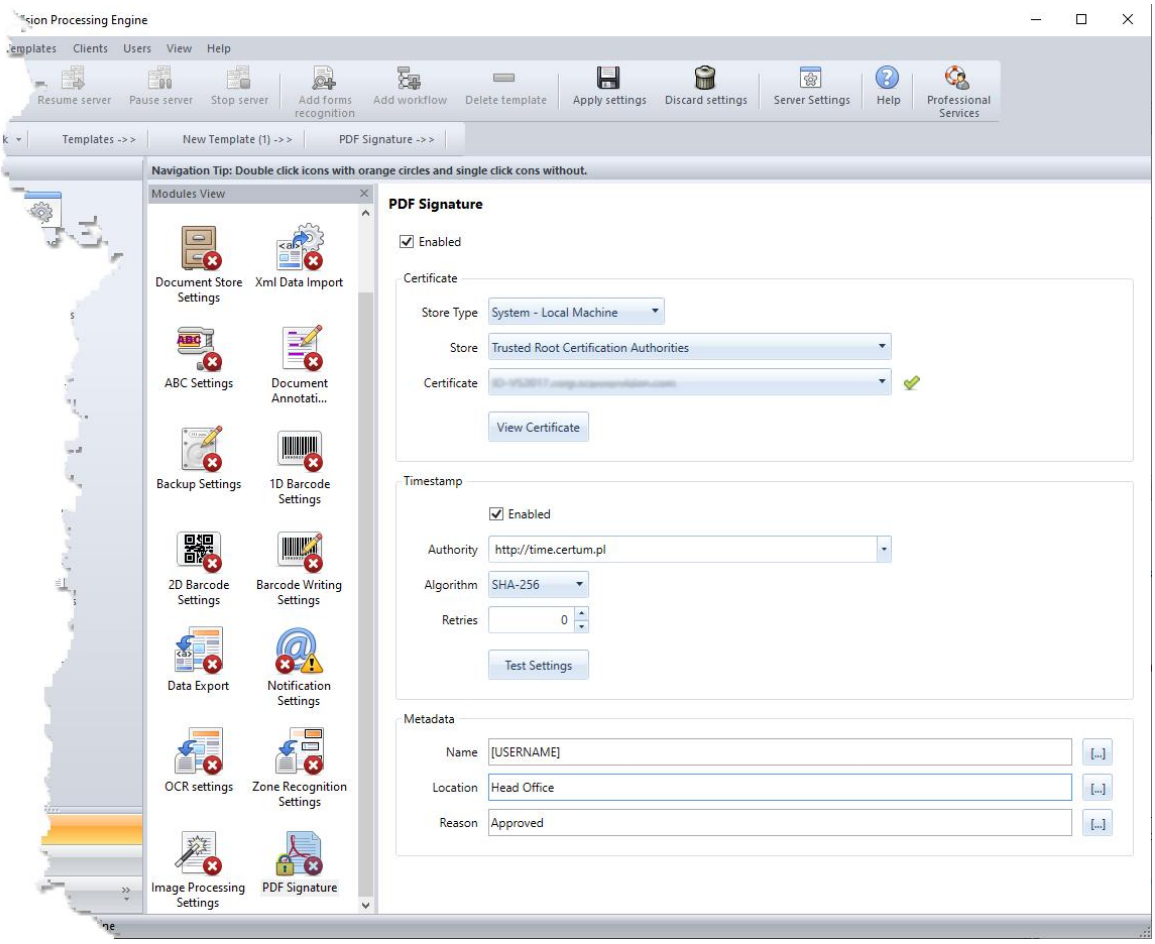
The options above provide a lot of control over the quality and size of the output document. ScannerVision offers 3 presets that provide a combination of these options suited to three types of output:

Quality	Higher quality, larger document size.
Balanced	Good quality, moderate document size.
File Size	Lower quality, smaller document size.

Once you have clicked on one of the preset buttons and the controls have been updated with the relevant options, you can customize it further to suit your needs.

8.6.1.2 PDF Signing

PDF Signing allows you to digitally sign a PDF document with certificate which guarantees the authenticity of the document. The PDF Signature screen is shown below.



Enabled

Enables or disables the PDF signing function.

Store Type

Allows you to select the location of the certificate with which the PDF will be signed. Options include:

File	You select a Personal Information Exchange (.pfx or .p12) file on disk.
System - Current User	You select an existing certificate from the current user's certificate store.
System - Local Machine	You select an existing certificate from the local system's certificate store.

In all case you have to select a certificate that was specifically created for the signing of PDF documents. For more information you can search online for "pdf digital id". If you select an invalid Personal Information Exchange file or installed certificate the "Certificate" drop down box will not list the certificate.

Password

This field is only available when the "File" store type is selected. This is the password of the Personal Information Exchange file.

View Certificate

Once you have selected a valid certificate the "View Certificate" button enables. If you click it the details of the certificate is displayed.

Timestamp

The time stamp function allows you to add a time stamp to the signed PDF document which means that the PDF signature is considered to be valid even if the certificate expires after the document was signed. The certificate has to be valid at the time the PDF document is signed though.

Enabled

Enables or disables the time stamp function.

Authority

The time stamp service to use.

Algorithm

The algorithm that is used to hash the time stamp with.

Retries

The number of retries in case an error occurs contacting the time stamp service.

Test Settings

Tests that the time stamp services can be contacted.

Metadata

The digital signature can be enriched with metadata which gives more information about the signature. The following values are provided:

Name

The name of the person or authority signing the document.

Location

The CPU host name or physical location of the signing.

Reason

The reason for the signing such as "I agree".

For more information you can see here: https://www.globalsign.com/en/blog/what-is-timestamping-how-does-it-work/

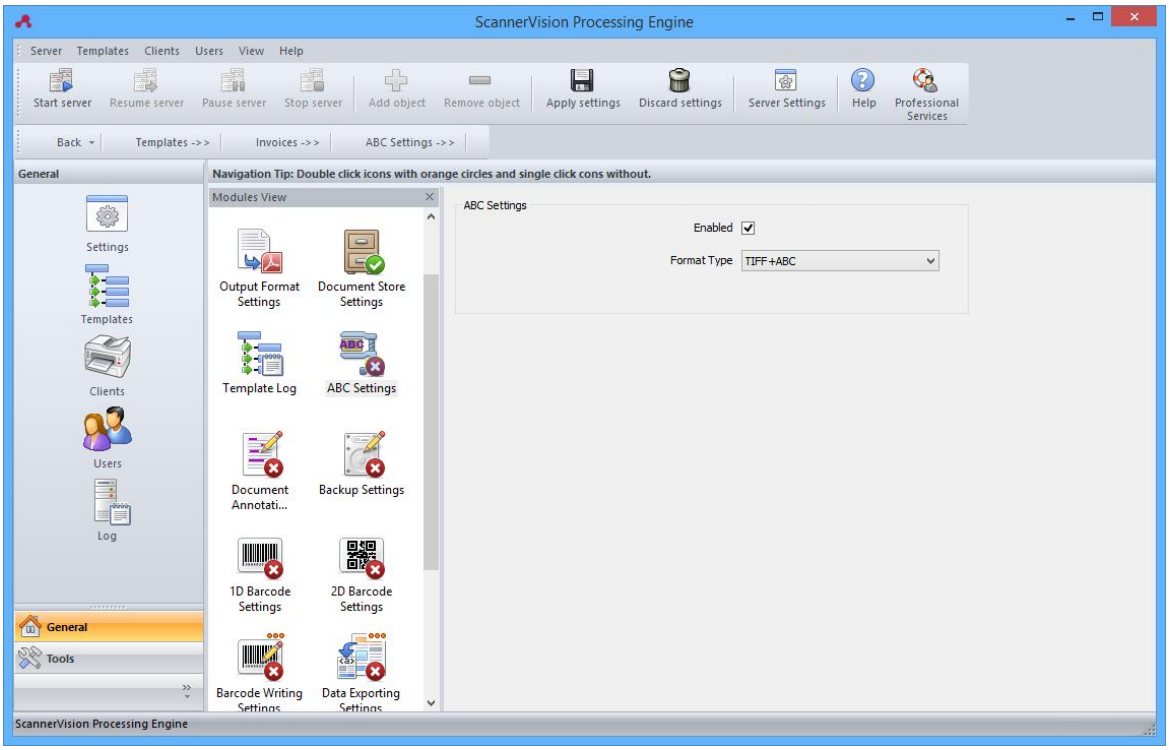
8.6.2 ABC Settings

Leadtools ABC (Advanced Bitonal Compression) is a proprietary 1 bit, lossless compression algorithm that produces file sizes that are up to 4 to 5 times smaller than the files sizes generated by standard 1 bit compression algorithms such as CCITT Group 4. ScannerVision can load ABC compressed Tiff documents as input documents as well as compress output Tiff documents using the ABC algorithm.

Note

Since ABC compression is a proprietary technology not all image viewing applications are able to read Tiff files that employ this compression algorithm. ScannerVision provides the ABC Manual Decompressor tool to de-compress ABC compressed Tiff documents. When ABC compressed Tiff documents are used as input ScannerVision can load them without the need to decompress them first.

The ABC settings screen is shown below:



Enabled

Enables/disables ABC compression on output documents.

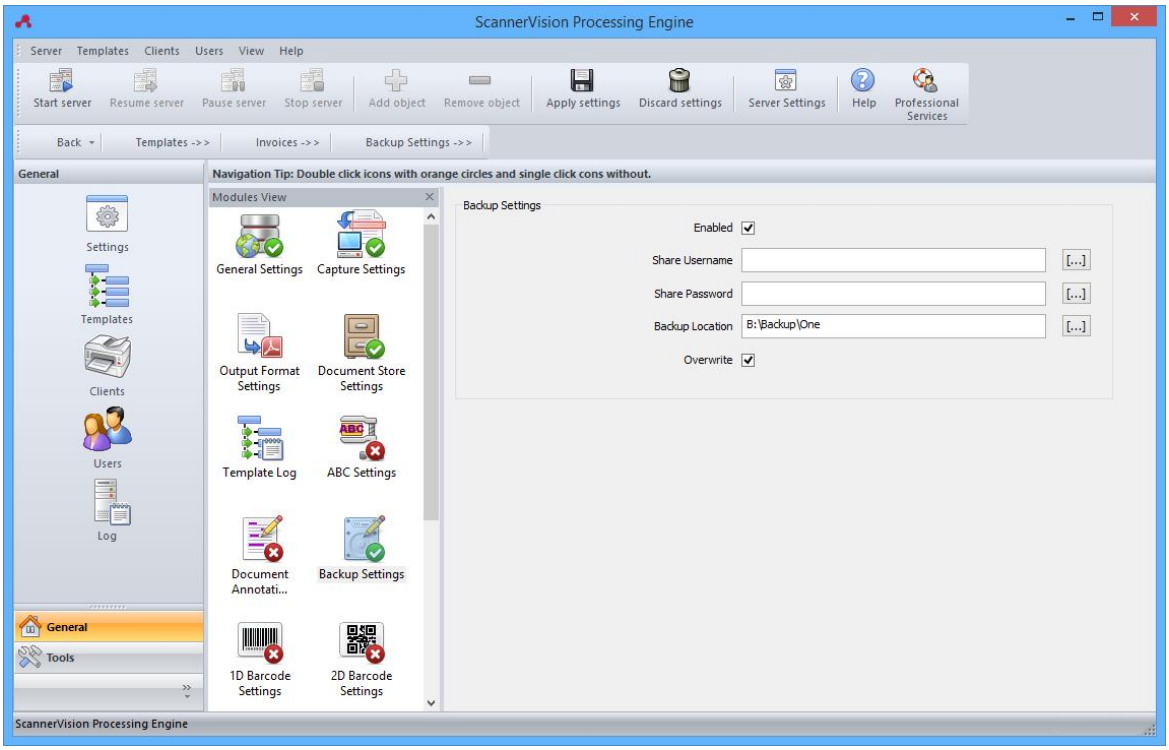
Format Type

ABC For single page files only (higher compression ratio).

ABC+TIFF For multi-page files (lower compression ratio).

8.6.3 Backup Settings

A copy of the incoming document can be stored in the location you specify on the Backup Settings screen shown below:



Enabled

Enables/disable backup.

Share Username & Password

Credentials to backup documents to a protected network share. To insert metadata click the "[...]" button to the right of the respective edit boxes.

Backup Location

The path to the backup location. To insert metadata click the "[...]" button to the right of the edit box.

Overwrite

Select this option to overwrite existing files. If this option is disabled and a file with the same name already exists in the backup location the original file is not copied.

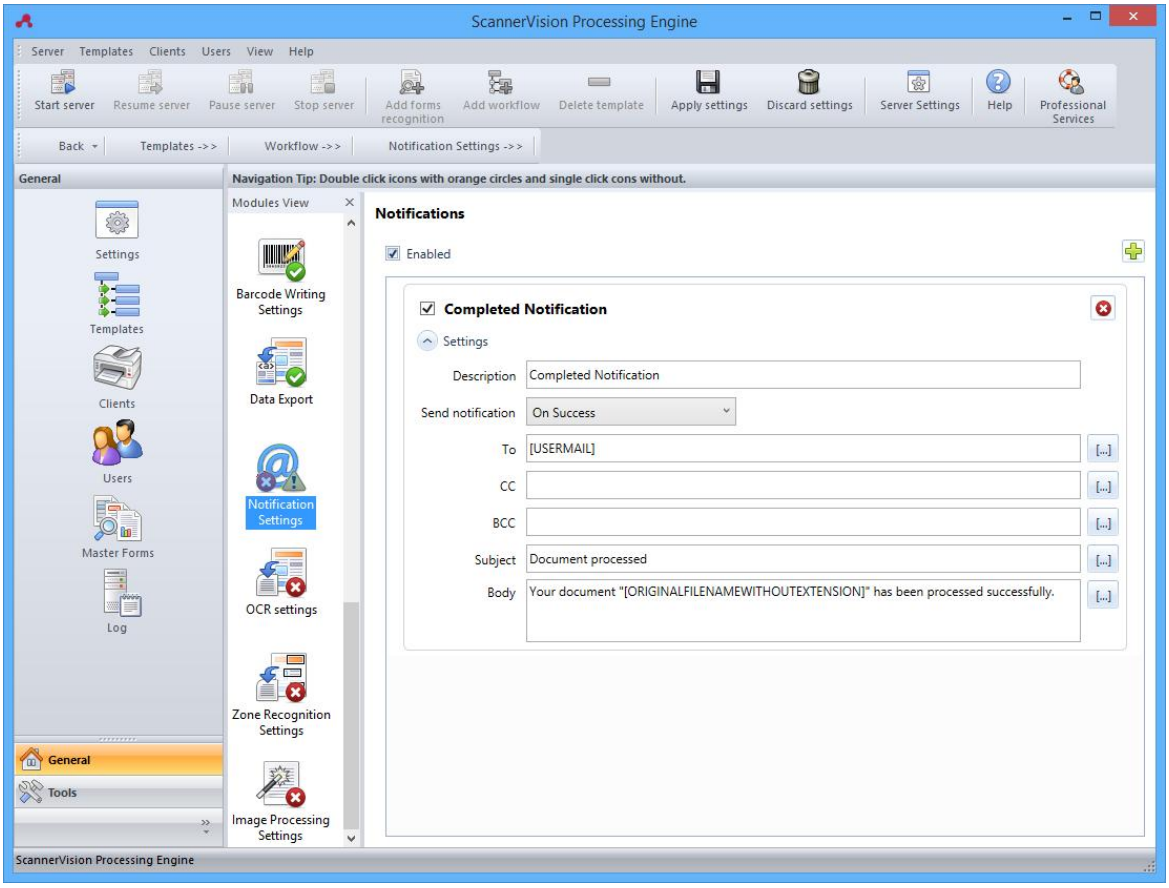
8.6.4 Notification Settings

Notification emails can be sent to any number of email addresses under following conditions:

1. Every time the template executes
2. Only when the template execution was successful
3. Only when the template execution failed

The SMTP settings that are used to send notifications can be configured in the SMTP Server Settings section.

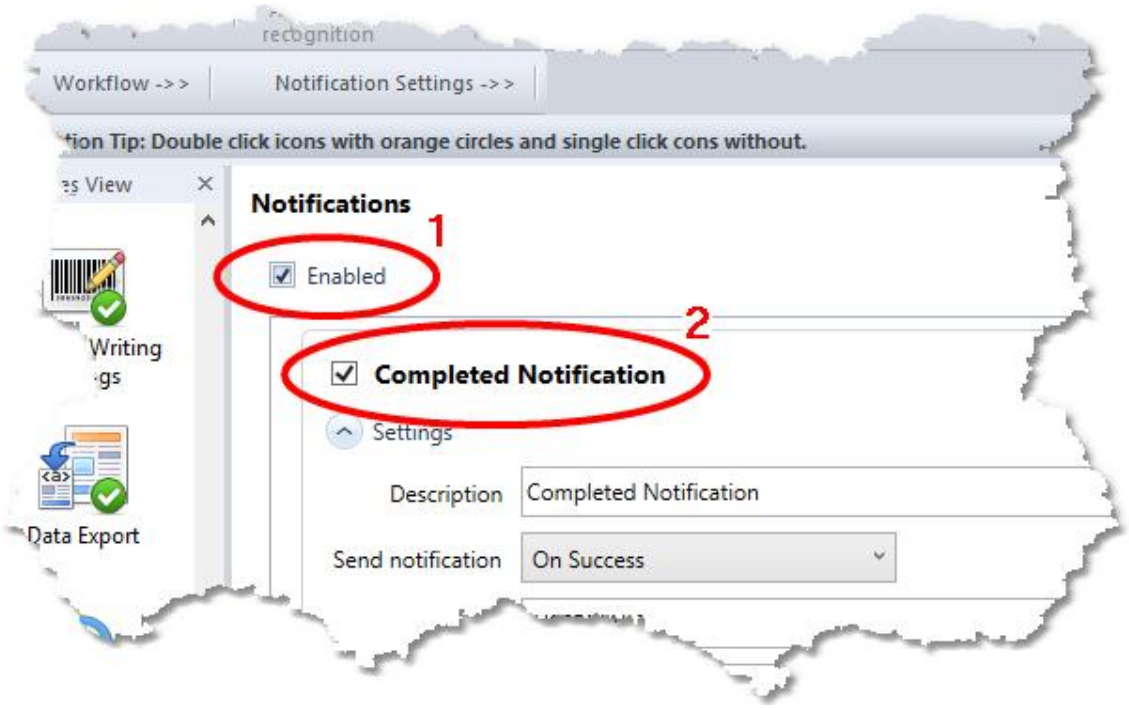
Email notification a configured on the "Notification Settings" screen shown below. To add a notification right click in the notification area and select "Add notification rule" from the context menu.



On the notification settings screen you configure the relevant properties of the email notification. The "[...]" buttons on the screen indicate that metadata can be used in the respective edit controls.

Enable/Disable notifications

To disable all notifications disable the "Enabled" check box circled and marked as 1 below. To disable an individual notification disable the check box circled and marked as 2 below.



Description

A description of the notification. The value you enter here is for your own use and has no effect on the functioning of the export.

Send notification

Select when the notification is sent. Possible values are:

Always	Every time the template executes
On Success	Only when the template execution was successful
On Failure	Only when the template execution failed

To

The recipients' email addresses separated by semicolons ";".

CC

CC recipients' email addresses separated by semicolons ";".

BCC

BCC recipients' email addresses separated by semicolons ";".

Subject

The subject of the email.

Body

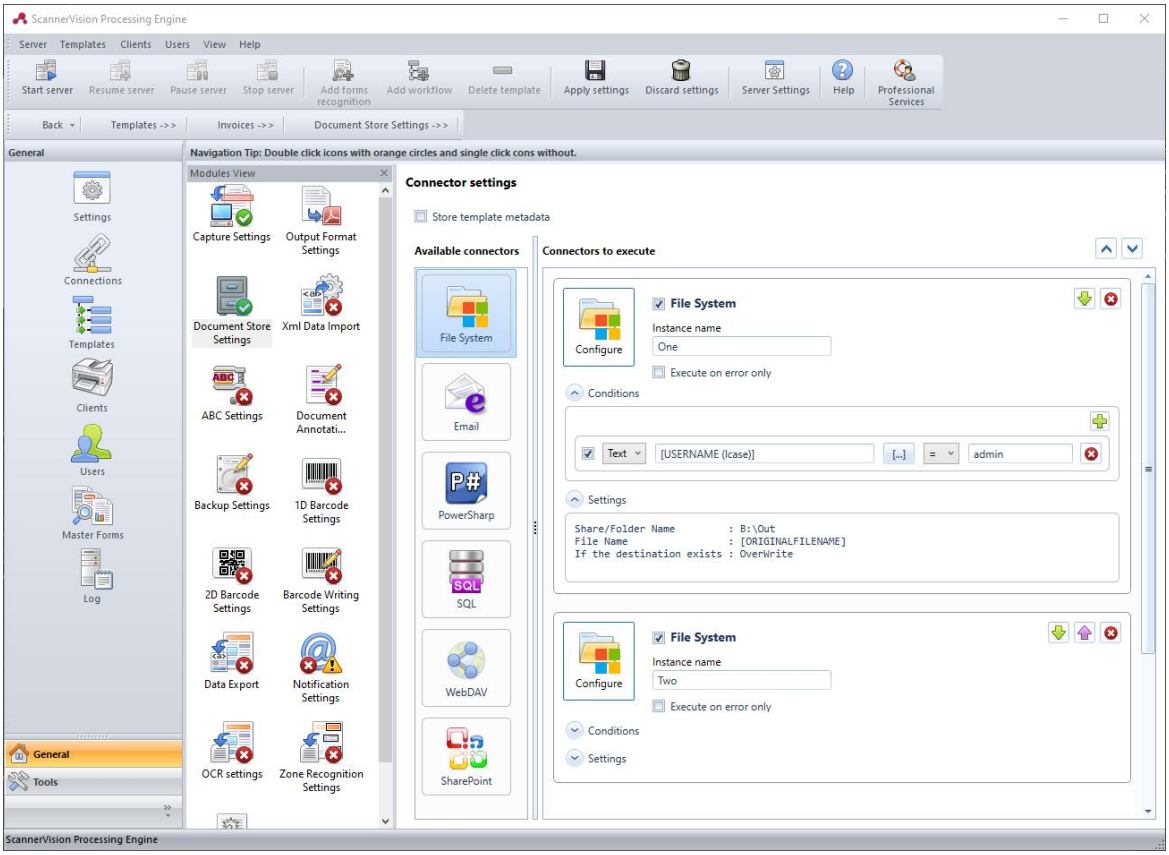
The email body.

8.6.5 Connector Settings

Connectors link ScannerVision to external systems such as relational database engines for which an ADO/ODBC driver is available and document management systems such as SharePoint, Laserfiche and AivikaOne. ScannerVision gives you the ability to put conditions on the execution of connectors which allows you to make decisions on where documents ultimately end up.

ScannerVision ships with a range of connectors which include the Windows File System Connector, Email Connector, Script Connector, SharePoint Connector, SQL Connector and more. To get access to more connectors please contact your ScannerVision reseller.

Multiple connectors can be added to a template by dragging the relevant icon from the "Available connectors" list onto the "Connectors to execute" list shown below.



Store template metadata

When this option is enabled ScannerVision uploads a text file containing the ScannerVision metadata in a Microsoft Ini file format before it uploads the document being processed. The data in the file looks like this:

```
[METADATA0]
Name=VVRD
Value=MjAxNC0wNi0yNVQxMjozMToxOC44ODI1Mzc0Wg==
[METADATA1]
Name=VUNU
Value=MjAxNC0wNi0yNVQxMjozMToxOFo=
[METADATA2]
Name=VFpE
Value=KzAyOjAw
...
[METADATA]
Count=58
```

Each ScannerVision metadata tag is represented in a section called "METADATA" with a number appended e.g. "METADATA0". The entries in each section represent the tag as a Name/Value pair with the Name entry holding the tag name and the Value entry holding the tag value. The data is Base64 encoded.

A section called "METADATA" **without** a number is included with a single value called Count. The Count entry contains the number of metadata entries contained in the ini file which can be used to construct the "METADATAx" section names.

Organizing connectors

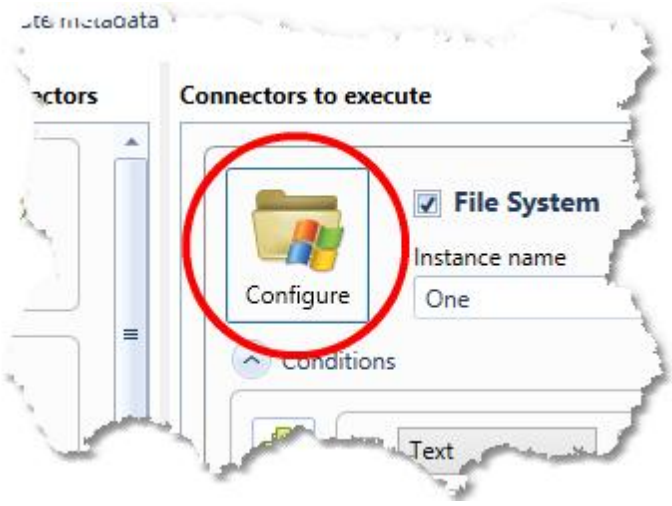
Connectors execute in the order which they appear in the "Connectors to execute" list. A connector's position in the list can be changed by using the arrow buttons in the top right hand corner of the connector's border.

The arrow buttons are context aware so only the arrows that are applicable are visible. For example if only one connector is added to the "Connectors to execute" list none of the arrow buttons will be visible. If more than one connector exists, the top most connector will only show the "Down" arrow and so forth.

To delete a connector click the  button.

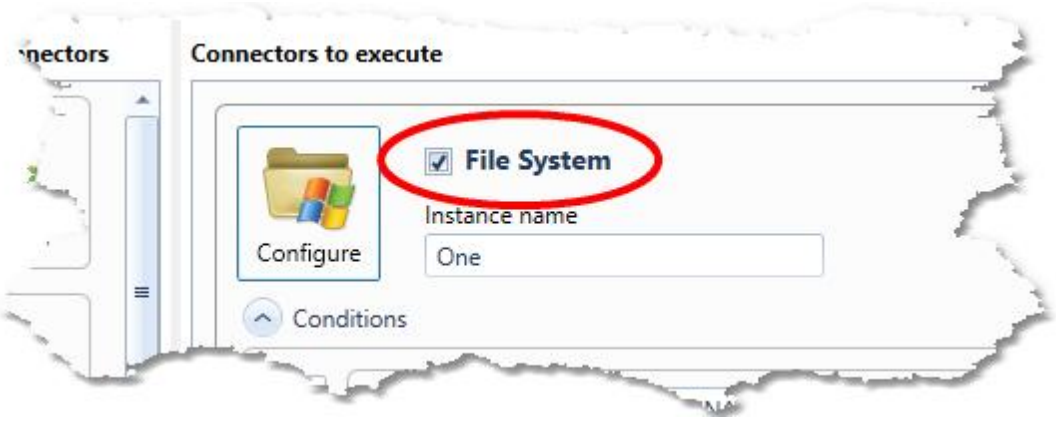
Configuring a connector

Connectors are added to a template by dragging the respective connector icon onto the "Connector to execute" area to the right. Connectors that are added this way must be configured by clicking the "Configure" button (shown below). The connector's setup screen is shown which is different for each connector. Each of the standard connectors' setup screens are discussed later.



Enabling/disabling a connector

When a connector is added it is enabled by default. To disable it deselect the check box to the right of the "Configure" button.



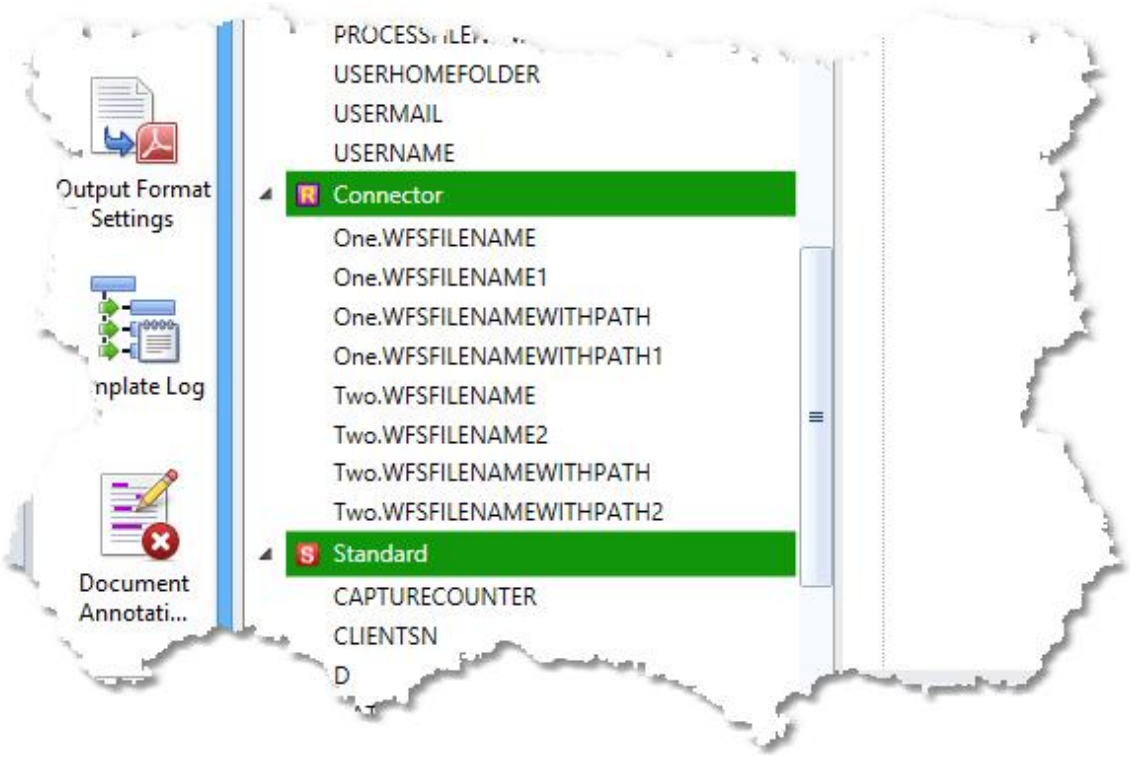
Instance name

A connector's instance name identifies the connector instance when more than one of a particular type has been added to the template. Certain connectors publish metadata that can be used by other connectors such as the File System connector which publishes the tags WFSFILENAME and WFSFILENAMEWITHPATH. When you add two instances of the Windows File System Connector and you don't specify an instance name you will not be able to use the ScannerVision Expression Editor and the template will not execute successfully.

You have to ensure that each instance of a particular connector has a unique instance name. The instance name is prepended to the published tag name as shown in the screen shot below. Two File System connectors have been added with the instance names "One" and "Two" respectively.

NOTE

Spaces are not allowed in instance names.



Important

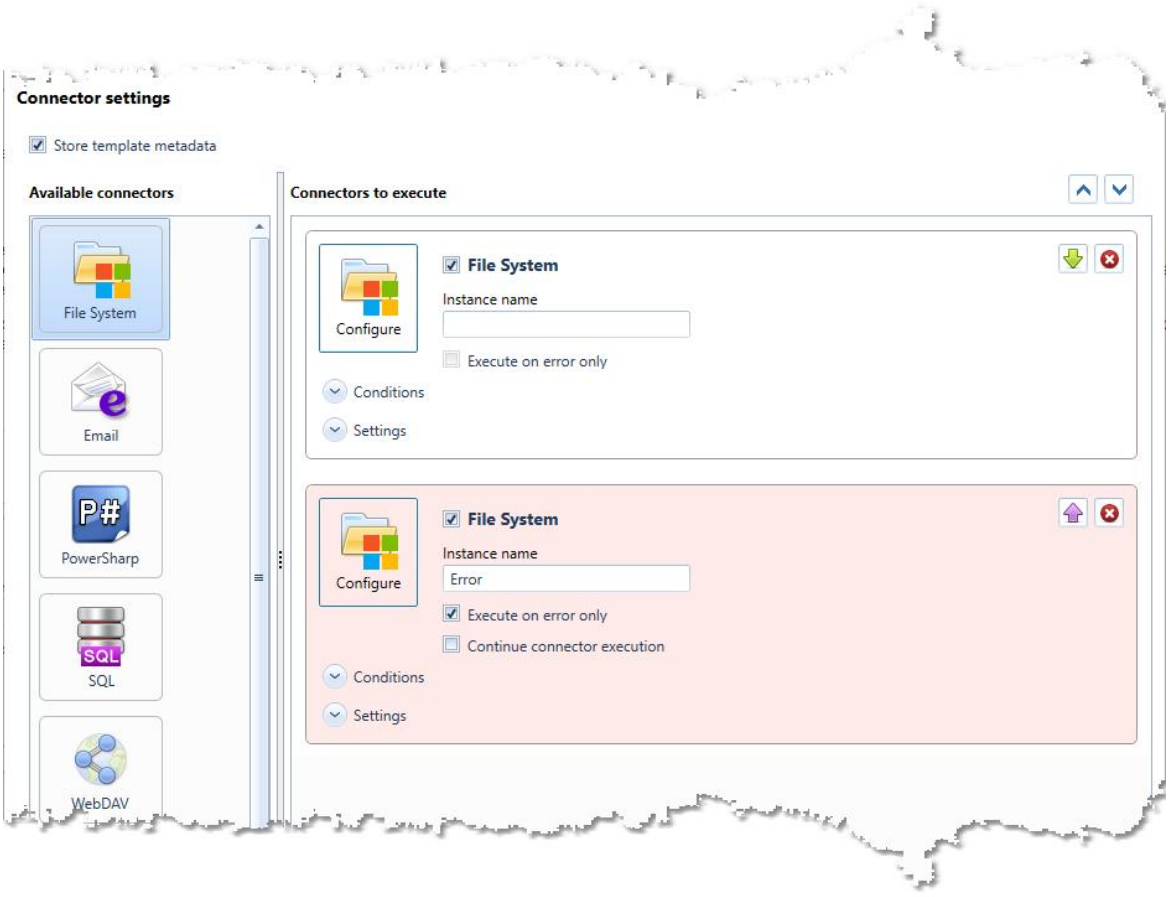
Whenever metadata that is published by connectors are used in the template the instance name must be included.

Execute on error only

If this option is selected the connector will only execute if the preceding connector encounters a problem that would otherwise have caused the template to fail. This option allows you to "catch" the error and execute the error handling connector which could be of any type. The "TEMPLATEERRORS" tag will be updated with the error message of the failing connector which you could for example use in an email connector.

When a connector appears at the top of the connectors' list, the "Execute on error only" option is disabled and if it was selected on a connector which is then moved to the top, the option is cleared automatically. As you can see below, when this option is enabled the connector is colored red so that it stands out as an error handling connector.

You also have the option to let template execution continue even if an error occurs in the error handling connector itself by checking the "Continue connector execution" check box. A use case for this is presented below. In any scenario where an error handling connector completes successfully, the error that has caused it to execute is deemed to have been dealt with by error handling connector and the template as a whole is deemed to have executed successfully - assuming the remainder of the template (including connectors) did not encounter any problems.



Here are a few usage scenarios:

1. You have 2 normal connectors named Upload1 and Upload2 followed by an error handling connector named ErrorUpload which is finally followed by a normal connector named Upload3.

   a. If Upload1 fails, Upload2, ErrorUpload and Upload3 are skipped and the template fails.

   b. If Upload1 succeeds but Upload2 fails, ErrorUpload executes. If ErrorUpload fails, the template fails. If ErrorUpload succeeds, Upload3 is executed and the template continues until completion.

   c. If both Upload1 and Upload2 succeeds, ErrorUpload is skipped but Upload3 executes.

2. Let's say you are sending processed documents via email to some recipients but you want to have fail-over that if the email cannot be sent to SMTP server 1, you want to send it via SMTP server2, and if that also fails you want to finally try SMTP server 3. You would then configure 3 email connectors called e.g. Email1, Email2 and Email3 with each naturally configured to point to their respective SMTP server. Email1 would be a standard connector while BOTH Email2 and Email3 will be error handling connectors, i.e. "Execute on error only" would be selected. In addition, Email2 will have "Continue connector execution" enabled while Email3 will have it disabled (automatically). Finally, you want to upload the document to a network share via a WFS connector called Share1 but only if the email was sent successfully. Here is how it will work:

   a. If Email1 succeeds, Email2 and Email3 are skipped and Share1 executes. If Share1 fails, the

template fails.

b. If Email1 fails, Email2 executes and if successful Email3 is skipped but Share1 executes.

c. If Email1 and Email2 fails, Email3 executes. If Email3 fails, the template fails because "Continue connector execution" was disabled. If Email3 succeeds, Share1 is executed.

3. An extension of scenario 2. above can be to add an exception handler connector after Share1 to catch any error that might occur with Share1. This could be another WFS connector that routes the processed document to another hot folder on the server that is monitored by a second template. This second template could then retry the upload to the share without duplicating sending of the email.

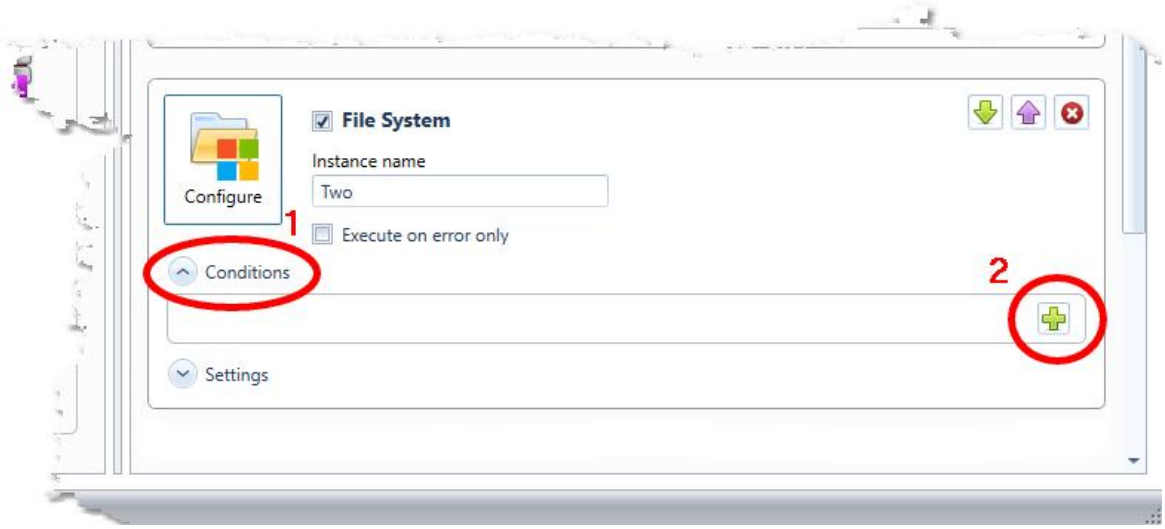
Conditions

Conditions allow you to decide when a connector should execute. Conditions could be based on textual or numeric comparisons or on ScannerVision expressions that evaluate to true or false. The screen shots below have been taken in the context of the Windows File System Connector but they work the same way and have the same meaning with all connectors.

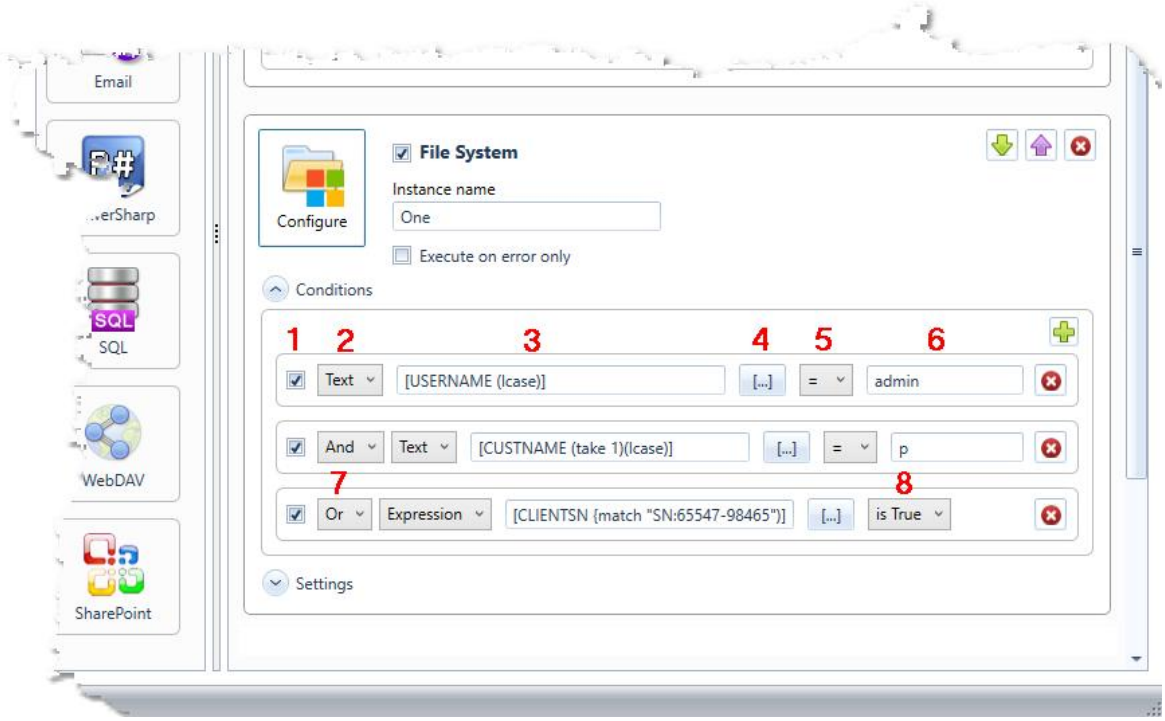
To add a condition:

1. Expand the "Conditions" section.

2. Press the "+" button in the top right hand corner of the Conditions border.



In the screen shot below three conditions have been added.



The bold red numbered parts of the conditions in the screen shot above have the following meaning:

1. Enables/Disabled the condition. Disabled conditions are not evaluated - it is as if they don't exist.
2. Condition type. See below for more information.
3. Argument one. The argument you want to compare - typically a metadata expression.
4. Displays the ScannerVision Expression Editor to construct the expression of the first argument.
5. Comparison operator. See below for more information.
6. Argument two. The argument you want to compare against. In the case of Text conditions this is a text phrase and in the case of Number conditions this is a rational number.
7. Logical operator. See below for more information.
8. Invert logic. See below for more information.

In the discussion that follows references are made to ScannerVision Metadata Expressions. If you don't know what they are or how they work please study Appendix A - Metadata first.

Condition type

Three types of conditions are supported namely:

1. Text

Textual comparison is based on the ordinal value of each character in the text. For an explanation of what the ordinal value of a character is please refer to the Tags section. It is important that you understand this concept otherwise you may not understand why your conditions don't evaluate the way you expected.

The characters 'a' and 'A' are not the same because they have different ordinal values. In the screen shot above you will see that in the first condition we have included the "lcase" function which converts all the characters of the USERNAME tag to lowercase. This ensures that the condition does not give a false negative when the username contains uppercase characters. So even if the username is "Admin", "ADMIN" or "aDMin", the comparison that is performed would be "admin" = "admin". When you don't care what the case is of the characters in the text that you compare we recommend that you follow this approach to make your conditions more robust.

When it comes to textual comparison there is no notion of numbers and symbols. The phrase "5 > 1" does not mean 5 greater than 1 in the context of textual comparison because the ordinal value of the characters '5', '1' and '>' has nothing to do with the numeric values 5 and 1 or the "greater than" operator.

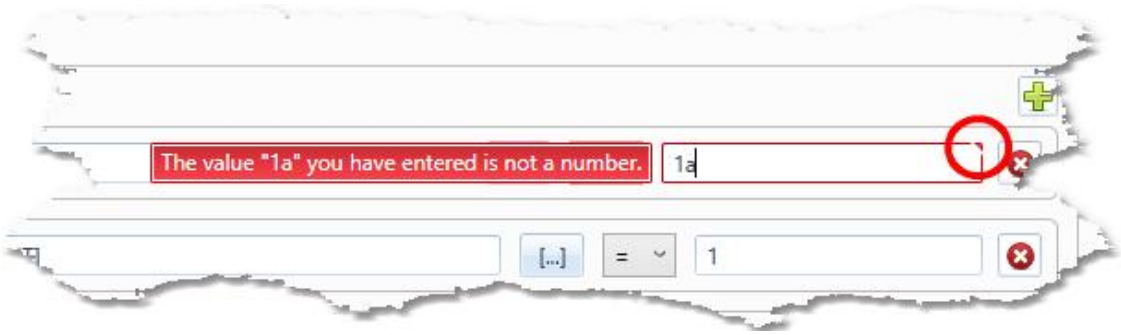
2. Number

Numeric comparison is based on rational numbers (numbers that can contain decimals). When numeric conditions are evaluated the first argument is converted to the rational number that is represented by the text. For example if the text is "515.90" it is converted to the number 515.90.

When using numeric conditions you have to be sure that the value of the tag CAN be converted to a number. In other words, the text must represent a number only. If for example a tag value includes the currency symbol, say "$", to hold the value "$515.90" the conversion to a number would fail. This does not mean that you can't use a numeric condition in this situation, you just have to strip away the "$" character by adding a "(take 2-)" function to the tag. ScannerVision trims all leading and trailing non-printable character off the first argument before the conversion process. So if the tag (let's call it TAGNAME) value was "$ 515.90" you don't have to do the following:

[TAGNAME (take 3-)] or [TAGNAME (take 2-)(trim)]

Numeric conditions require a number as the second argument. As soon as you enter a value that cannot be converted to a number a red border will appear around the edit box and a small red triangle will appear in the top right hand corner of the edit box. Hover the mouse over this triangle to see a tool tip with a description of the error as shown below:



3. Expression

Expression conditions are evaluated as TRUE or FALSE. A TRUE expression is one that contains one or more characters and a FALSE expression is one that does not contain any characters, i.e. it is completely empty. The phrase " " is not empty, it contains 3 space characters and therefore is a TRUE expression. The space character is one of the non-printable characters (see the Tags section for more information) and they are not stripped automatically when expressions are evaluated. Make sure you do this yourself to eliminate difficult to find errors.

You would typically use the "match" function in your expressions as this gives you the full power of regular expressions to look for any text or sub-text in metadata such as barcodes and OCRed text.

Inverted logic

You could also choose to invert the logic of the expression by selecting "Is False" in the drop down box shown below:

By inverting the logic you are instructing ScannerVision to execute the connector when the expression is FALSE. How is this useful? You would use this technique when you want to execute a connector only when another did not execute, i.e. "Execute connector 'B' only when connector 'A' did not execute".

As an example let us assume you want to route documents into two different folders. Your template is processing Invoices, Orders and Quotations and you want the invoices to go the "Processed Invoices" directory on a network share and Orders and Quotations to go to your SharePoint server. For the sake of this example let's assume you are OCRing the whole document with the idea to look for the phrase "INVOICE" in the OCRed text. To route invoices to the "Processed Invoices" share you would add a File System connector with the condition:

[OCRTEXT (ucase)(match "INVOICE")] with "invert logic" option set to "Is True". So, if the phrase "INVOICE" is found in the OCRTEXT tag the connector is executed.

To route Orders and Quotations to your SharePoint Connector you would add a SharePoint Connector with the EXACT SAME condition that you used in the Windows File System Connector except the "invert logic" option set to "Is False". When ScannerVision evaluates the expression and the phrase "INVOICE" is found the expression evaluates to TRUE but since the logic is inverted the TRUE now becomes a FALSE and the connector is NOT executed. Inversely, if the phrase "INVOICE" is **not** found in the OCRTEXT tag the expression evaluates to FALSE but with the inversion of the logic the FALSE becomes TRUE and the connector is executed.

Comparison operator

The comparison operator determines how arguments of Text and Number conditions are compared. The comparison operators are:

= Equal to

!= Not equal to

< Less than

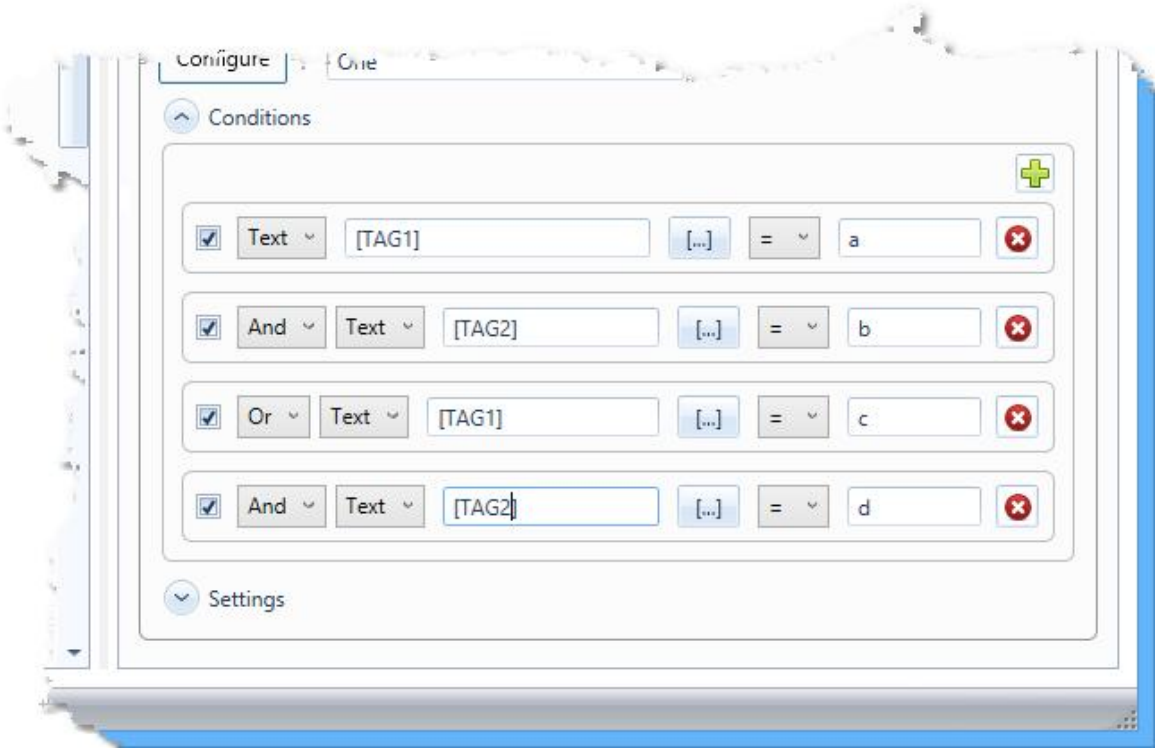
<= Less than or equal to

> Greater than

>= Greater than or equal to

Logical operator

When you have more than one condition you have to tell ScannerVision whether all conditions must be true before the connector executes or whether any one could be true. In the screen shot below we have four conditions:



ScannerVision evaluates the conditions above as follows:

(([TAG1] = 'a') and ([TAG2] = 'b'))

or

(([TAG1] = 'c') and ([TAG2] = 'd'))

In English this means the connector will execute when:

Tag 1 is equal to 'a' AND Tag 2 is equal to 'b'

OR when:

Tag 1 is equal to 'c' AND Tag 2 is equal to 'd'

Every "And" condition that follows an "Or" condition up to end of the conditions or the next "Or" condition are evaluated together.

8.6.5.1 File System Connector

The Windows File System Connector allows documents to be stored in a local system folder or on a network share. You can use the connector in either of two modes namely "Legacy" and "Version 8".

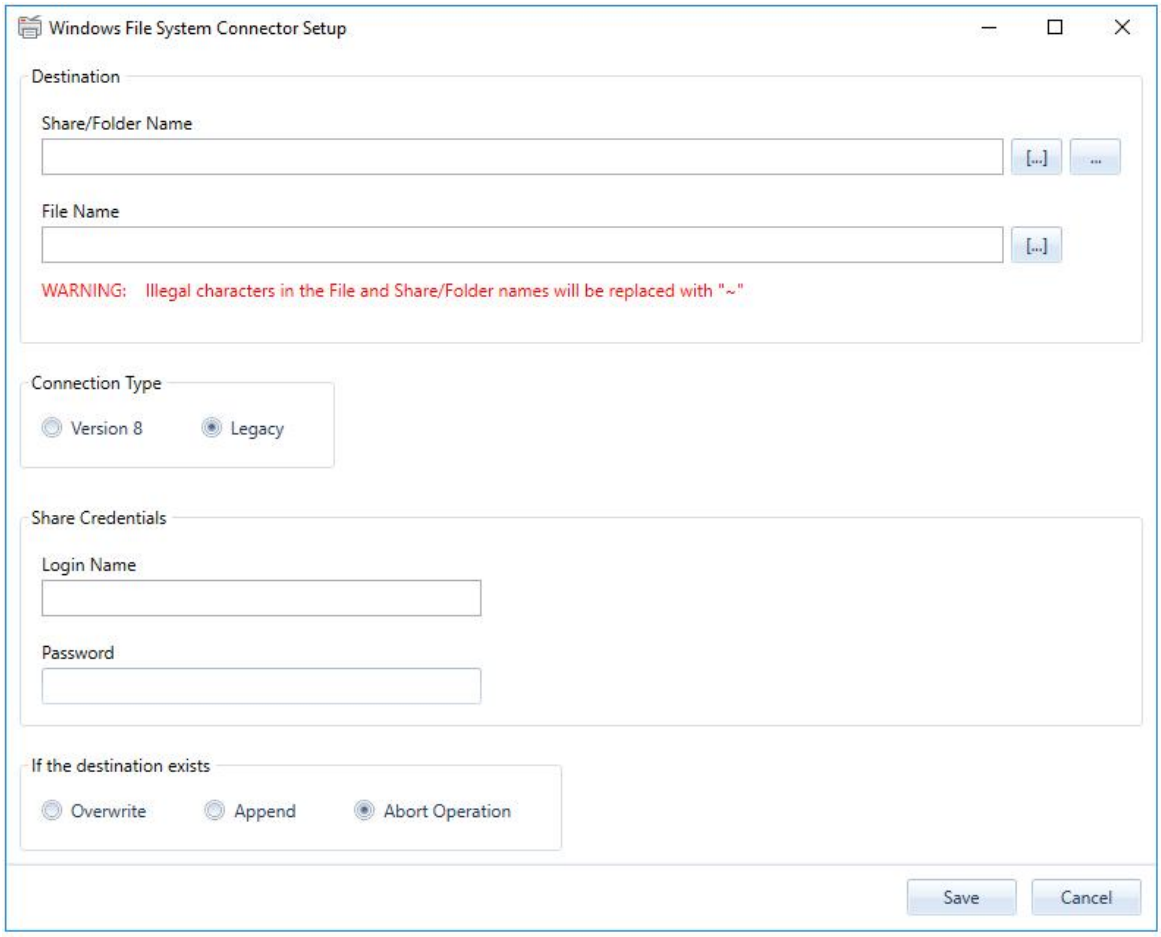
Legacy

Legacy mode means that you will use the connector like earlier versions where you explicitly provide the full "Share/Folder Name", "Login Name" and "Password" in the connector settings. You could use this method where you know you are going to use the particular connector configuration only once. If on the other hand you are going to need to upload documents with multiple connector instances to the same folder or network share, whether in the same template or different templates, you would be better served by using the "Version 8" connection type.

Version 8

In this mode you can use a Connection instead of having to provide the full "Share/Folder Name", "Login Name" and "Password" for every instance of the connector you create. With Connections you can centrally manage Folders/Shares and connection credentials.

The Windows File System Connector connection screen is shown below:



Anywhere that you see the "[...]" button you can make use of ScannerVision metadata tags.

Share/Folder Name

The path to the shared drive or folder name.

File Name

The file name of the document that is stored in the share or folder. There are several ways to

generate file names dynamically. See Appendix H - Unique Suffix for details.

Connection Type

See above.

Share Credentials (Legacy)

If documents are to be stored in a network share that requires authentication the login name and password is specified here.

Connection Name (Version 8)

The name of the Connection to use.

If the destination file exists

The file name already exists in the destination folder/share you have three options to handle the situation.

- Overwrite
- Append (Tiff and PDF only)
- Abort operation

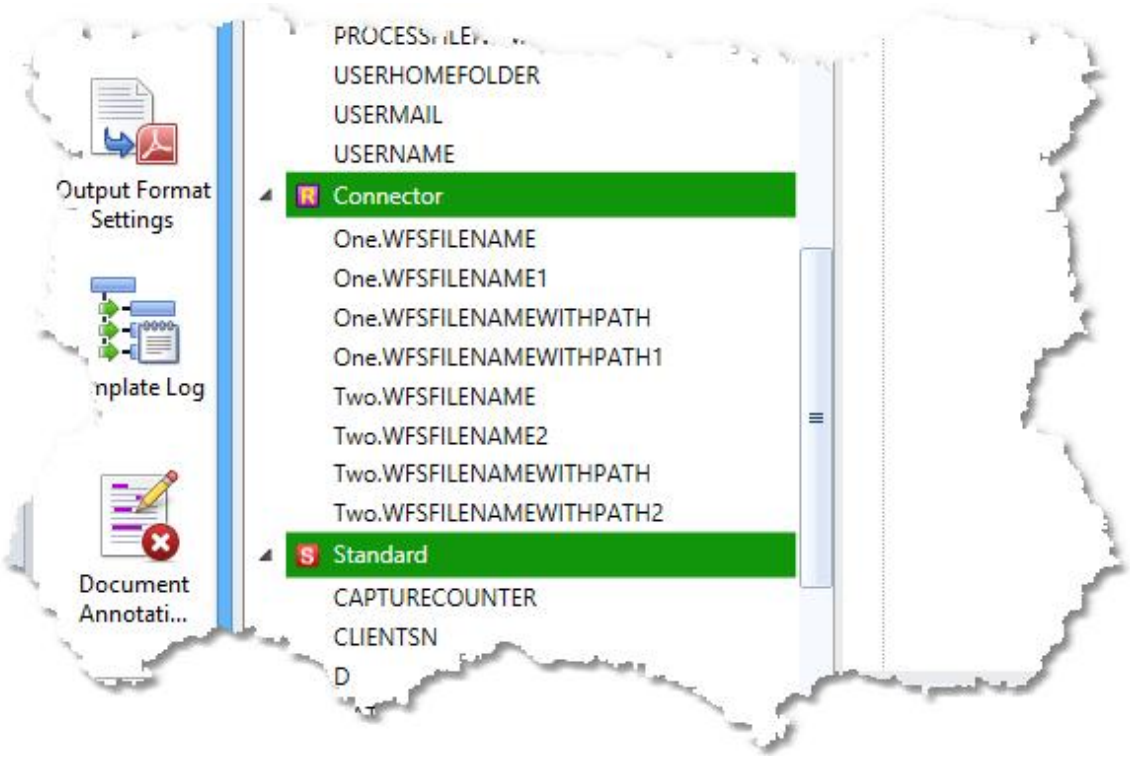
If the [UNIQUESUFFIX] tag is used in the file name none of these options apply.

Published tags

The File System connector publishes the following tags:

Tag Name	Description
WFSFILENAME	The file name of the document that is stored by the connector excluding the path.
WFSFILENAMEWITHPATH	The file name of the document that is stored by the connector including the path.

In addition to the tags above you will see the same tags with a number appended to the end in the ScannerVision Expression Editor as shown below:



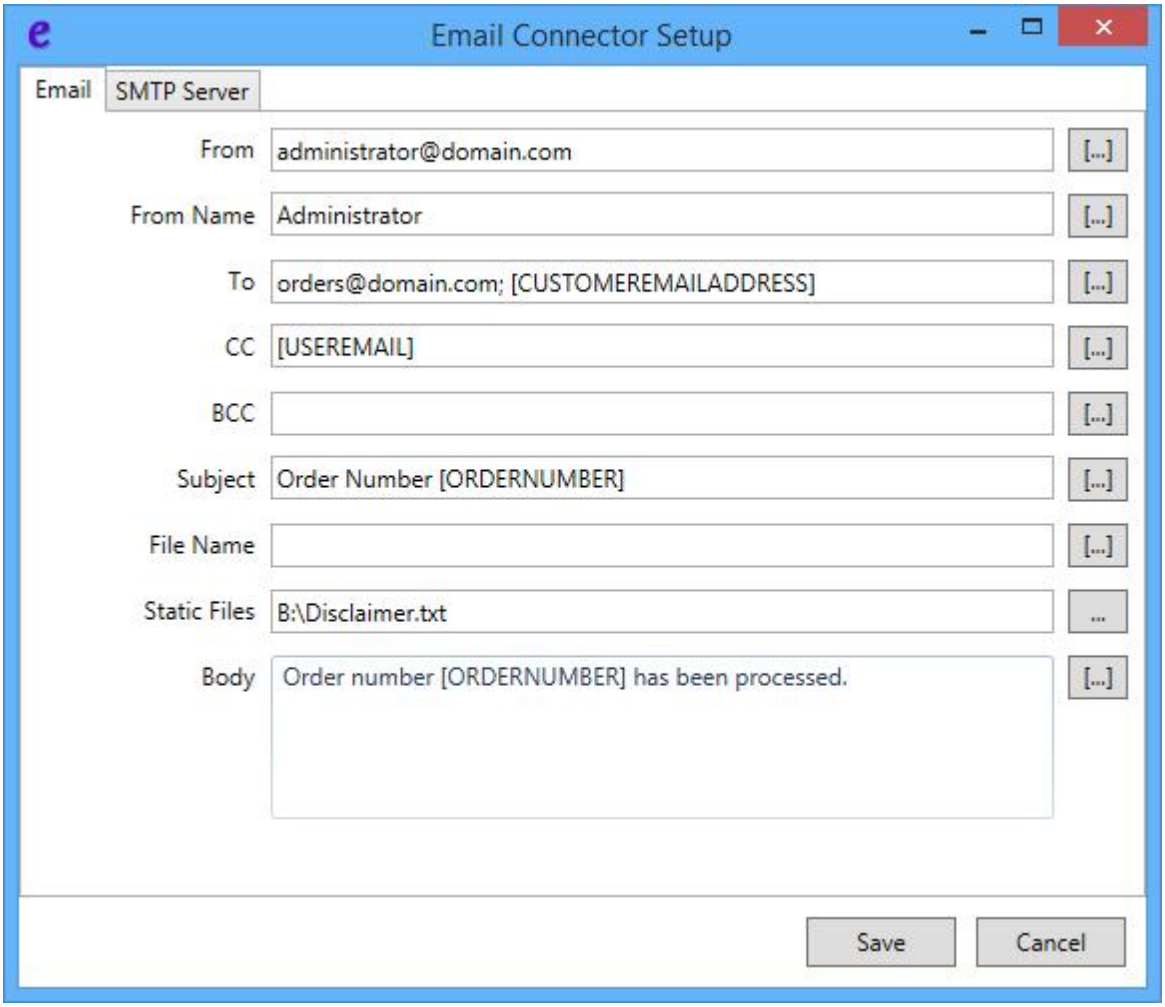
The numbered tags have been superseded by the Instance Name of the connector and it is recommended that you don't use them.

Illegal characters

The Windows File System Connector has ability to dynamically identify illegal characters in file and directory names. Any illegal characters that are identified are stored to disc for future reference.

8.6.5.2 Email Connector

The Email connector sends documents as email attachments to one or more email recipients.



Anywhere that you see the "[...]" button you can make use of ScannerVision metadata tags.

From

The sender's e-mail address.

From Name

The sender's name.

To

The recipients' email addresses separated by semicolons.

CC

Carbon copy email addresses separated by semicolons.

BCC

Blind carbon copy email addresses separated by semicolons.

Subject

The email's Subject line.

File Name

The name given to the attached document. If this field is left blank the original document name is used.

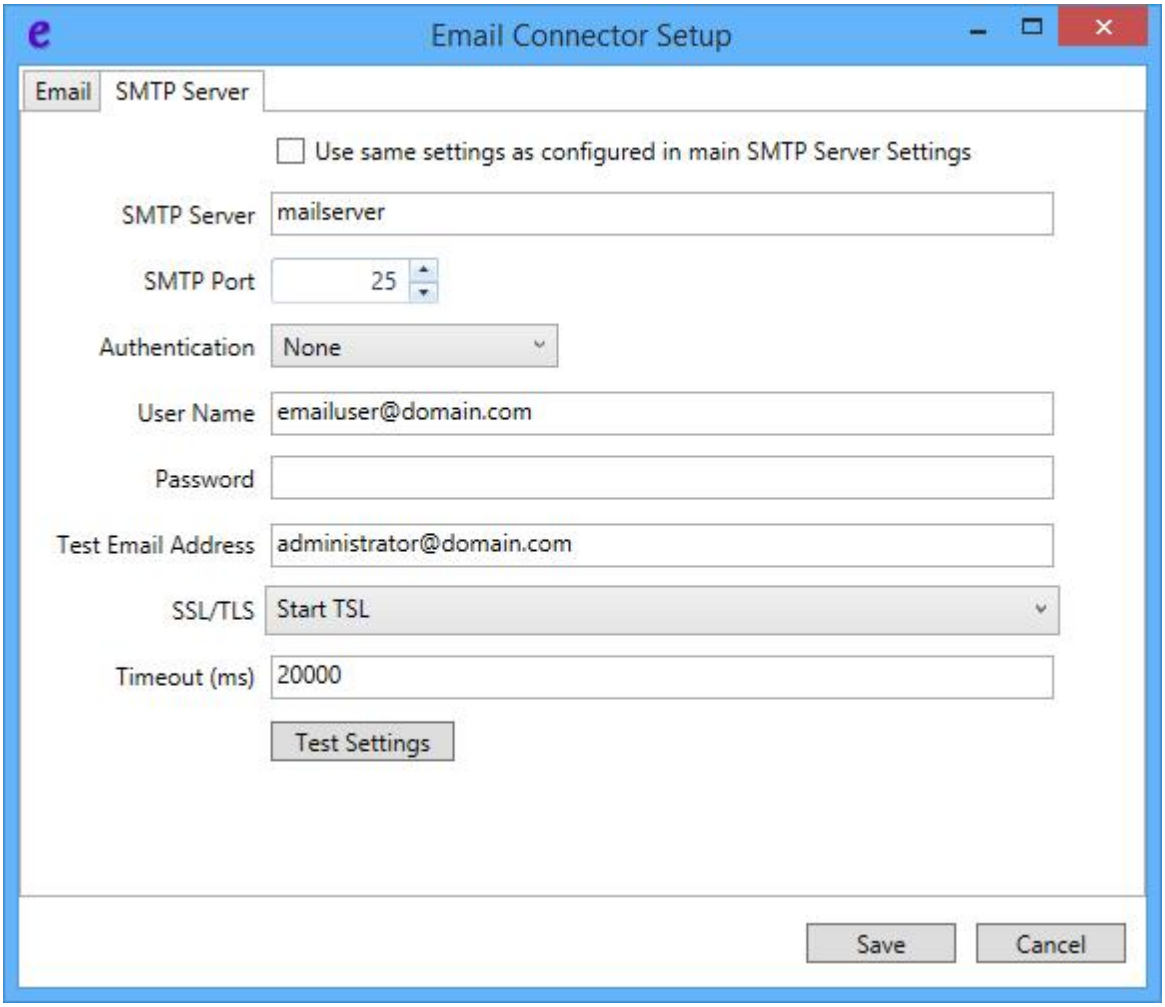
Static Files

List of files to be attached to the email separated by semicolons. This can be used to attach for example a disclaimer to every email that is sent.

Body

The message body of the email.

The screen shot below shows the SMTP server configuration screen.



Use same settings as configured in main SMTP Server Settings

If this check box is selected the email connector uses the SMTP server settings as configured in ScannerVision's "SMTP Server Settings" screen.

SMTP Server

The SMTP server address.

SMTP Port

The SMTP server port.

Authentication

The authentication method to use when contacting the SMTP server. Options include:

1. None
2. CRAM MD5
3. SMTP
4. SMTP Clear
5. Microsoft NTLM
6. POP3 before SMTP
7. Autodetect

User name

The account with which the connector connects to the SMTP server.

Password

The password with which the connector connects to the SMTP server.

Test Email Address

The email address to which a test email is sent when the "Test Settings" button is clicked.

SSL/TLS

Select the encryption option for communication between ScannerVision and the SMTP server.

Timeout (ms)

The time allowed for sending the email message.

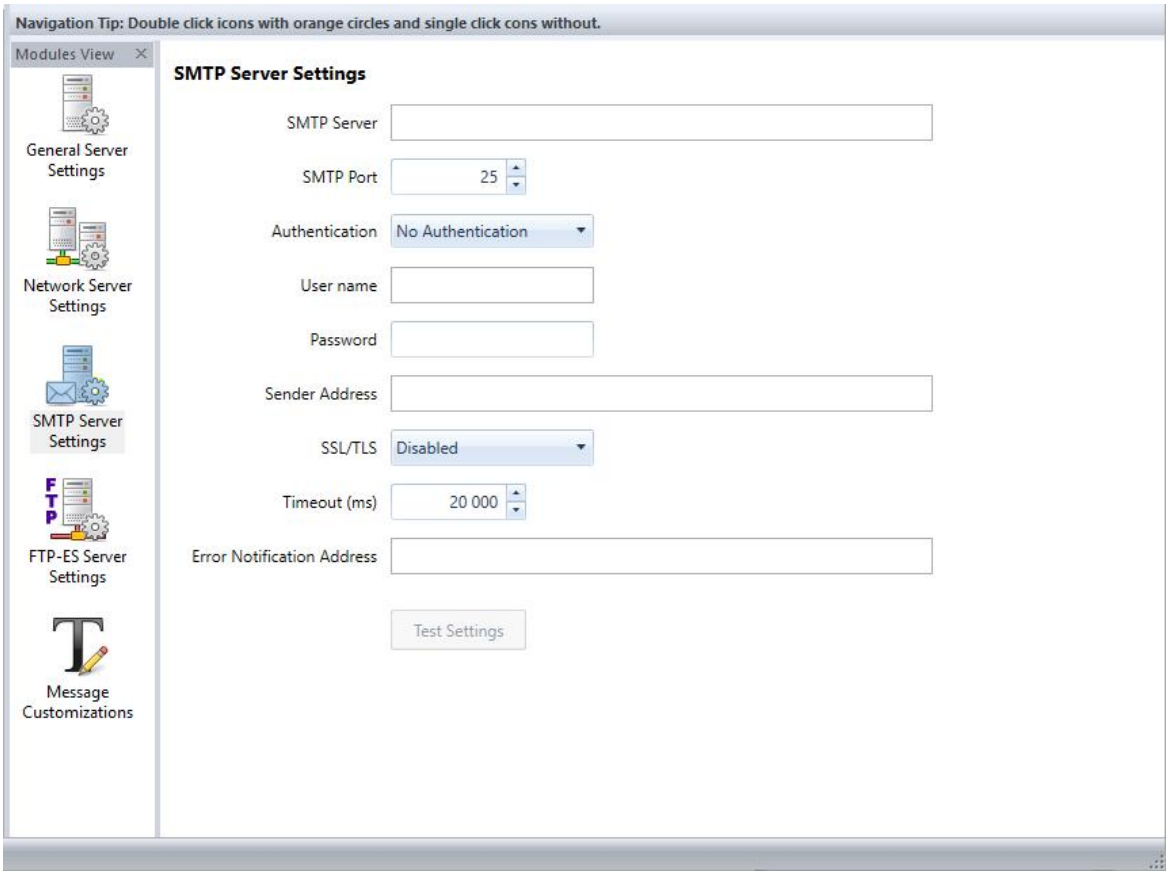
Test Settings

You can test that your SMTP server settings are correct by clicking the "Test Settings" button upon which a test email is sent to the provided email address.

When the "Use same settings as configured in main SMTP Server Settings" is checked the "Sender Address" is used as the "From" address and the "Error Notification Address" is used as the "To" address of the test email.

8.6.5.2.1 SMTP Server Settings

When you select the "Use same settings as configured in main SMTP Server Settings" option on the "SMTP Server" tab of the Email Connector the connector maps the settings of the "SMTP Server Settings" page shown below to its own settings.



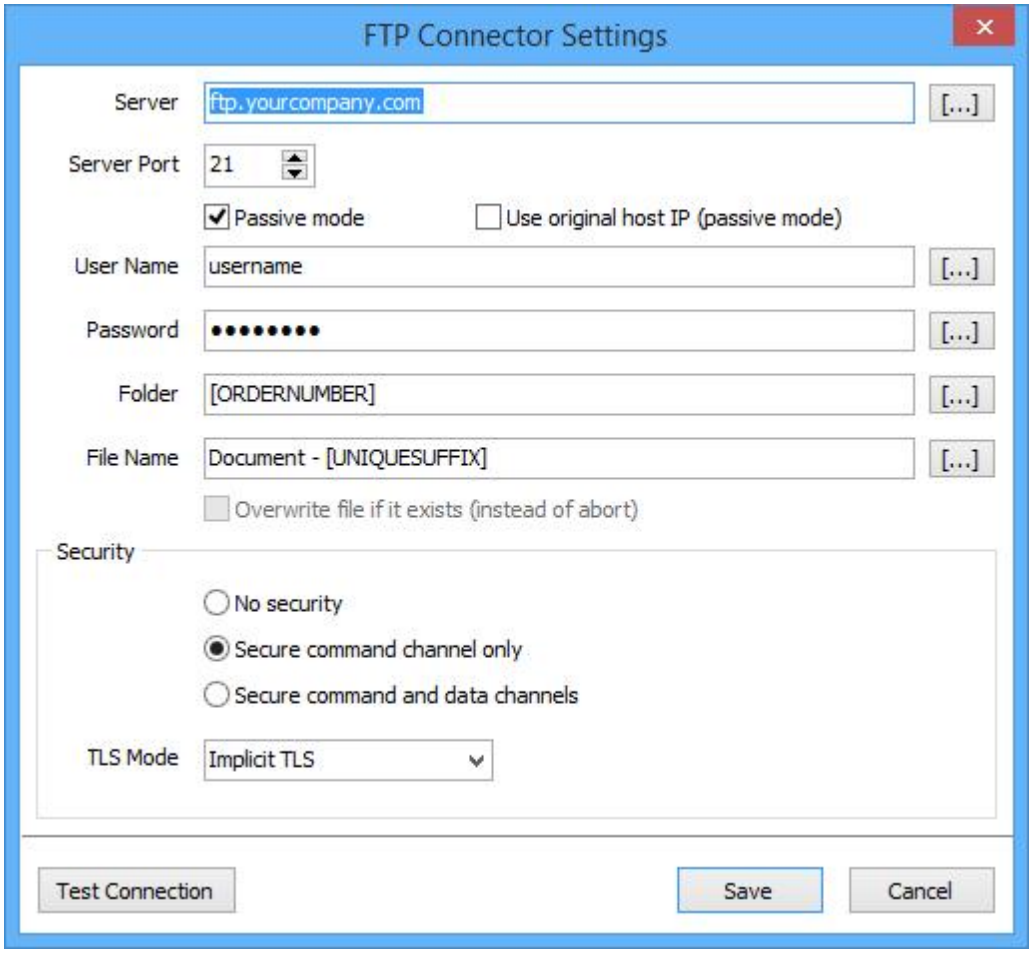
The "SMTP Server Settings" are mapped to the connector's settings as follows:

SMTP Server Settings	Email Connector
SMTP	
SMTP Address	SMTP Server
SMTP Port	SMTP Port
Authentication	
No Authentication	None
CRAM MD5	CRAM MD5
SMTP Authentication	Login
SMTP Clear Authentication	Regular
MS NTLM Authentication	Microsoft NTLM
POP	Auto Detect
Autodetect (Sperimental)	Auto Detect

SSL/TLS	
-	Manual
Credentials	
User name	User Name
Password	Password
Addresses	
Sender Address	From (Email tab)
Error Notification Address	Test Email Address

8.6.5.3 FTP Connector

The Ftp connector uploads documents to an ftp server.



Anywhere that you see the "[...]" button you can make use of ScannerVision metadata tags.

Server

The address of the FTP server.

Server Port

The port of the FTP server.

Passive Mode

Select this check box to connect in passive mode. A discussion of passive mode is beyond the scope of this document. For more information please see Passive mode.

Use original host IP

In passive mode this option is used when the data channel is created for the client in the Get and Put methods. This option allows the client to avoid problems resolving the host name or IP address for FTP services where the workload may be distributed or load-balanced across several servers. When the "Use original IP" option is selected the value in the "Server" property is ignored for the data channel connection and the IP address for the peer connection in the control channel is used instead.

User Name & Password

The credentials to use when connecting to the ftp server.

Folder

The folder name on the FTP server in which the document needs to be stored. If the folder does not exist on the server it is created if rights allow.

File Name

The name given to the document that is uploaded to the FTP server. There are several ways to generate file names dynamically. See Appendix H - Unique Suffix for details.

Overwrite file if it exist

If this option is selected existing files on the FTP server is overwritten.

Security

You can specify 3 levels of security:

No security	Data is sent in the clear.
Secure command channel only	Only control data - which includes the user name and password - sent on the command channel is secured. The file data is not secured.
Secure command and data channels	Both the control and data channels are secured.

When choosing one of the secured modes you must further specify the Transport Layer Security (TLS) mode to employ:

No TLS support	Data is sent in the clear.
Implicit TLS	Communication to the server is secured from the beginning and connection must be to a SSL enabled port.

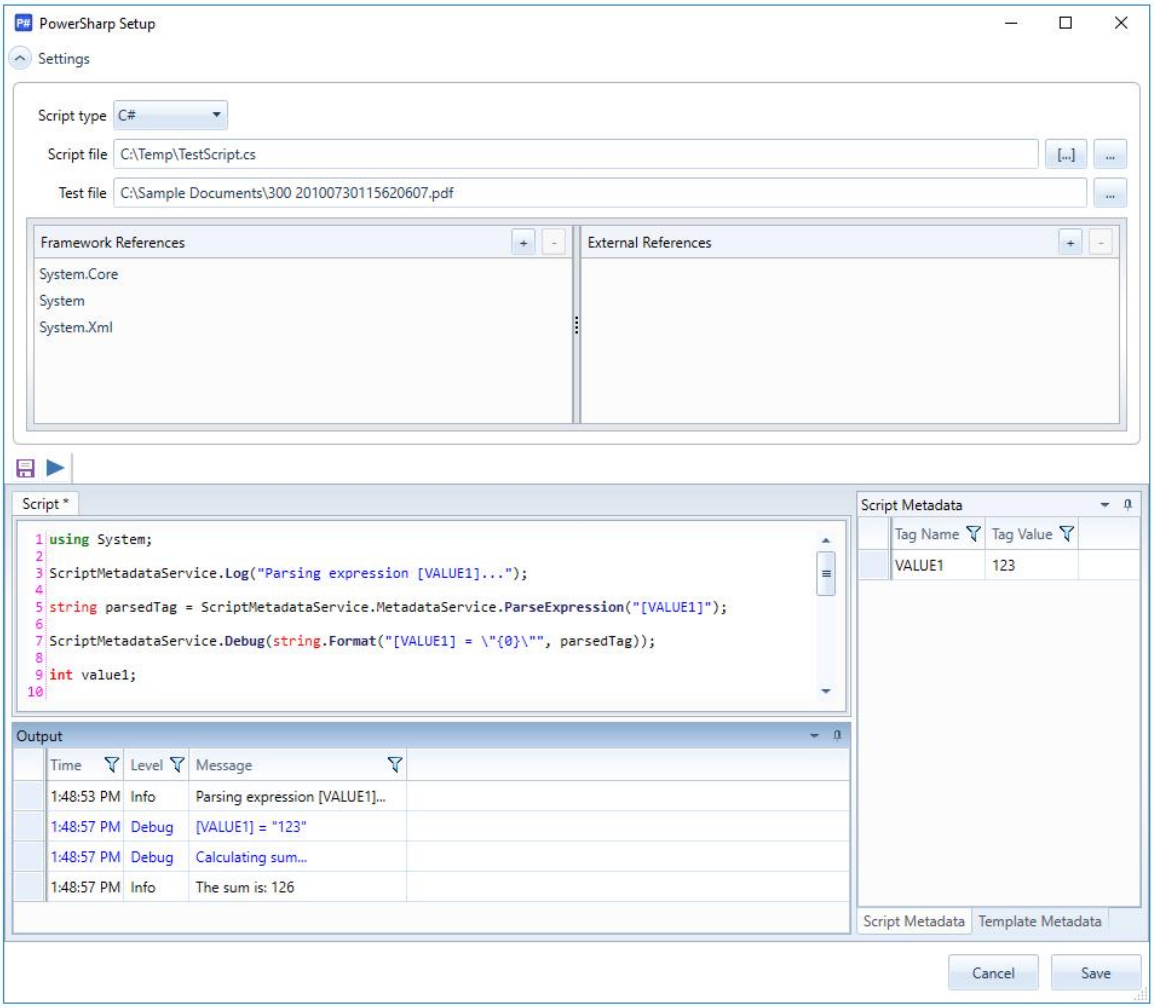
Explicit TLS	Security is not initially enabled but negotiated between the client and server and both ends support SSL/TLS then it will be enabled before any sensitive information such as user name and password is sent. If the server does not support SSL/TLS the connection will fail.

Test Connection

You can test that your FTP server settings are correct by clicking the "Test Connection" button. A connection to the ftp server is attempted but no file data is transferred.

8.6.5.4 PowerSharp Connector

The PowerSharp Connector allows you to execute Microsoft C# or PowerShell code.



Script type

Selects the type of script. Options are "C#" or "PowerSharp".

Script file

If you specify a file name, the contents of the code editor window is loaded from and saved to this file. Only the file path you specify here is saved in the template and not the code itself. This means that if the code in the file changes, the updated code will execute when the connector executes.

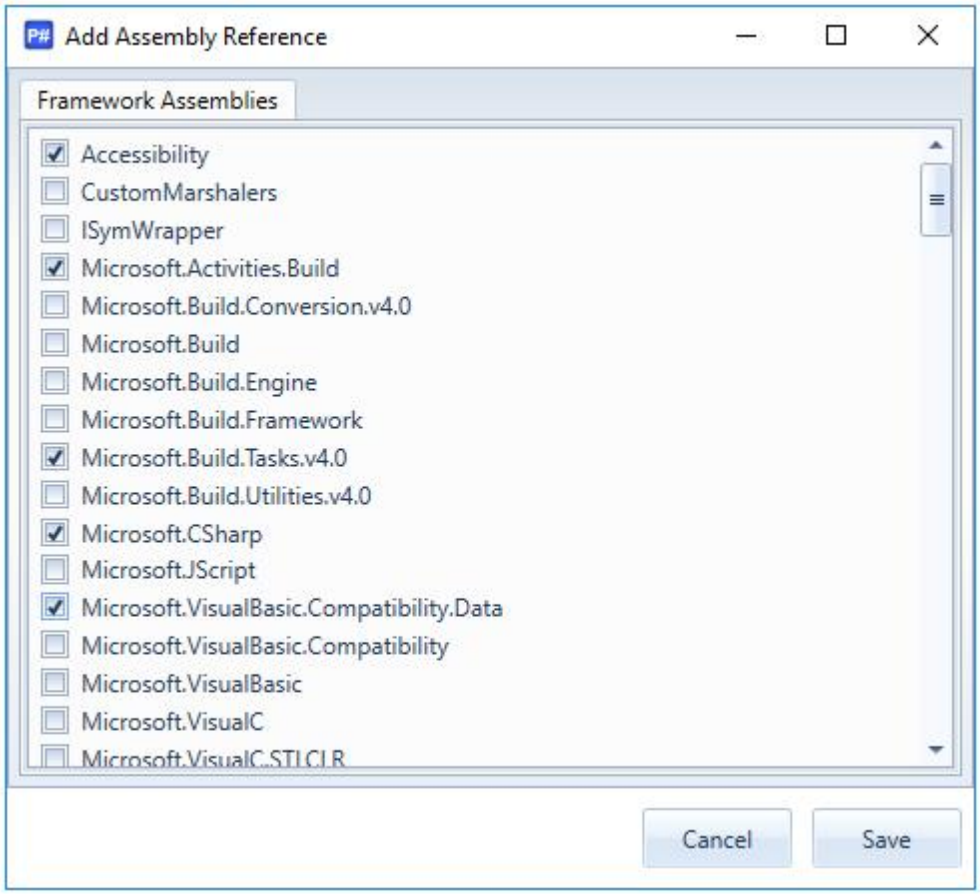
If no file path is provided, the code itself is stored in the template.

Test file

If the script you are creating needs to work on the document that is being processed during run-time, you can still test your script at design-time by specifying a path to test document that will be used by the script.

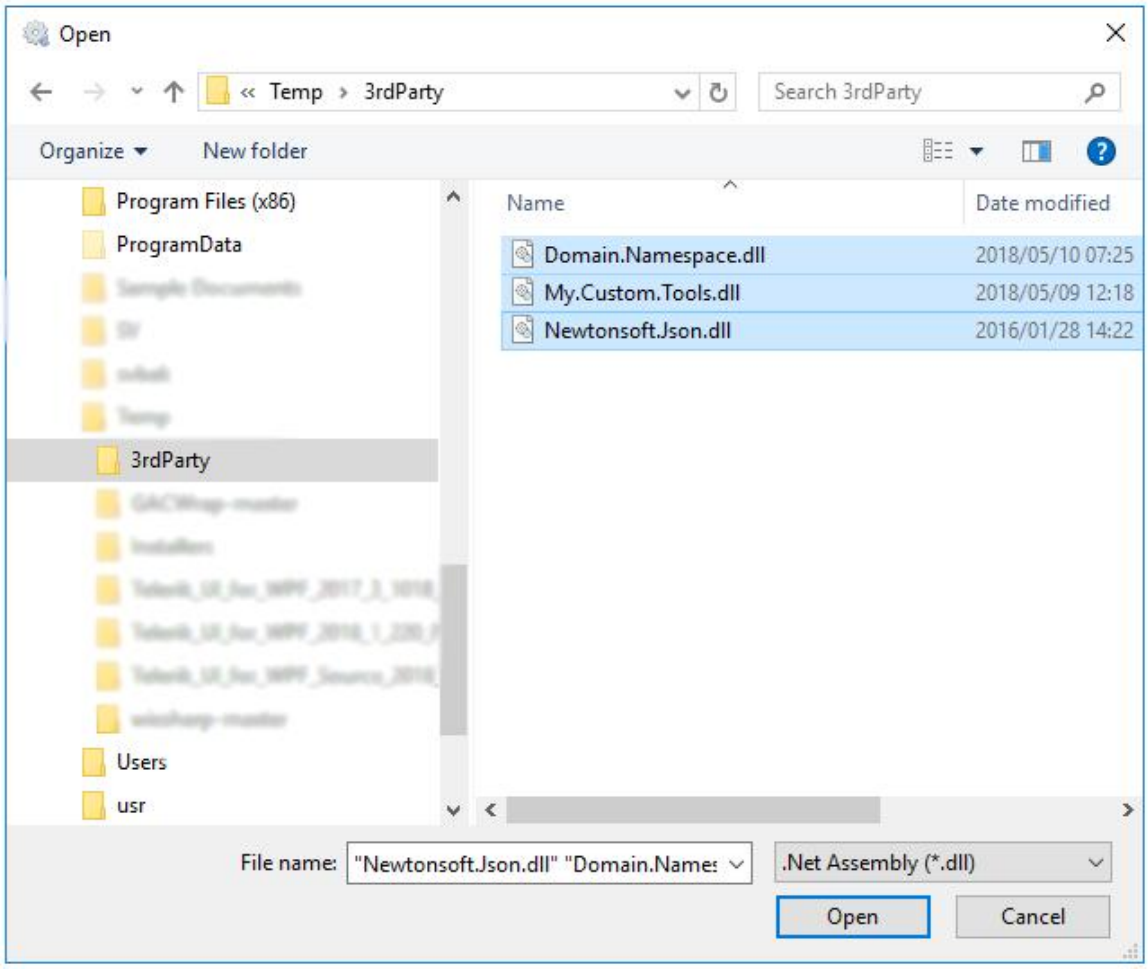
Framework References

These are references to .Net assemblies that are part of the .Net Framework. The list of assemblies which are available is based on .Net 4.7.1. To add a framework assembly click the "+" button in the section heading or select the "Add" context menu option. A dialog is displayed with a list of the available assemblies.

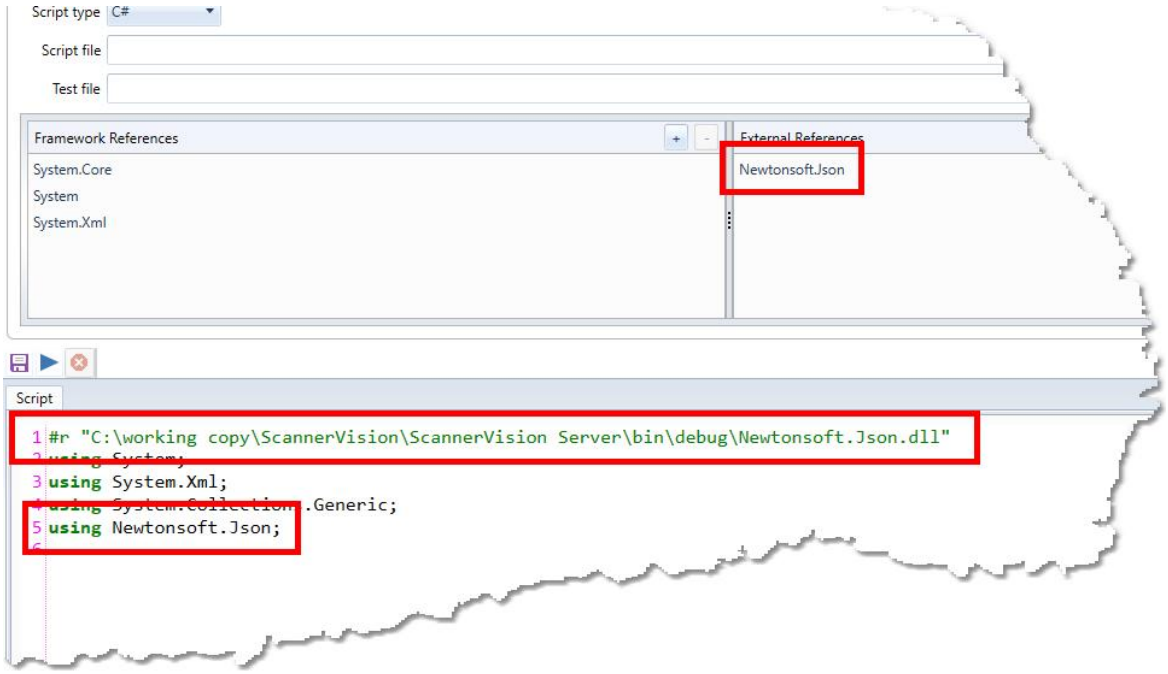


External References

These are references to .Net assemblies that are **not** part of the .Net Framework. These could be assemblies from third parties or your own custom assemblies. To add an external assembly click the "+" button in the section heading or select the "Add" context menu option. A selection dialog is displayed with which you can select the assemblies to reference.

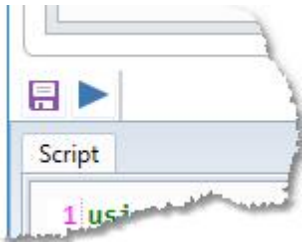


When an external reference is added, a line is added to your script that starts with "#r". This is so that Intellisense works in the editor. This line will not have any impact when the script executes. You still have to add the usual "using" clauses of the namespaces in the assembly you want to use.



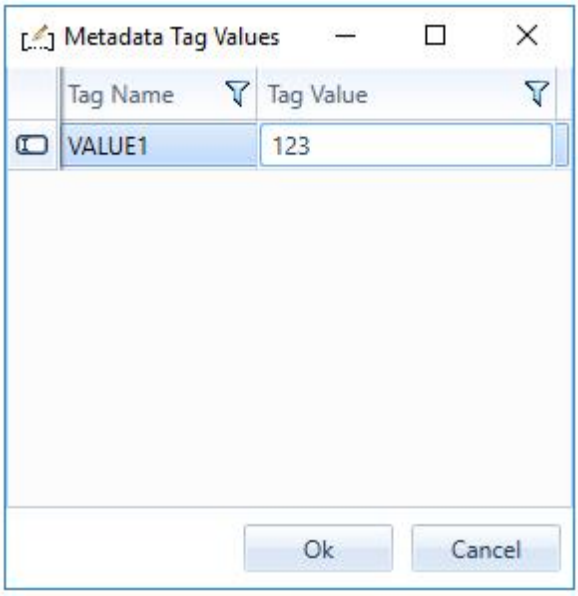
Script

The script that you want to execute. If you have specified a script file path you can save your code without closing the connector window by clicking the button shown below or pressing CTRL+S on your keyboard. To execute the script, click the button or press F5 on you keyboard.

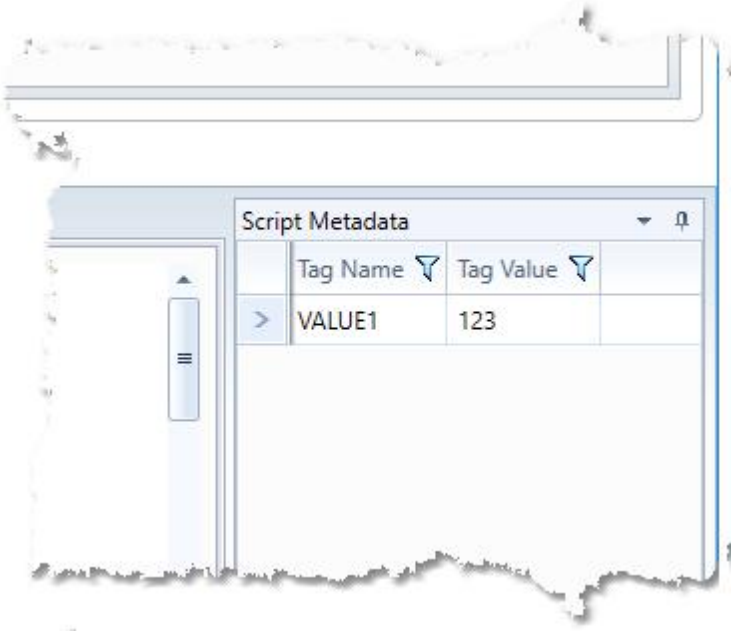


Metadata

A list of metadata tags that are referenced in the script. The first time you execute the script a dialog is displayed that allows you to fill in sample values to use for testing/debugging purposes.



The values you enter here are displayed in the Metadata grid:

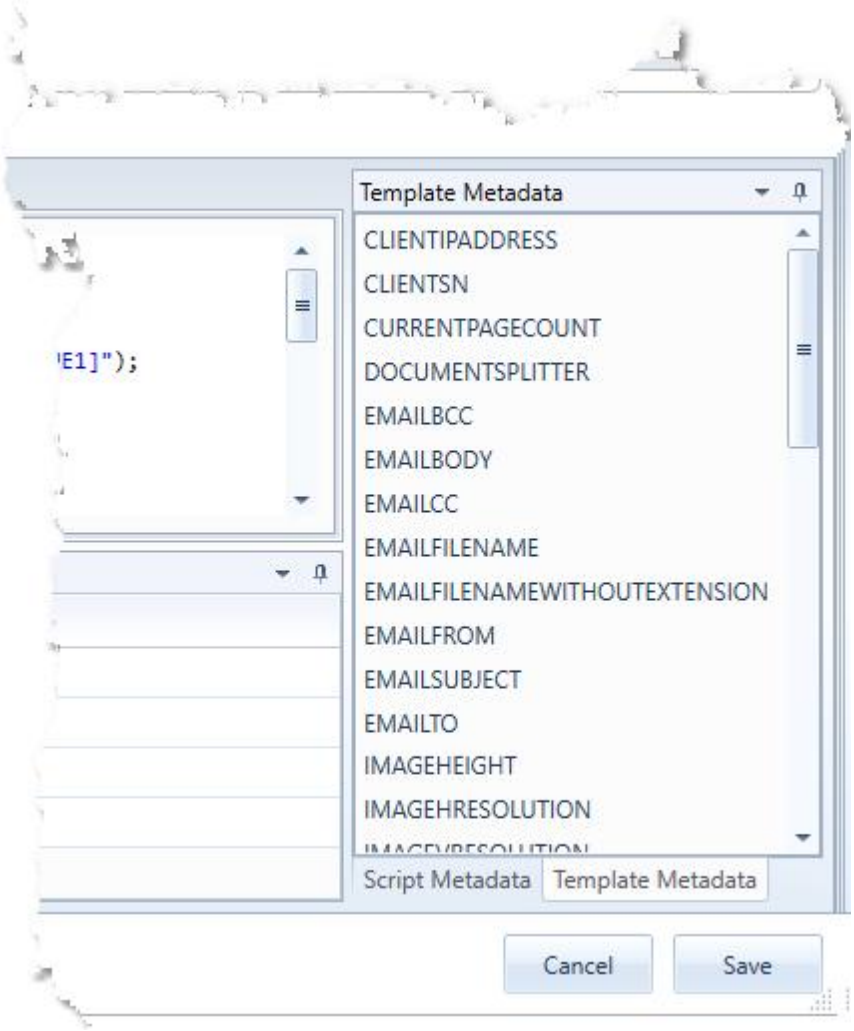


Once metadata tags appear in the grid you can edit their values directly in the grid. These values are stored in the template so you will not have to enter them again when you edit the script.

Template Metadata

The Template Metadata grid contains a list of all the metadata tag names of the template. When you double click a tag name in the grid, it is inserted in your code at the caret position or if text is

selected, the selection is replaced. The tag name is also copied to the clipboard.



Output

The output window shows the result of your script.

8.6.5.4.1 Scripts

The following section applies to both C# and PowerShell scripts except where explicitly stated otherwise.

ScriptMetadataService

ScannerVision provides you with a special class called "ScriptMetadataService" through which you could interact with ScannerVision.

The ScriptMetadataService is static which is to say you do not have to create an instance of the class in order to access it's members. The members of the ScriptMetadataService can be inserted into your code through the context menu of the code editor.

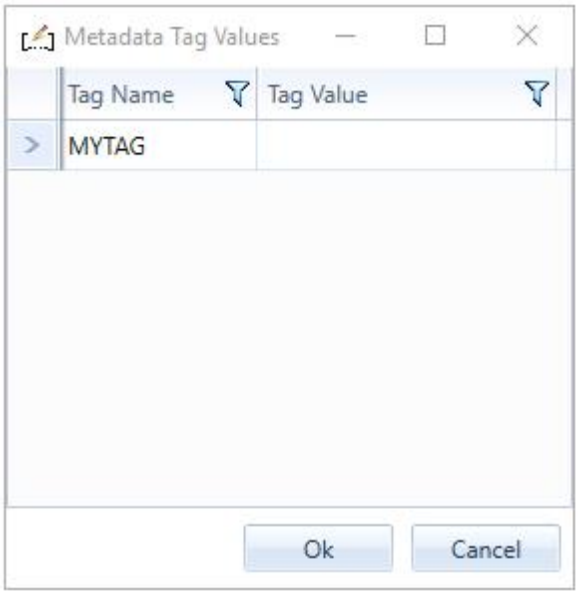
The class provides the following members:

ScriptMetadataService.MetadataService

The MetadataService class is another static class which provides access to ScannerVision metadata tags. The class provides the following members:

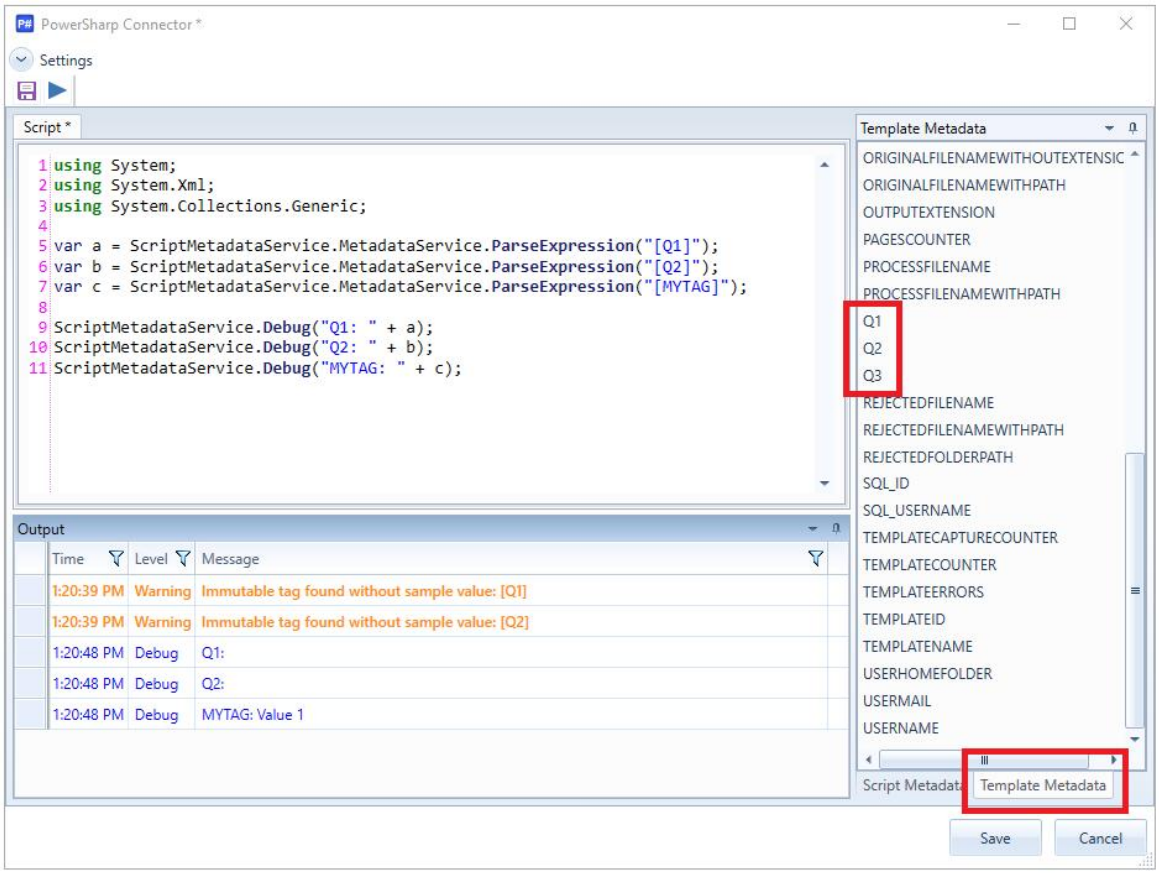
string ParseExpression(string expression);	Parses a ScannerVision expression.
void SetMetadata(string tagName, string value);	Sets the value of a **Connector** metadata tag. You

When you debug a script, i.e. you execute it from the script editing screen, you need to make sure all the metadata tags you reference in the script have sample values. For new metadata tags or ones that have been created/published by other connectors you will be prompter to provide values as shown below:



If you reference template metadata tags such as the tags "Q1", "Q2" and "Q3" - which are the tag names of template questions - that do not have sample values specified, you will receive a warning in the Output window as shown below and the tags will remain empty.

It is important therefore to always provide sample value for all template metadata tags that you create such as template questions, barcodes and OCR zones.



ScriptMetadataService.DocumentName

DocumentName is a read only property that holds the name of the file that is being uploaded or in the case of the debugging, the path you specified in the "Test file" edit box.

Logging methods

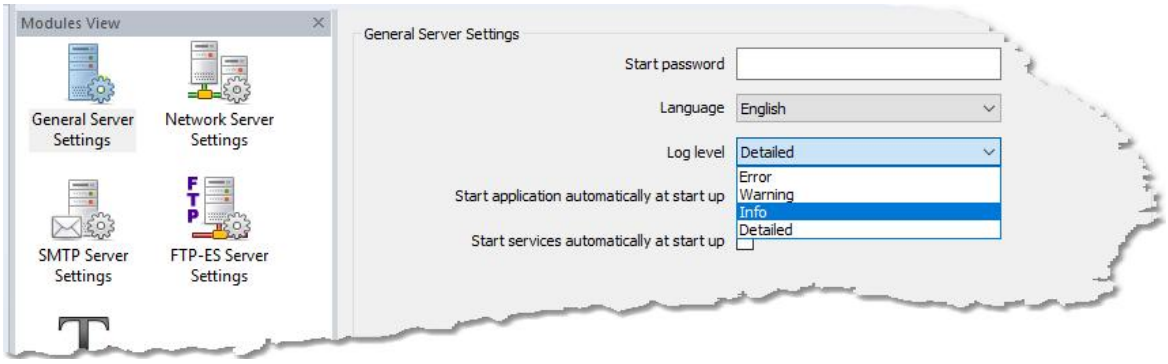
The ScriptMetadataService class provides 4 logging functions:

```
void Debug(string message)
void Log(string message)
void Warning(string message)
void Error(string message)
```

NOTE

When you call the "Error" logging function the log message is wrapped in an exception which is then thrown by the connector at run-time.

The different logging functions relate to the verbosity and severity or level of the messages you want to emit. During debugging the log messages are always written to the output window. During run-time however log messages are emitted according to the "Log level" set in ScannerVision's "General Server Settings" screen:



See the General Server Settings for more information.

8.6.5.4.2 C# Example

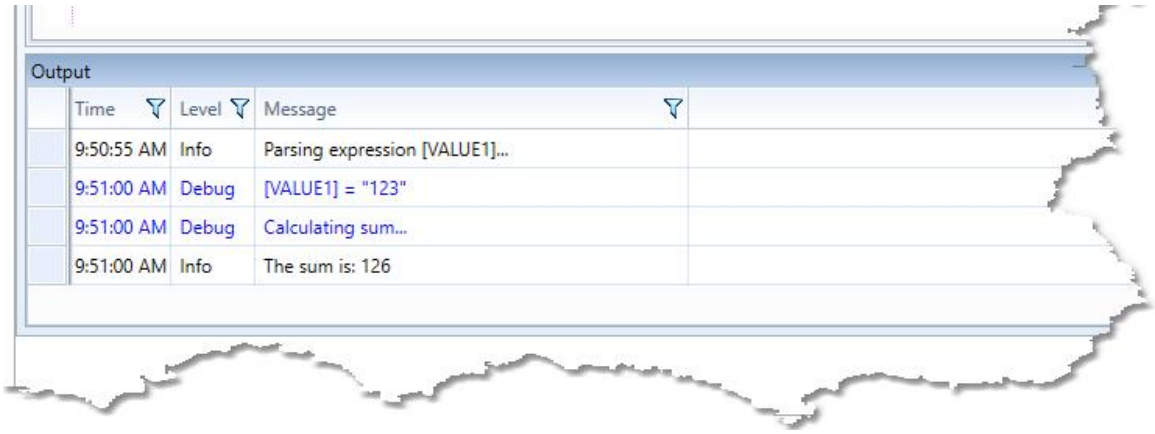
The scripting engine in ScannerVision is based on the Microsoft Roslyn compiler. As such it has enormous power and flexibility which you can read about here.

Below is a trivial example of what you can do with C# scripting. Note specifically the use of the ScriptMetadataService class's members.

```
using System;

ScriptMetadataService.Log("Parsing expression [VALUE1]...");

string parsedTag = ScriptMetadataService.MetadataService.ParseExpression("[VALUE1]");

ScriptMetadataService.Debug(string.Format("[VALUE1] = \"{0}\"", parsedTag));

int value1;

if (!int.TryParse(parsedTag, out value1))
{
   ScriptMetadataService.Warning("The value provided for [VALUE1] is not a valid integer.");

}
else
{
   ScriptMetadataService.Debug("Calculating sum...");
```

```
    var calculation = Sum(value1, 3);

    ScriptMetadataService.MetadataService.SetMetadata("SUM", calculation.ToString());
    ScriptMetadataService.Log("The sum is: " + calculation);
}

int Sum(int value1, int value2)
{
    return value1 + value2;
}
```

Here is the output:

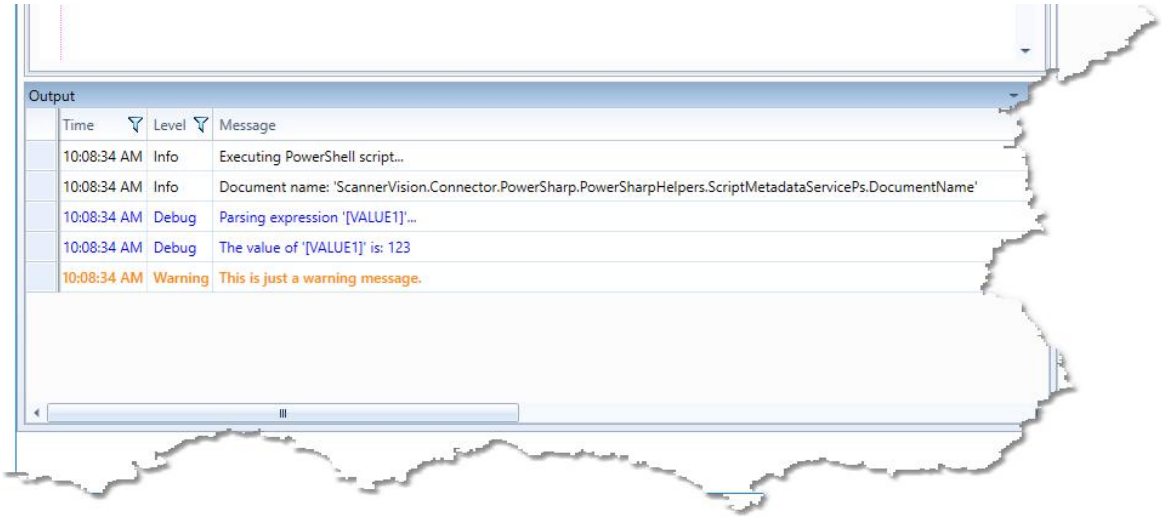


8.6.5.4.3 PowerShell Example

Below is shown how to call the ScriptMetadataService functions.

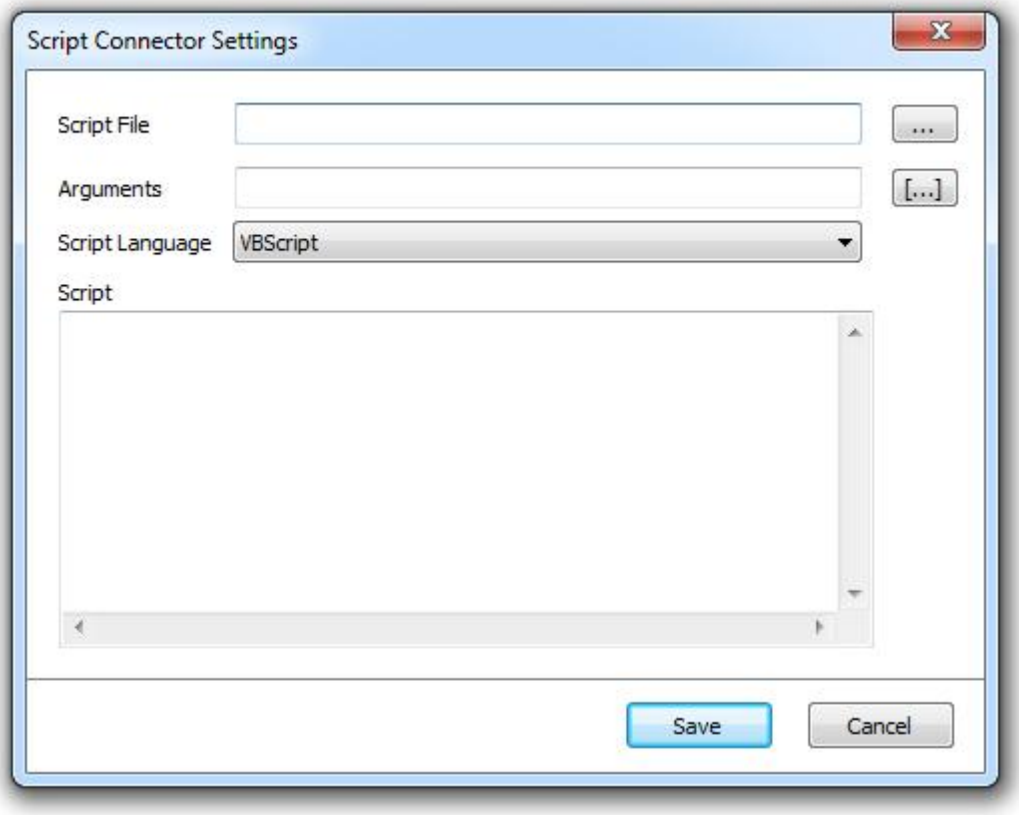
```
$ScriptMetadataService.Log("Executing PowerShell script...");
$ScriptMetadataService.Log("Document name: '$ScriptMetadataService.DocumentName'");

$ScriptMetadataService.Debug("Parsing expression '[VALUE1]'...");

$Value1 = $ScriptMetadataService.MetadataService.ParseExpression("[VALUE1]");

$ScriptMetadataService.Debug("The value of '[VALUE1]' is: $($Value1)");

$ScriptMetadataService.Warning("This is just a warning message.");
```

Here is the output:



8.6.5.5 Script Connector

The Script connector executes VBScript or JScript scripts. A script could either be loaded from an external file or it could be entered into the "Script" edit field directly.



Anywhere that you see the "[...]" button you can make use of ScannerVision metadata tags.

Script File Name

The fully qualified file name of the script.

Arguments

Script arguments are any values that you need to send through to your script other than ScannerVision metadata. You can send as many arguments as you need. Arguments are separated by one space character followed by one forward slash (" /"). You therefore cannot have " /" in your data. The first argument is indicated with a forward slash without the leading space i.e. "/First argument". Here is an example of passing and using arguments:

```
/c:\test.txt /Argument 2 /Another argument with/hashes
```

Given the following VB script:

```
outFile = Arguments.Item(0)

Set objFSO = CreateObject("Scripting.FileSystemObject")
Set objFile = objFSO.CreateTextFile(outFile, True)

objFile.Write Arguments.Item(1) & vbCrLf & Arguments.Item(2)
objFile.Close
```

a file called "test.txt" is created in the root of drive "C" with the contents:

```
Argument 2

Another argument with/hashes
```

Arguments are referenced through the "Item" method of the "Arguments" object. The Item method takes the zero based index of the argument you want to reference.

Notice that since the hash used in the 3rd argument does is not preceded by space character it forms part of the 3rd argument and consequently appears in the output.

There is no need to feed ScannerVision metadata into your script through arguments since you can access them directly in your script through the "Metadata" object described below.

Script language

Select the language of your script. VBScript and Jscript are supported.

Script

If you don't want to make use of an external file you can enter your script here.

Metadata

ScannerVision metadata can be referenced in your script through the "Values" method of the "Metadata" object. The Values method takes a string parameter that represents the metadata tag you want to reference. To reference the [DATETIME] tag you would write the following:

```
Metadata.Values("DATETIME")
```

You can also modify or create new metadata as follows:

```
Metadata.Values("WFSFILENAME") = "MyDate.txt"
```

If a value is assigned to a metadata tag that does not exist it will be created.

Note

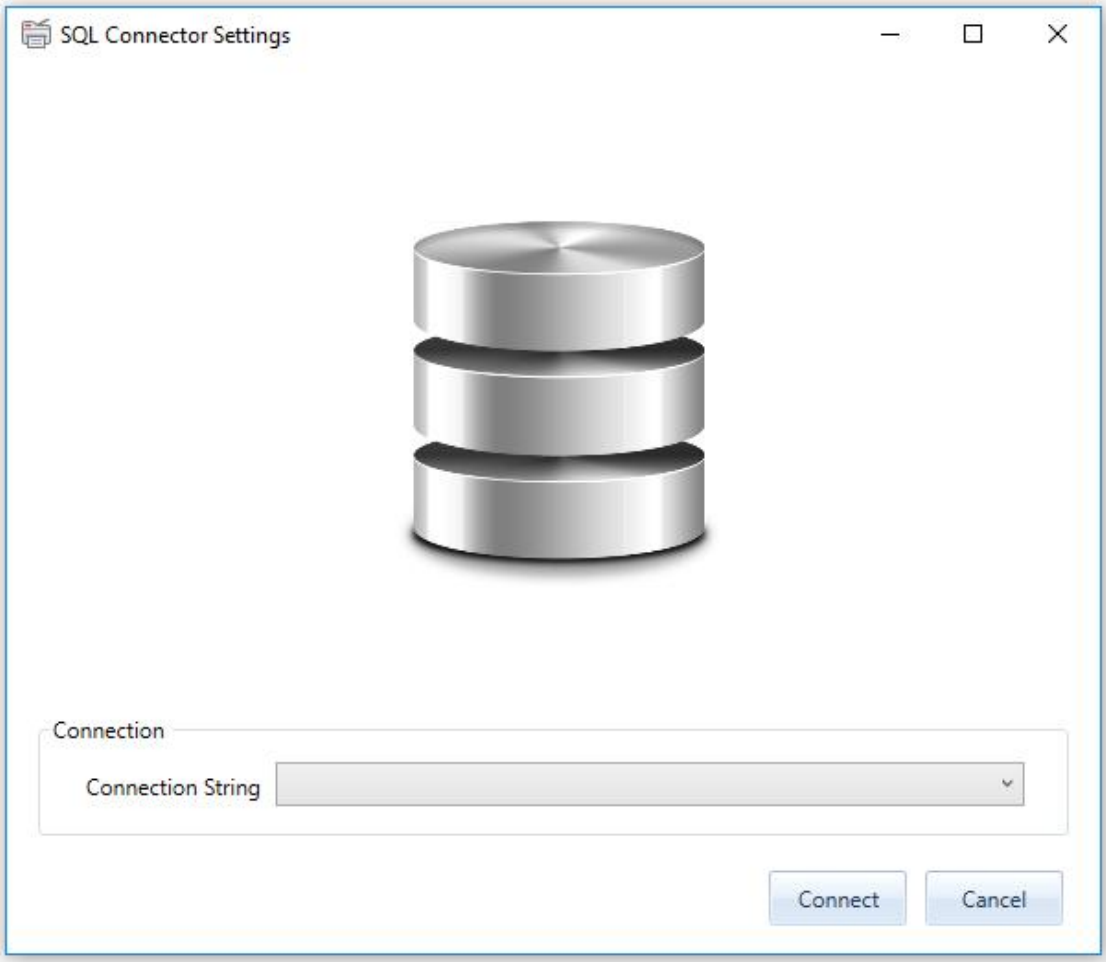
Only metadata tags that have been published by connectors such as the [WFSFILENAME] tag - which is published by the Windows File System connector - can be modified by a script.

8.6.5.6 SQL Connector

The SQL connector you can perform the following tasks:

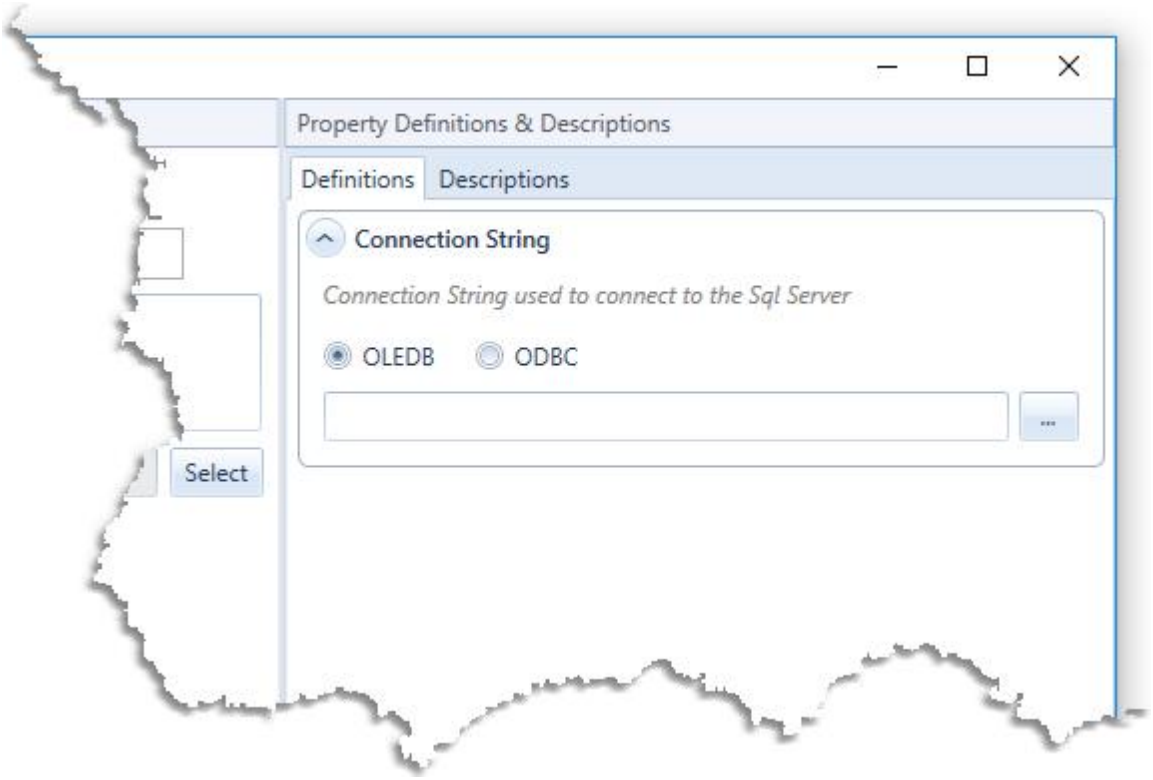
1. Select fields from database tables and publish those fields as ScannerVision metadata for use by other connectors or for export.
2. Execute a query to determine if a record exists.
3. Depending on the outcome of 2. above insert a new record or update an existing one.

The SQL connector connection screen is shown below:



Anywhere that you see the "[...]" button you can make use of ScannerVision metadata tags.

Before you can configure a new SQL connector you have to configure a Connection. The connection settings screen is shown below:



Select type of connection you want to make and then click on the "..." button to open the appropriate connection editor dialog box where you configure the database connection string or select an ODBC data source. Please refer to Appendix C - Database Connection Strings for more information.

After you have configured the connection, selected the admin user, configured the permissions for the admin user and authenticated him/her, you are ready to configure a SQL connector.

Connection String

Select the Connection instance you want to use for this connector.

Once you have configured the connection string click the "Connect" button. This will take you to a screen where you can configure:

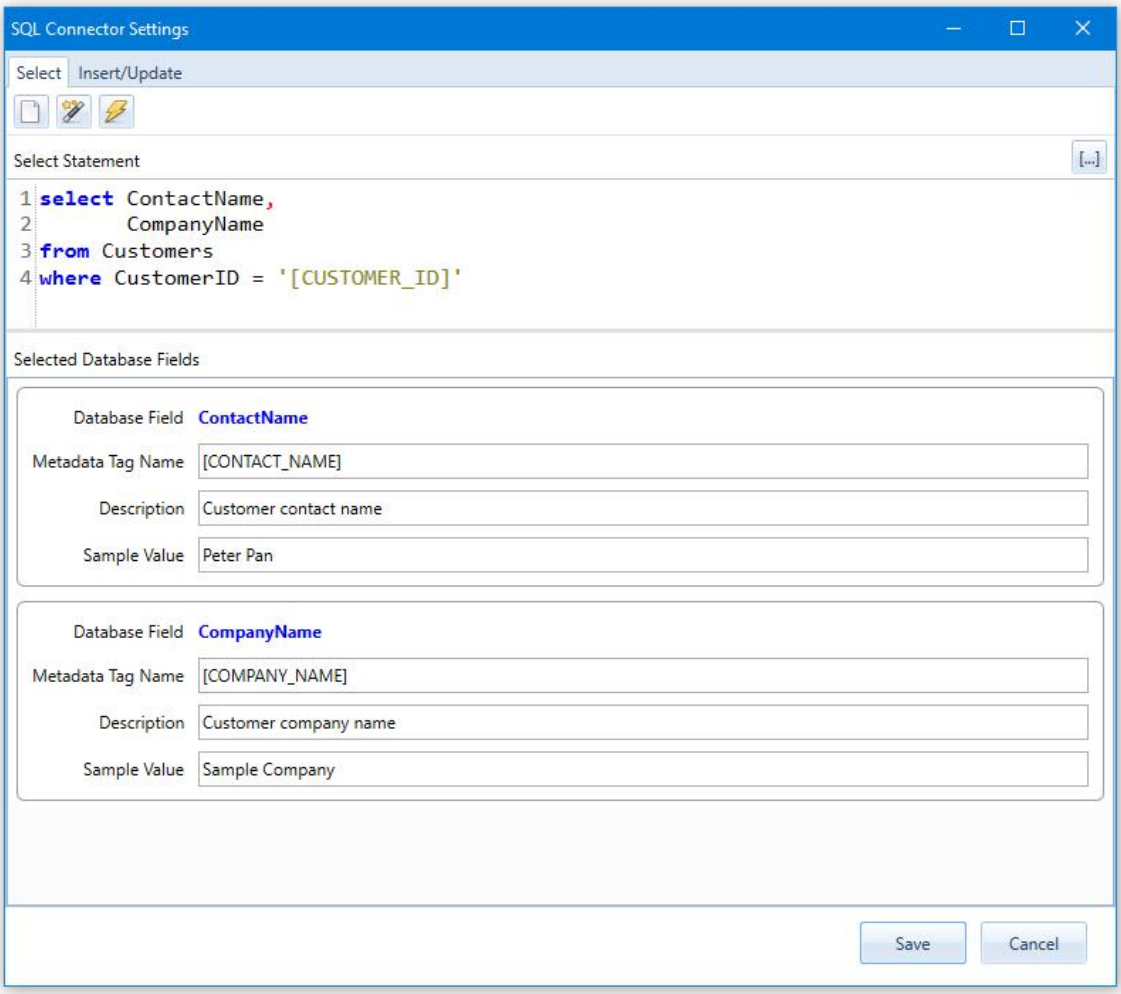
1. The Select statement for publishing fields from the database as metadata.

2. The Insert/Update statements to update or insert data in the database.

8.6.5.6.1 Select

The Select query screen is shown below. To configure the connector for selecting and publishing database fields follow these steps:

1. Enter the SELECT SQL statement in the "Select Statement" window.
2. Press the execute ( ) button.
3. The fields in the "Select Statement" will be added to the "Selected Database Fields" list.
4. Provide values for "Metadata Tag Name" (required), "Description" and "Sample Value" (optional).



Toolbar

Icon	Description
	Clears the Select Statement, Sample Values and Database Fields.

Icon	Description
	Generates a generic SELECT SQL statement.
	Executes the Select Statement.

Select Statement

Any legal SELECT SQL statement. The statement may include ScannerVision metadata tags in the WHERE clause.

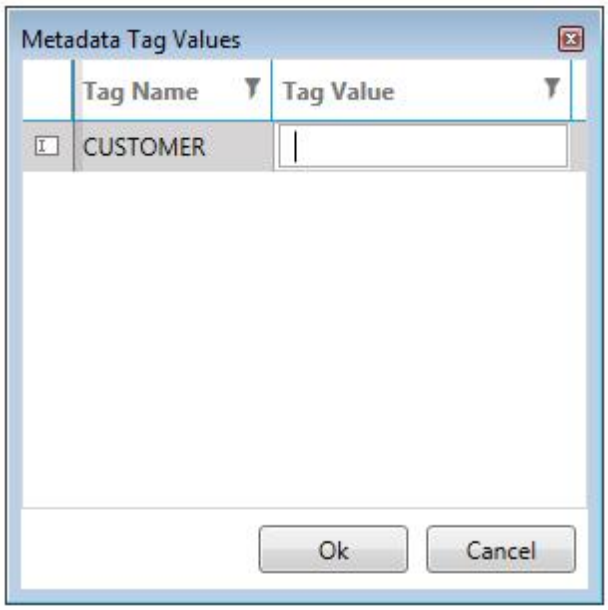
Selected Database Fields

A list of fields selected in the "Select Statement".

Database Field	The name of the database field or alias as entered in the SELECT statement.
Metadata Tag Name	The name of the metadata tag under which the value of the database field will be published.
Description	A description of the metadata tag. This is used as a tool tip in the ScannerVision Expression Editor.
Sample Value	A sample value for the metadata tag. This is used as the value for the tag in the ScannerVision Expression Editor.

Providing temporary metadata tags

If you have referenced ScannerVision metadata tags in your query you will be prompted to provide temporary values for the metadata tags you've used when updating the columns or testing the picklist:



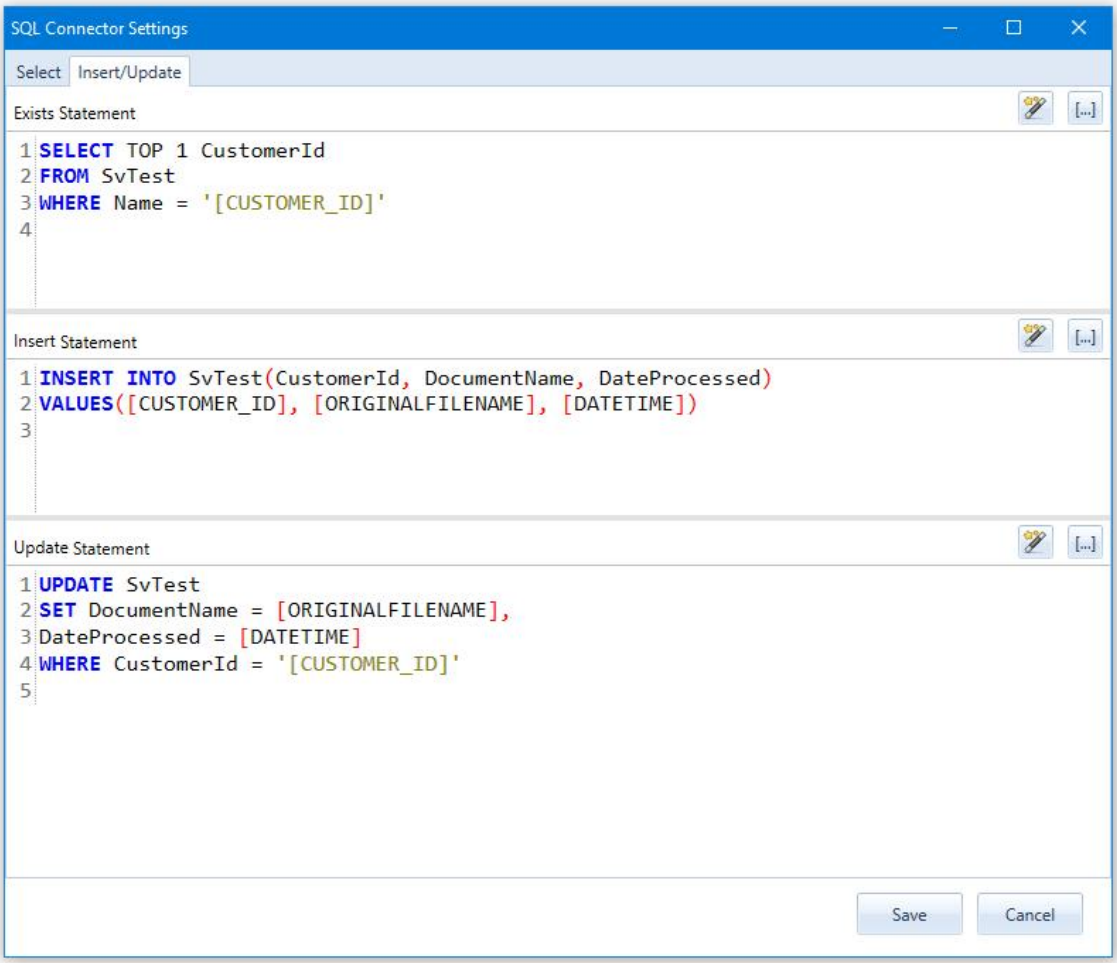
The values you specify here are used to by ScannerVision to substitute for the used metadata tags. If you enter legitimate values then the query should execute successfully.

8.6.5.6.2 Insert/Update

The Insert/Update query screen is shown below. To configure the connector for inserting or updating records follow these steps:

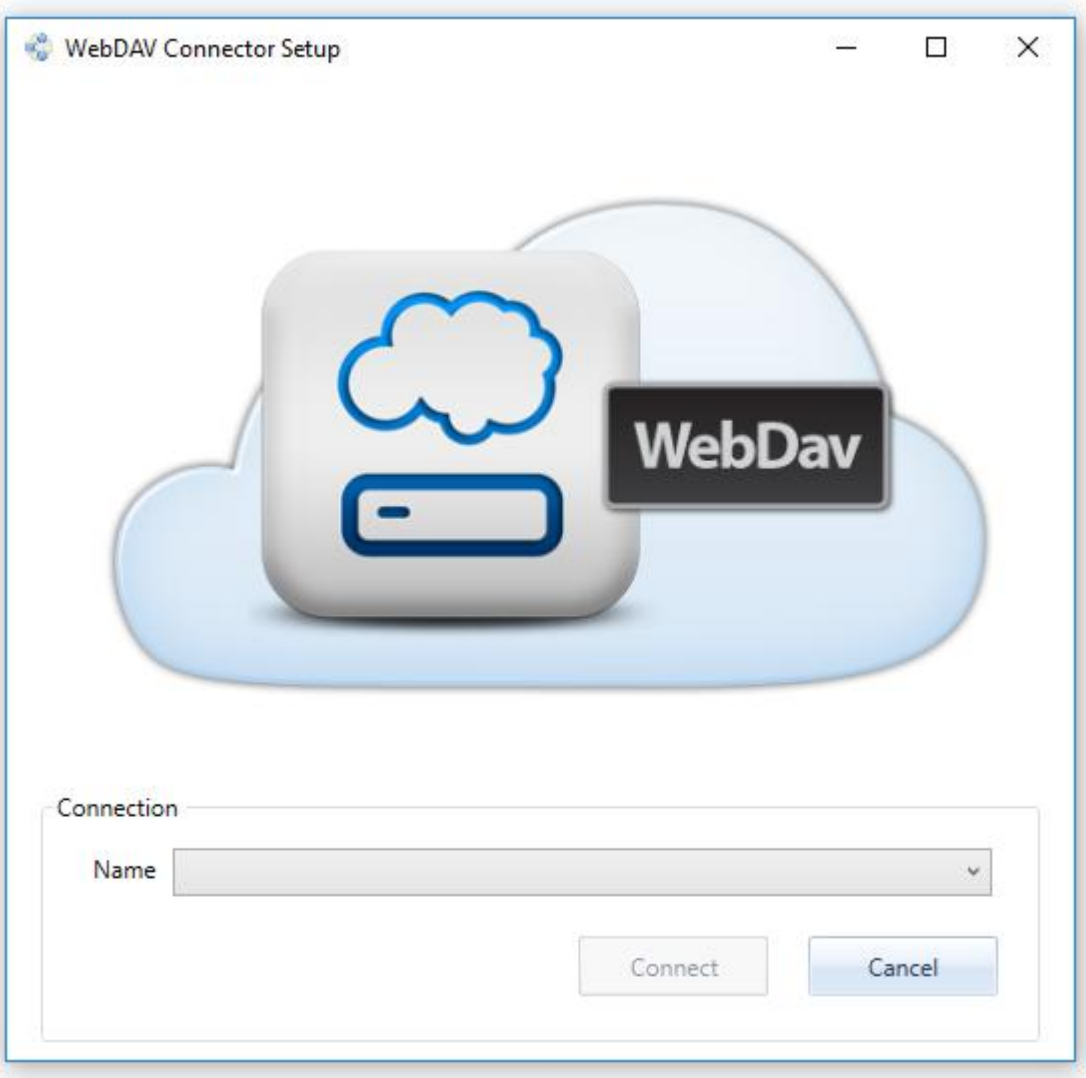
1. Enter the SELECT SQL statement in the "Exists Statement" window. The result of this query should return one or zero records. If the query returns more than one record the connector upload will fail.
2. Enter the INSERT statement in the "Insert Statement" window.
3. Enter the UPDATE statement in the "Update Statement" window.

The  buttons above each SQL statement window inserts a generic version of the respective SQL statement into the query editor. The "[...]" button displays the ScannerVision Expression Editor.



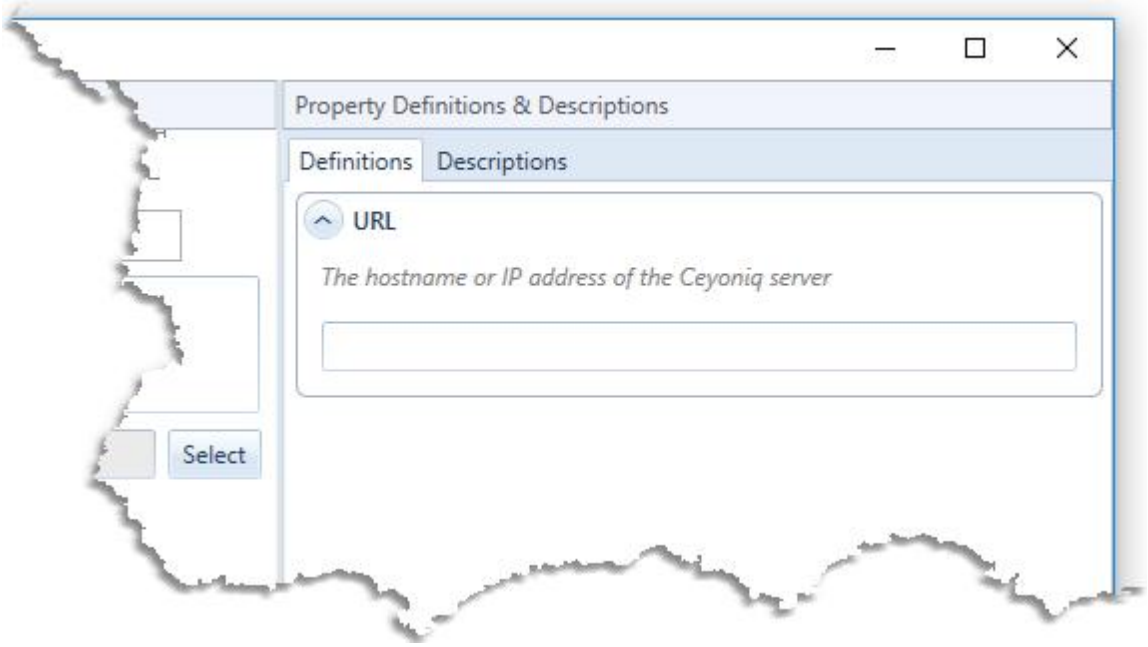
8.6.5.7 WebDAV Connector

The WebDAV connector uploads documents to a WebDAV enabled web server. The connector supports SSL encryption.



Anywhere that you see the "[...]" button you can make use of ScannerVision metadata tags.

Before you can configure a new WebDAV connector you have to configure a Connection. The connection settings screen is shown below:



URL

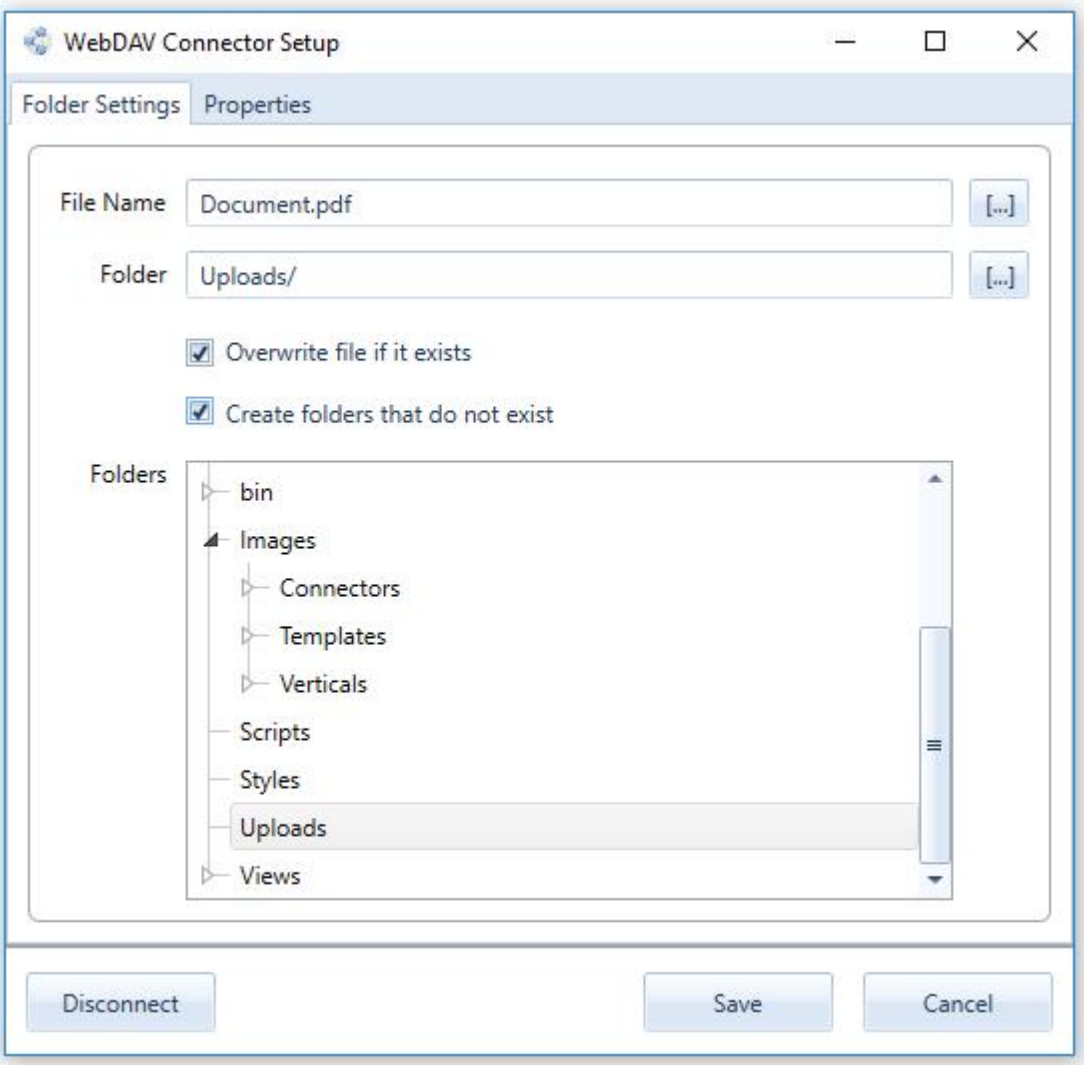
The URL of the web server. To secure the communication with the WebDAV server, specify the **https** protocol.

Please refer to Appendix I - SSL Certificates if you are presented with a SSL certificate warning dialog.

After you have configured the connection, selected the admin user, configured the permissions for the admin user and authenticated him/her, you are ready to configure a SQL connector.

8.6.5.7.1 Folder Settings

After you have selected a connection click the "Connect" button which will take you to the following configuration screen:



File Name

The file name of the document that will be stored on the WebDAV server.

Folder

The folder on the WebDAV server in which the document needs to be stored. Please ensure that the user connecting to the WebDAV server has sufficient file system permissions to create files in the selected folder.

Overwrite file if it exists

When this option is selected existing files will be overwritten.

Create folders that do not exist

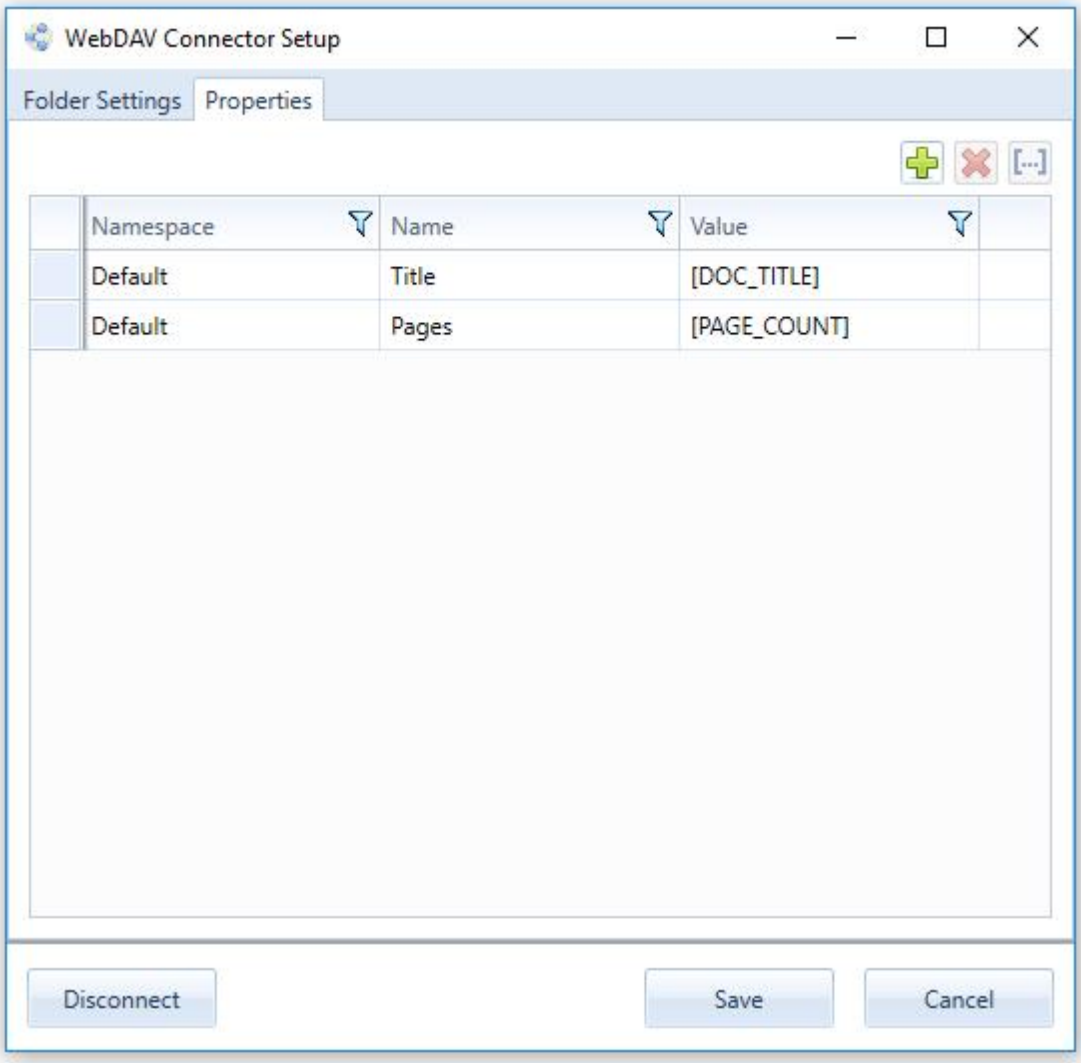
When this option is selected folders will be created automatically if they do not exist.

Folders

The folder view presents a list of the available folders on the web server to which documents could be uploaded. Double click the folder to select it as source.

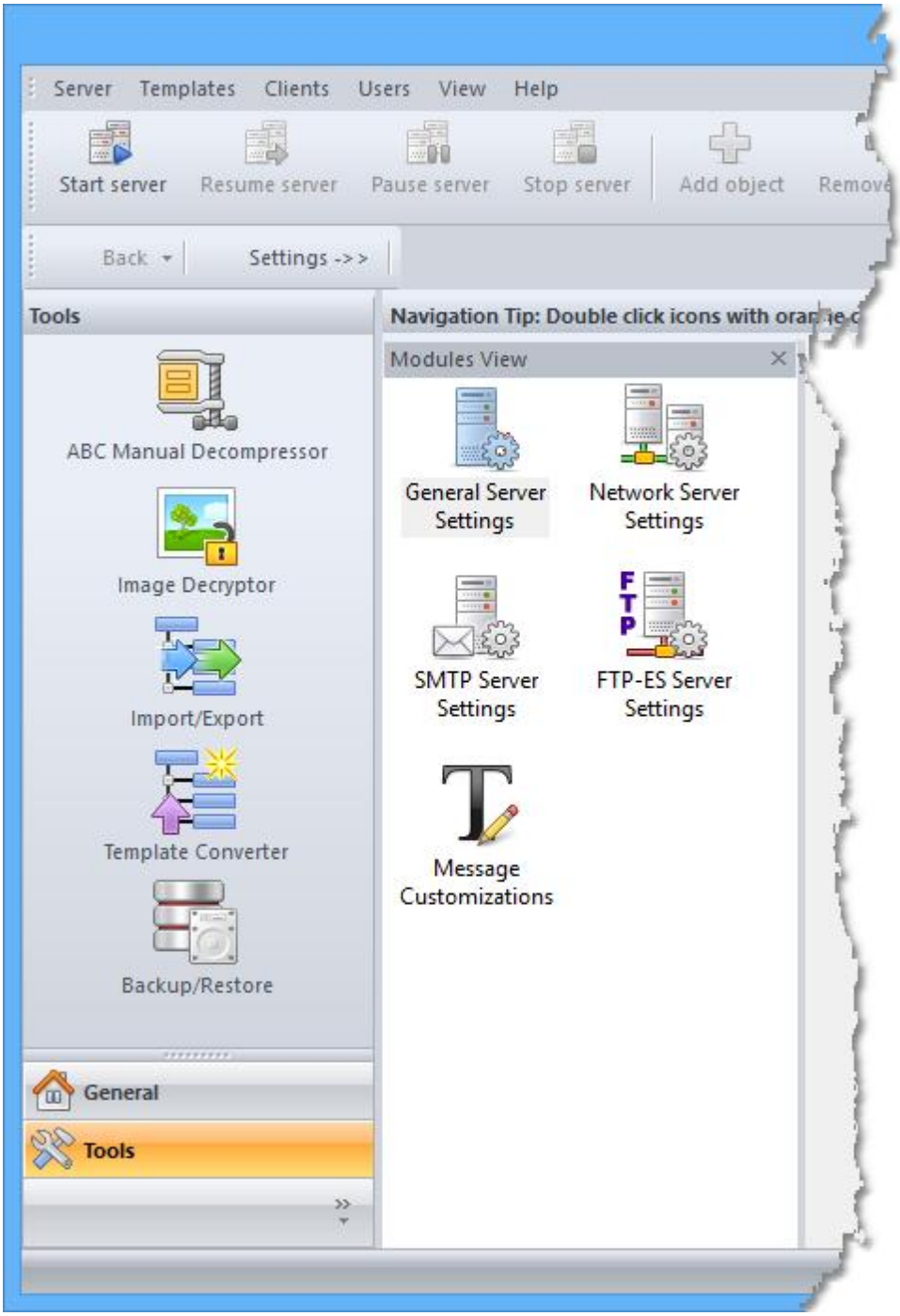
8.6.5.7.2 Properties

On the properties screen shown below you can configure properties to be set on the uploaded document. Refer to your WebDAV server's documentation for further details.



9 Tools

ScannerVision includes several stand-alone tools that can be launched via the ScannerVision Tools menu, from the Windows Start Menu or directly from the ScannerVision installation folder. The ScannerVision Tools menu is shown below:



- ABC Manual Decompressor
- Image Decryptor
- Import/Export
- Template Converter
- Backup/Restore

9.1 ABC Manual Decompressor

Documents that were compressed using ScannerVision's ABC Compression can be decompressed using the ABC Manual Decompressor command line utility. To decompress a document launch the ABC Manual Decompressor utility and select the compressed file.

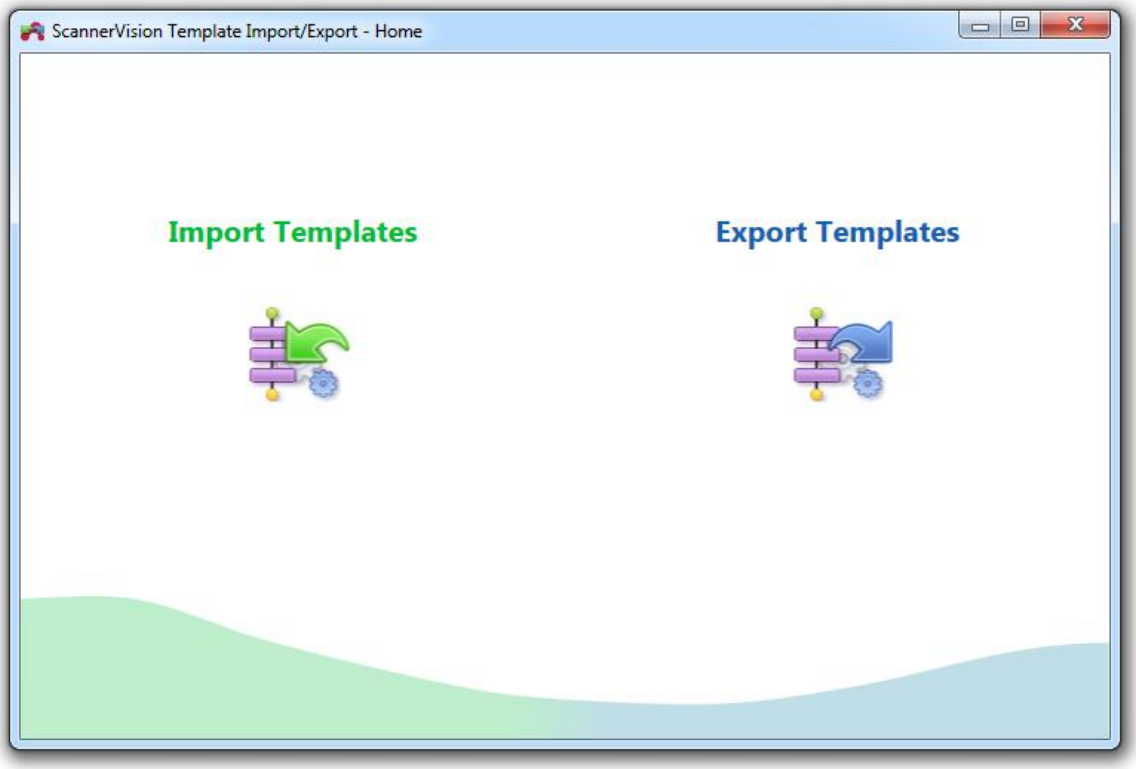
9.2 Image Decryptor

Images encrypted using the "Encrypt" function in ScannerVision's Annotation module can be decrypted using the Image Decryptor command line utility. When the Image Decryptor is launched you have to select the document you want to decrypt. If the Image Decryptor cannot a file with the same name as the document you have selected but with an ".ann" extension you will be prompted to select the .ann file manually. If you don't have the .ann file you will not be able to decrypt the file.

After you have selected the document to be decrypted you have to select the name and properties of the decrypted document after which the document is decrypted.

9.3 Template Import/Export

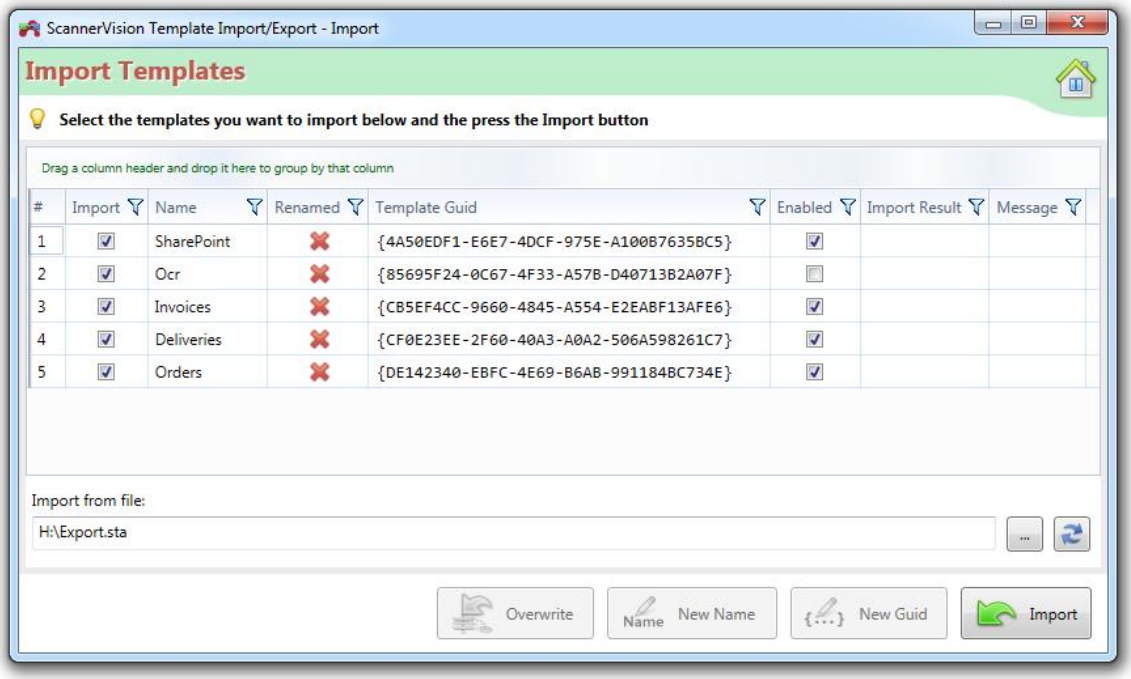
When ScannerVision templates reference external files such SQL or Visual Basic or JScript scripts they need to be copied with the template xml file in order for the template to be usable elsewhere. The paths to these files also have to be updated to reflect their new locations. The ScannerVision Template Import/Export (IE) application facilitates the transfer of templates and their dependency files by packaging them into a single template archive. The Import/Export application also provides a command line mode so that it can be used for scheduled backups of templates. The home screen of the IE application is shown below:



From the Home screen you can choose if you want to Import or Export templates.

9.3.1 Import

When a template archive file has been opened, the Import screen lists all the templates contained in the archive as shown below:

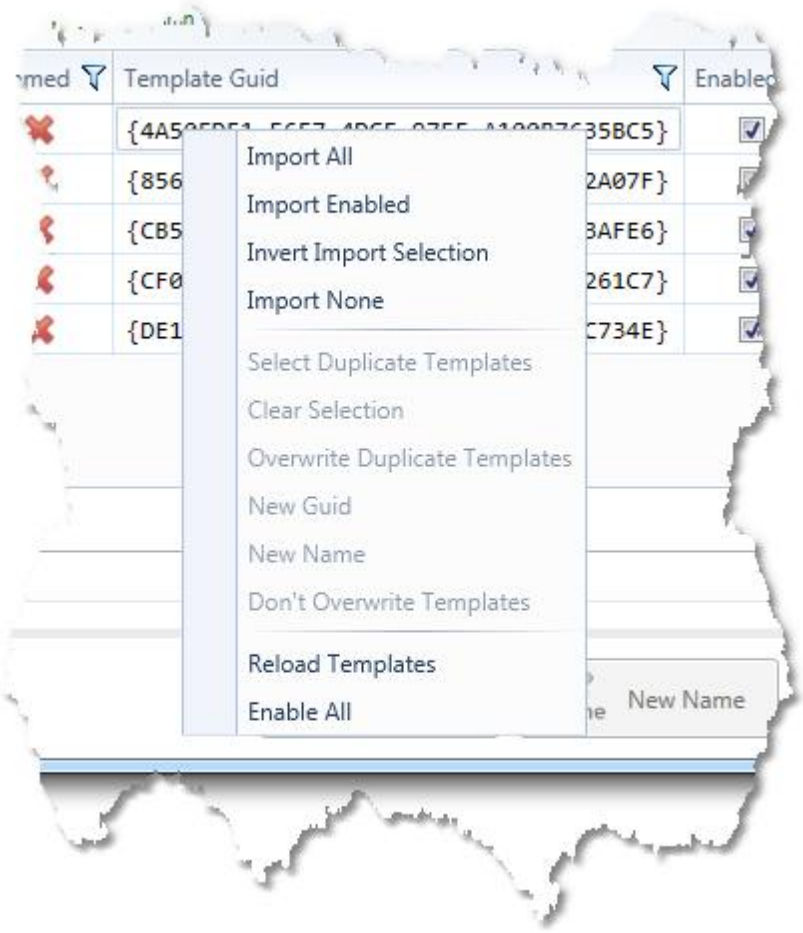


The grid contains 8 columns:

Column	Description
#	Template number
Import	Select which templates must be imported.
Template Name	The name of the template as configured in ScannerVision. You can change the name of the template by editing the template name before you import the templates.
Renamed	Indicates if the template has been renamed because of a naming conflict with an existing template. See the Duplicate Templates section for more information.
Template Guid	The unique identifier of the template that is generated by ScannerVision when the template is created
Enabled	Indicates whether the template was enabled or disabled on the machine on which the template archive was created. Templates can be enabled or disabled before importing them.
Import Result	Indicates if the template was imported successfully

Message	Shows an import status message after the template import process is complete

The grid's context menu:

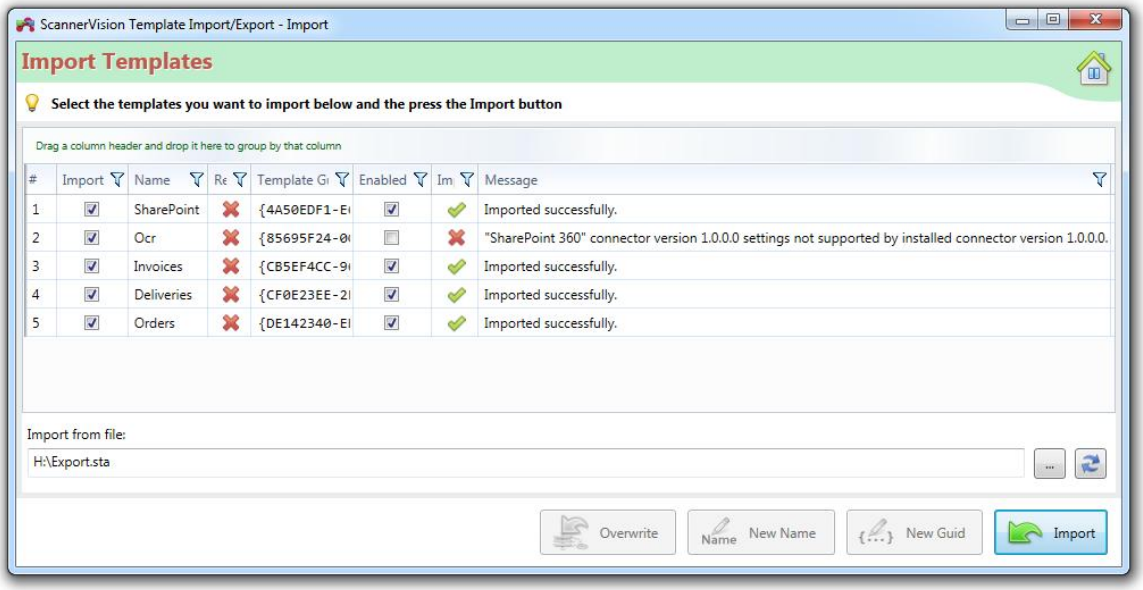


contains the following commands:

Command	Description
Import All	Marks all templates for import
Import Enabled	Marks only templates that are enabled for import
Invert Selection	Inverts the current import selection

Select None	Clears import selection of all templates
Grayed out items	The grayed out items are discussed in the Duplicate Templates section
Reload Templates	Reloads templates from the template archive file
Enable All	Enables all templates

Once you have selected the templates you want to import click the "Import" button to start the import process. Below is a screen shot of the Import screen after the import process has completed:



You will notice that one of the templates could not be imported because the settings of the SharePoint 360 connector as exported originally is not supported by the version of the connector that is currently installed on the system. You have to ensure that the connectors installed on your system are compatible with the settings of those on the original system.

An import log is generated for every import in the "C:\ProgramData\ScannerVision\Logs\Template Import" folder on machines with Windows 7 and later installed. The name of the import log file has the format "Import - 24 Jun 2013, 09-52-11.log" where the date and time part is the date and time at the time of the import.

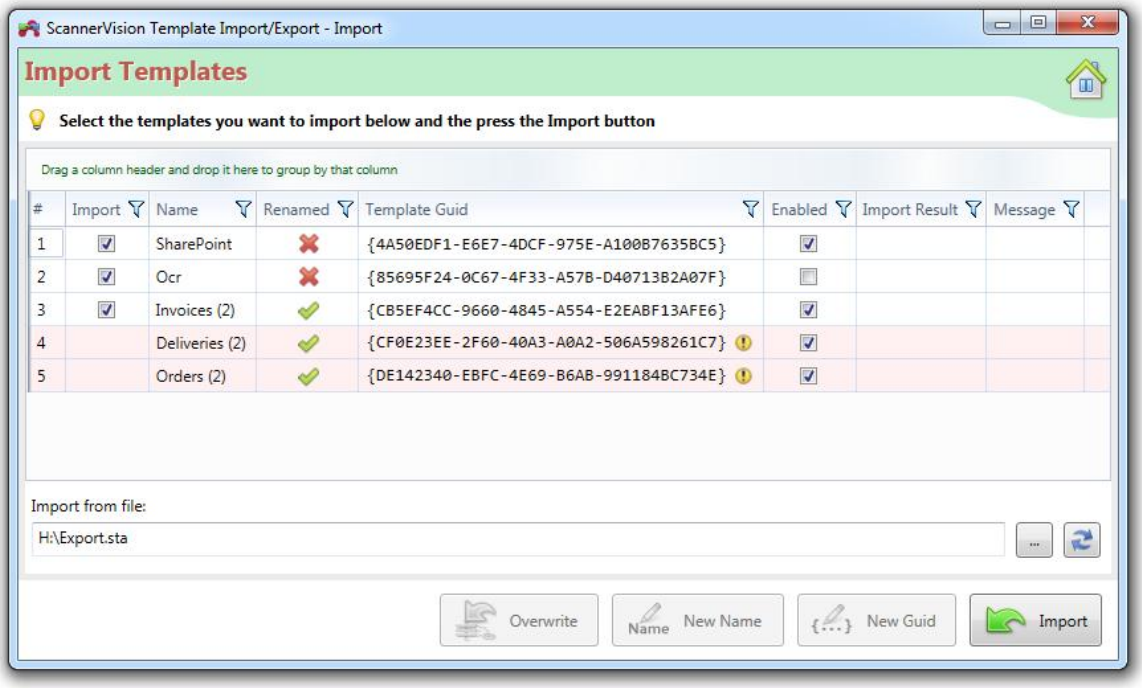
To return to the Home screen, click the "Home" button in the top right hand corner of the screen:

9.3.1.1 Duplicate Templates

When importing templates there is the possibility that one or more of the templates being imported may conflict with one already on the system. There are two possible points of conflict:

1. Guid conflicts will happen when templates are imported into the same system from which they were exported.
2. Name conflicts will happen when templates are imported that have the same name as ones already on the system.

The above conflicts are depicted in the screen shot below:



Guid Conflicts

When the IE application encounters a template Guid conflict (rows 4 & 5) several things happen:

1. The import check box is removed.
2. If the template's name conflicts with an existing template's (which would be true in most cases if the Guid is in conflict), the template is automatically renamed. The way in which the renaming is done is discussed below.
3. A check mark appears in the "Renamed" column of the template's row.
4. The background the template's row turns to red.

5. An exclamation mark appears after the Guid to indicate that there is a conflict.

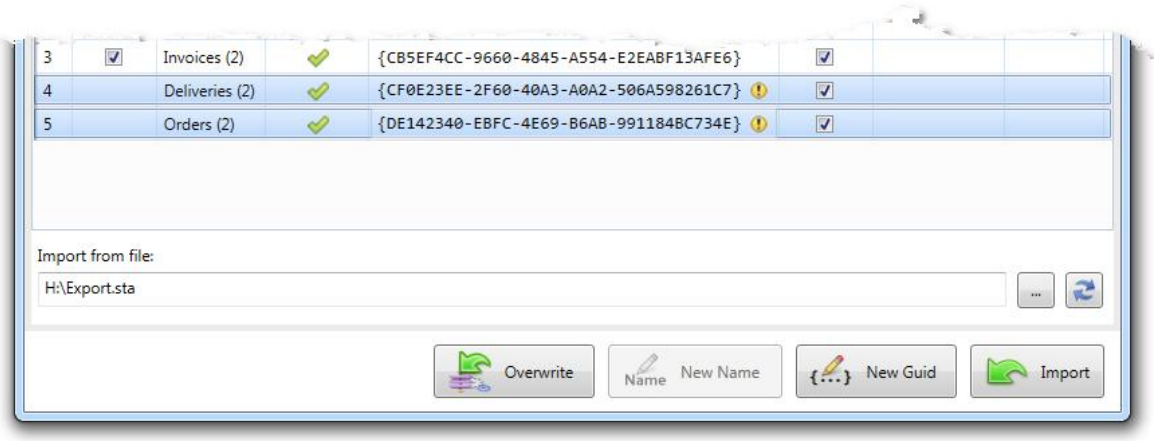
You have three choices to resolve a Guid conflict:

1. Don't import the template. This is the default option.

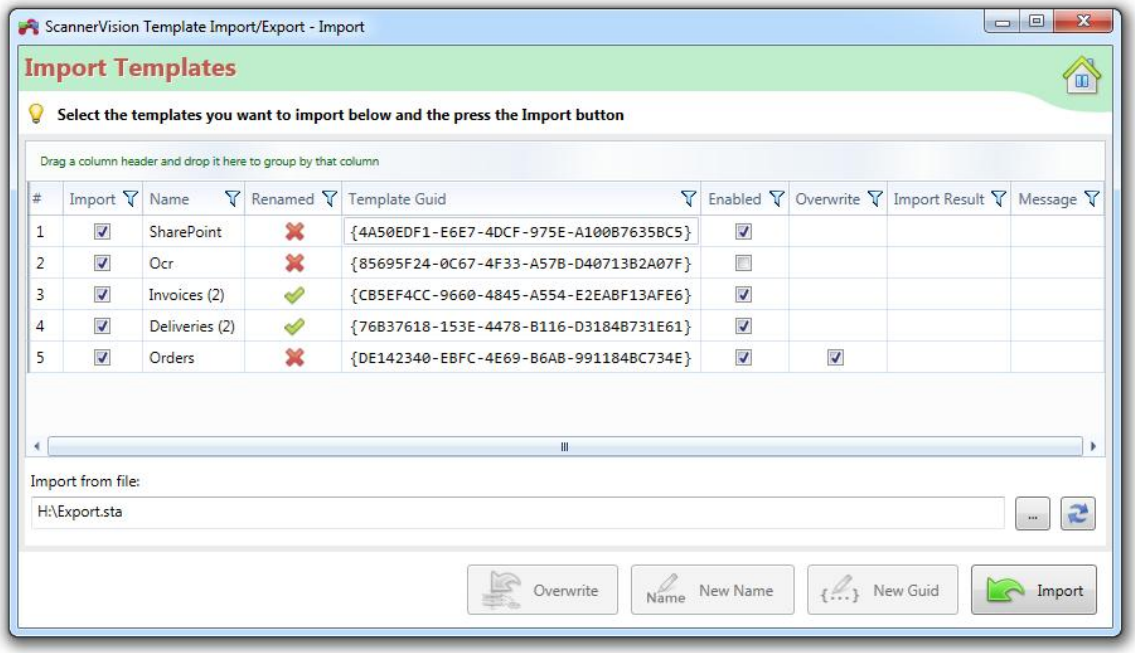
2. Resolve the conflict by generating a new Guid for the template.

3. Overwrite the existing template with the imported one.

To affect either of the choices 2. or 3. above, select the relevant templates in the grid by clicking their respective rows or by selecting the "Select Duplicate Templates" item in the grid's context menu (right click on the grid to show the context menu). Selected rows have a blue background like this:



Notice that the "Overwrite" and "New Guid" buttons enable when one or more duplicate rows are selected. The screen shot below shows the result of generating a new Guid for row 4. and overwriting the existing template of row 5.



Notice that when a new Guid is generated for a template it is still renamed. If an existing template is overwritten the original template name is retained.

Template Name Conflicts

When a template name conflict is encountered the following things happen:

1. The template is renamed.

2. A check mark appears in the "Renamed" column of the template's row.

If a template in renamed manually to a name that conflicts with another template in the template archive or with one on the system, an exclamation mark appears behind the name of the conflicting template. In the screen shot below the template in row 3. has been renamed from "Invoices (2)" (which was the name the IE application generated automatically) to "Invoices". A template with the name "Invoices" already exists on the system hence the exclamation mark. The template in row 4. was renamed from "Deliveries (2)" to "Ocr". There is already a template called "Ocr" in the template archive therefore both templates (rows 2. and 4.) are flagged as being duplicates with the exclamation mark.

To resolve name conflicts you can manually rename the relevant templates or you can select them and click the "New Name" button.

Auto Renaming

When the IE application generates a new name for a template, it adds a number in brackets to the end of the name. "Invoices" will therefore become "Invoices (2)". If the new name still conflicts with another template, the number is incremented until a unique file name is obtained e.g. "Invoices (3)", "Orders (4)" etc.

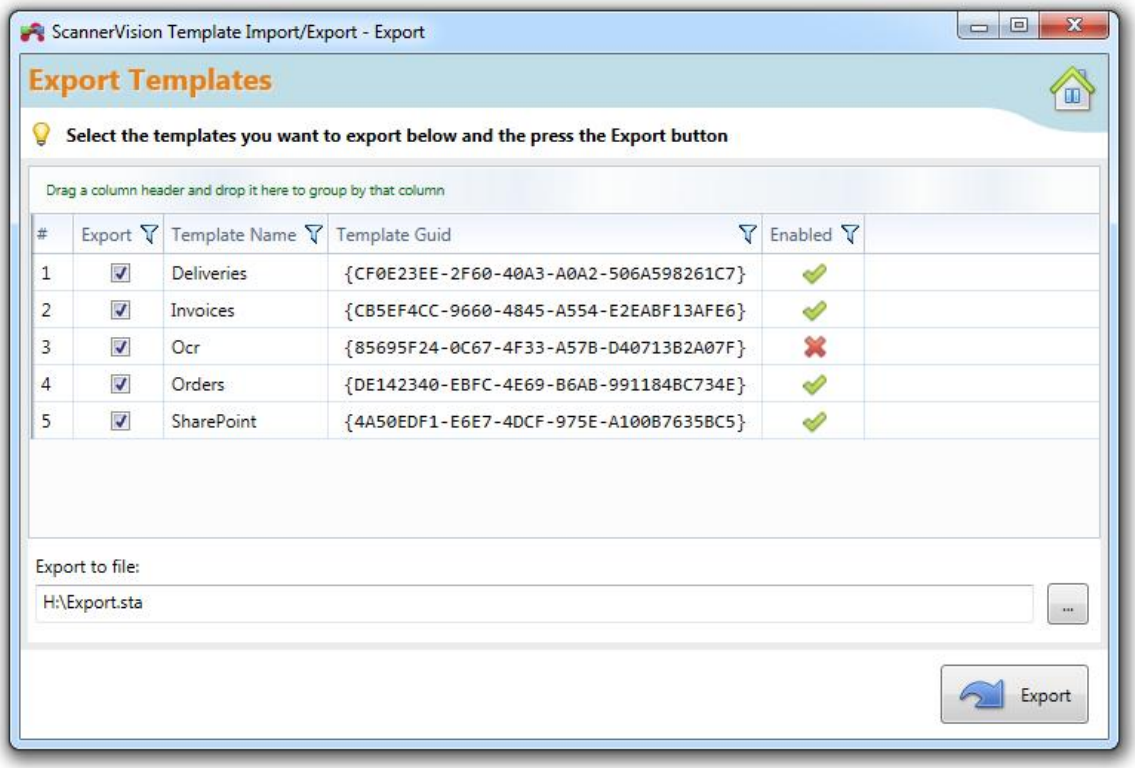
9.3.1.2 Importing .ini templates

Support for importing old .ini based template has been removed from ScannerVision V7. In order to import these templates into ScannerVision V7 or later follow these steps:

1. Install ScannerVision V6. No license is required to perform the steps below.
2. Import the template archive (.sta file) into ScannerVision V6.
3. Export the templates again and create a new template archive - which will now contain the converted .xml templates.
4. Import the template archive (.sta file) into ScannerVision V7 or later.

9.3.2 Export

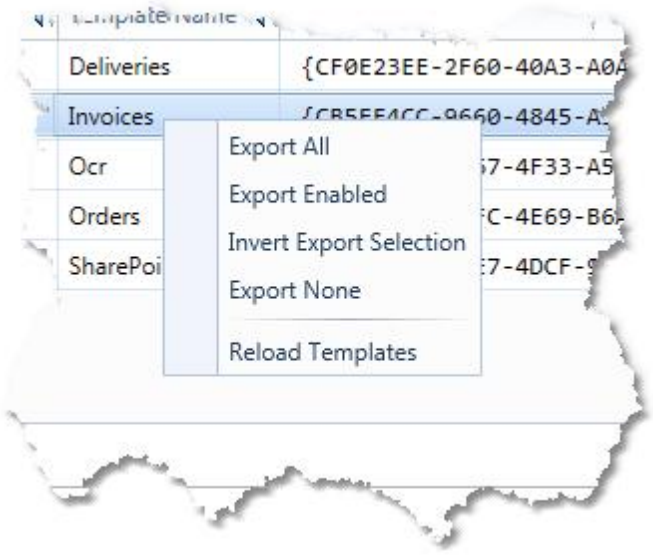
The Export screen lists all the templates in the ScannerVision Templates folder as shown below:



The grid contains 5 columns:

Column	Description
#	Template number
Export	Select which templates must be exported
Template Name	The name of the template as configured in ScannerVision
Template Guid	The unique identifier of the template that is generated by ScannerVision when the template is created
Enabled	Indicates whether the template has been enabled or disabled in ScannerVision

The grid's context menu:



contains the following commands:

Command	Description
Export All	Marks all templates for export
Export Enabled	Marks only templates that are enabled for export
Invert Export Selection	Inverts the current export selection
Export None	Clears export selection of all templates
Reload Templates	Reloads templates from the ScannerVision Templates folder

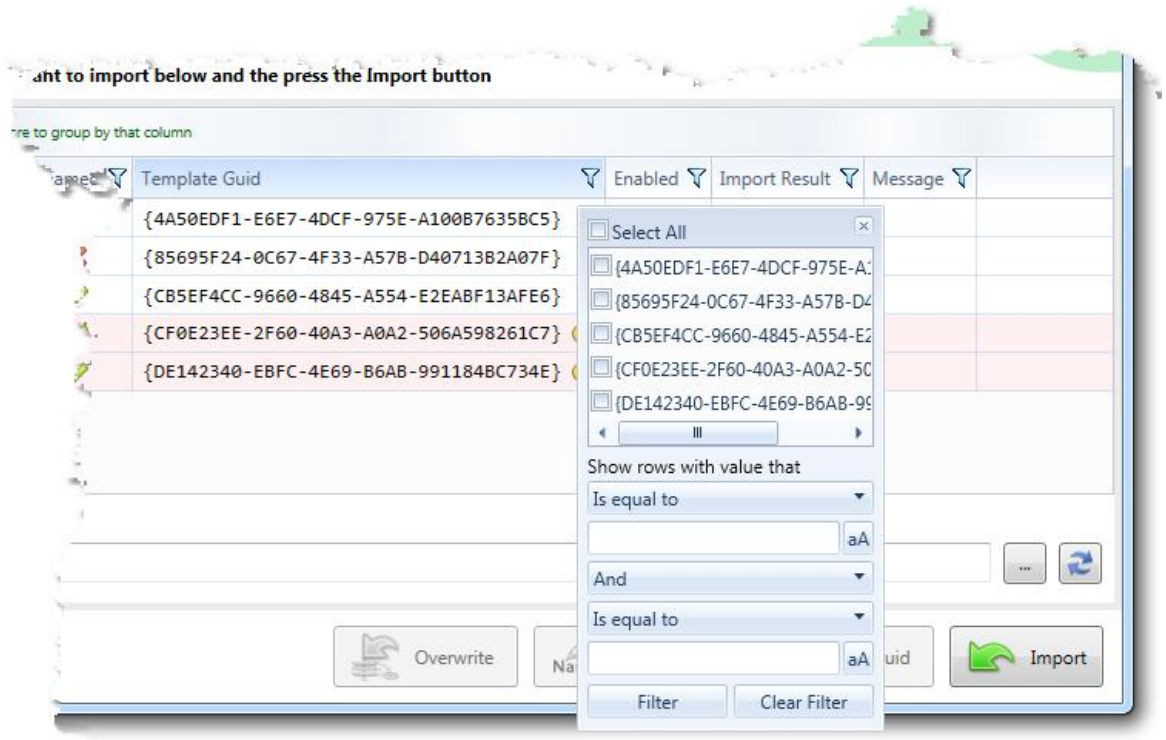
Once you have selected the templates to be exported, select the location and name of the template archive by typing a fully qualified file name into the "Export to file:" edit box. You can also use the "..." button to browse to the desired folder.

Click the "Export" button to start the export process.

To return to the Home screen, click the "Home" button in the top right hand corner of the screen:

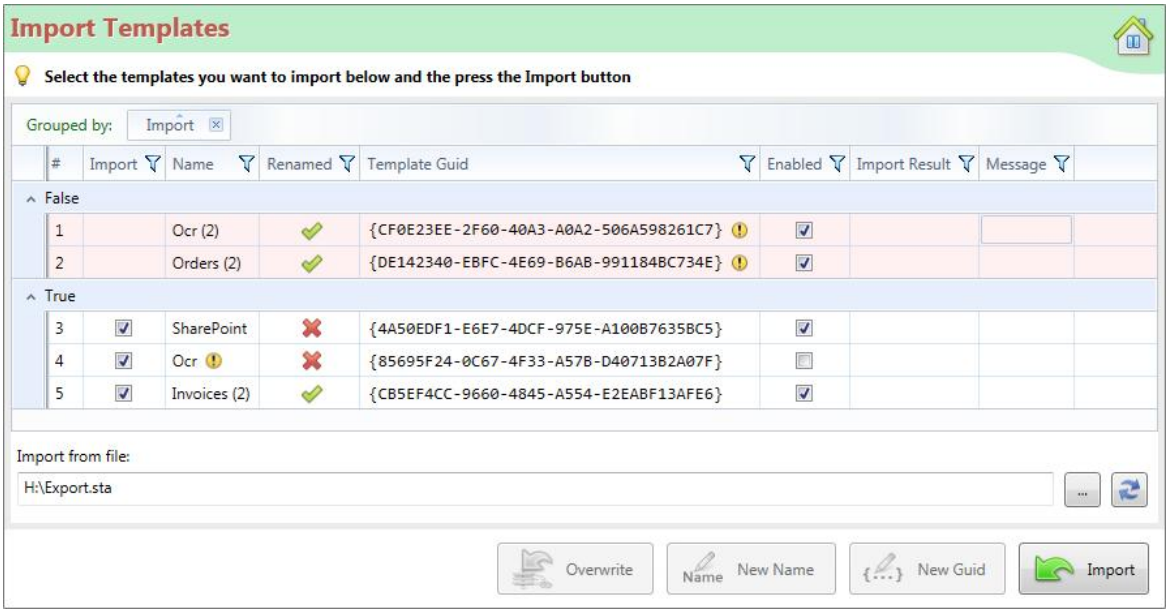
9.3.3 Filtering, Grouping and Sorting

The templates listed in the grids on the Import and Export screen can be sorted by any column with the icon in the header. If you click on the icon itself you are presented with filter menu as shown below:



Here you have many options for filtering the templates that appear in the grid.

You can also group templates by dragging the column header onto the area above the grid as shown below:



Here we have grouped the templates by the "Import" column. You can add as many groups as you want. To remove a grouping, click the "x" button in the group box.

9.3.4 Command Line

The "TemplateImporterExporter.exe" application supports a number of command line options. If any command parameter is passed to the application the normal GUI is not displayed and if the "/s" (silent) parameter is passed no user interface is shown at all.

When a command line switch indicates a file name is to follow e.g. /p, /l the format is as follows:

```
/p "c:\Path\Archive.sta"
```

The following command line switches are supported:

Switch	Description
/i	Perform import. Omit this switch to perform an export.
/p	Specify the fully qualified template archive file to import from or export to.
/l	Specify a fully qualified log file name.
/s	Silent - no console window is displayed.
/c	Specify the current culture/language information e.g. "en_US", "sp_ES" etc.
/o	Overwrite existing templates.

/n	Create new template Guids for duplicate templates. This switch has higher precedence than /o.

Examples

The following command performs a silent template export to the template archive "c:\Exports\Templates - 2013-06-25.sta" and outputs logging information to the file "c:\Exports\Export Log - 2013-06-25.log":

```
TemplateImporterExporter.exe /s /p "c:\Exports\Templates - 2013-06-25.sta" /l "c:\Exports\Export
```

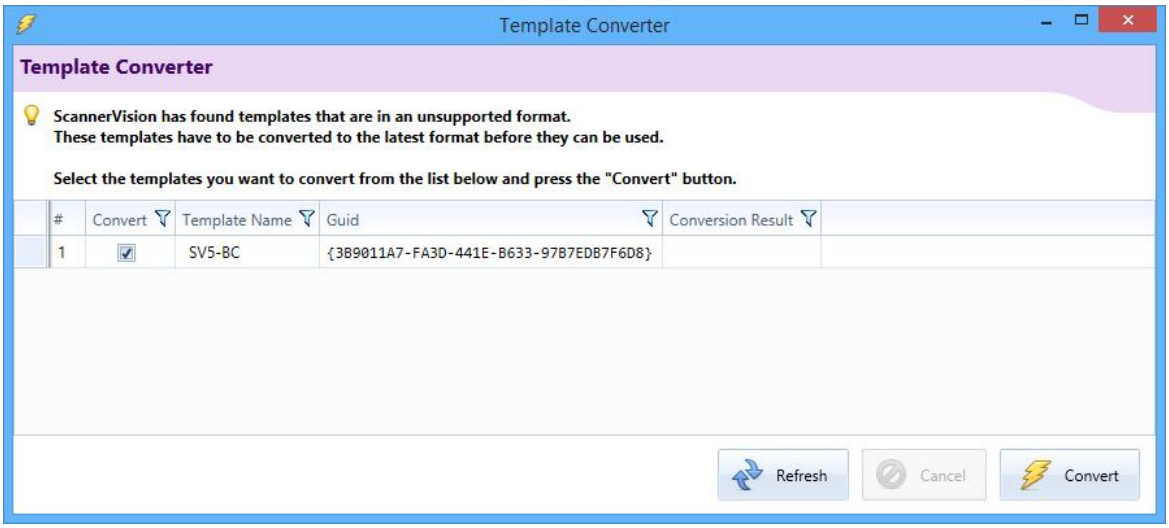
The following command performs a template import from the template archive "c:\Exports\Templates - 2013-06-25.sta" and outputs logging information to the file "c:\Exports\Import Log - 2013-06-25.log", duplicate templates are overwritten:

```
TemplateImporterExporter.exe /i /o /p "c:\Exports\Templates - 2013-06-25.sta" /l "c:\Exports\Impo
```

9.4 Template Converter

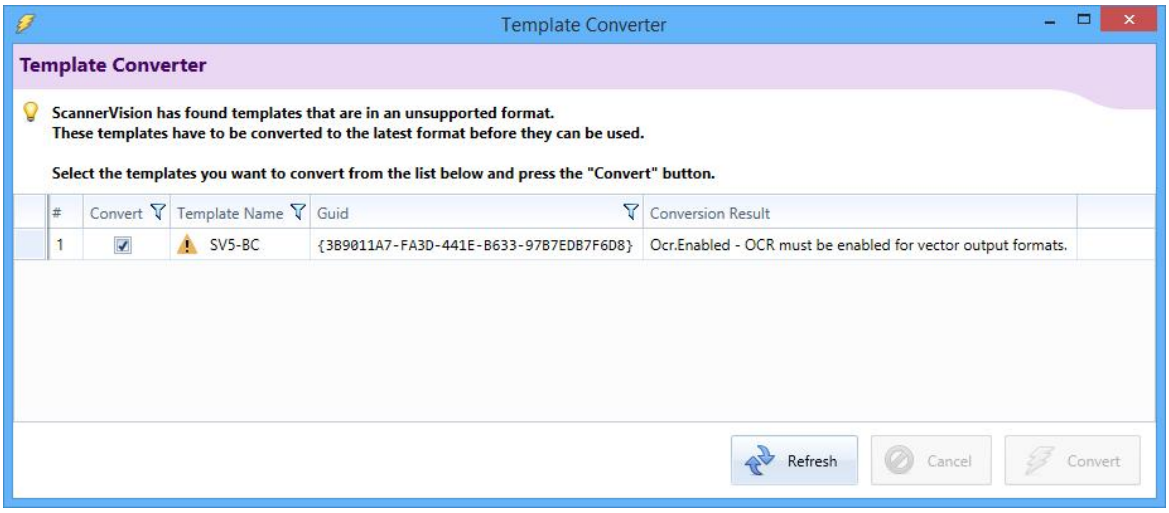
With the introduction of ScannerVision V6 the format in which templates are stored has changed from a multi-file ini and binary format to a single Xml based file.

When the Processing Engine User Interface is started it checks to see if there are any templates in ScannerVision's templates directory and if there are it launches the Template Converter utility shown below:



The Template Converter lists all the unconverted templates that exist in the templates folder. Select the templates that you want to convert and press the "Convert" button.

Take note of the "Conversion Result" column after the conversion process has completed. If there are problems with the template a description of the problem is shown here. In the screen shot below you can see that the template was converted but with a warning. This does not mean that the template was not converted, only that there is a problem with the template that needs to be corrected. When you open this template in ScannerVision and try to save it, the same error message will be shown.

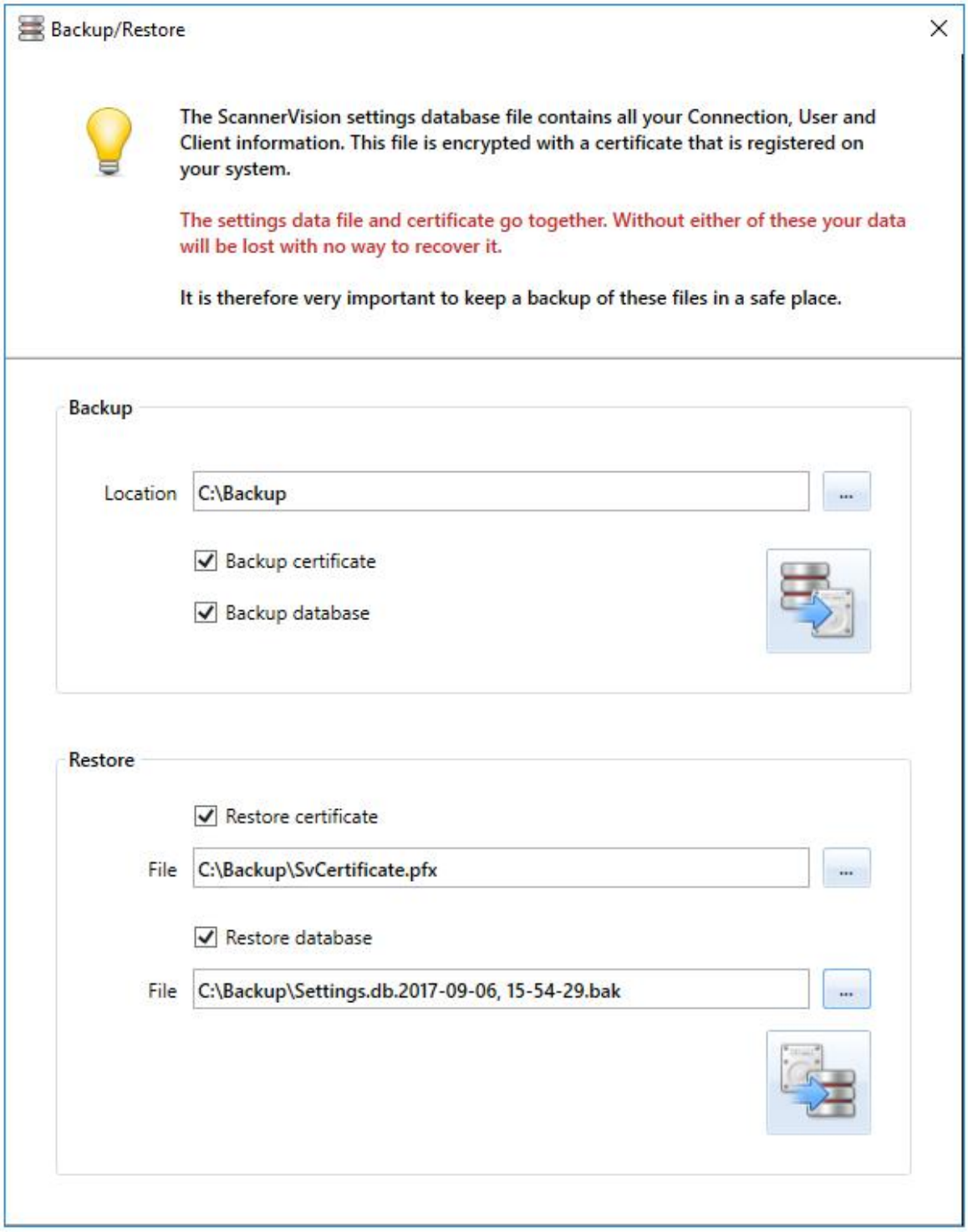


The "Refresh" button refreshes the list of templates.

9.5 Backup/Restore

The ScannerVision settings database file contains all your Connection, User and Client information. This file is encrypted with a certificate that is registered on your system. The settings data file and certificate go together. Without either of these your data will be lost with no way to recover it. It is therefore very important to keep a backup of these files in a safe place.

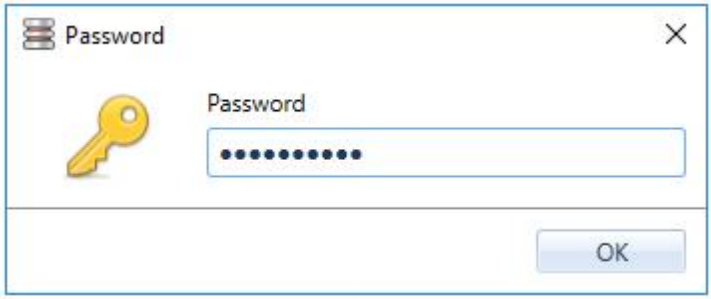
The Backup/Restore utility allows you to create backups of the settings database as well as the certificate.



Backup

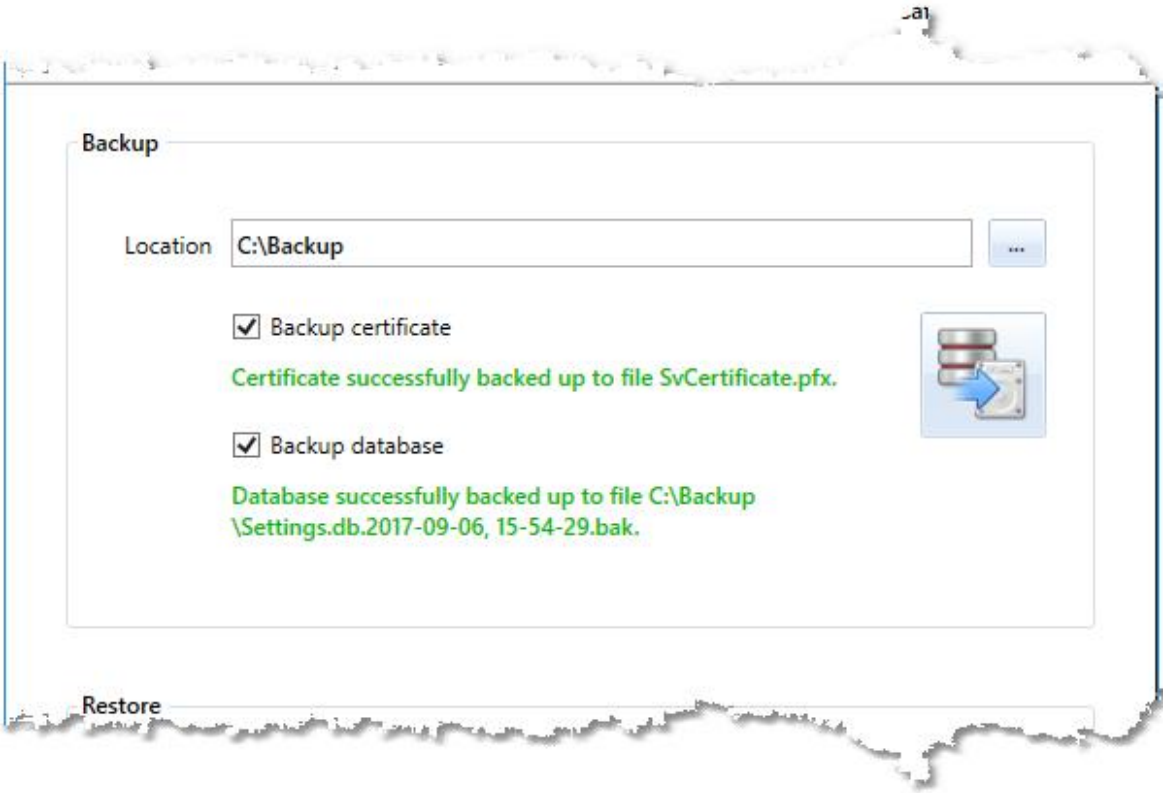
To backup your database and/or certificate:

1. Select the folder in which you want the backup files to be created in.
2. Select which files you want to backup.
3. Press the backup button to the right of the check boxes.
4. If you have selected the "Backup certificate" option you will have to provide a password for the certificate export file:



5. Enter a password and click the "OK" button.

Important You will need to provide this password when you import the certificate so do not lose it.

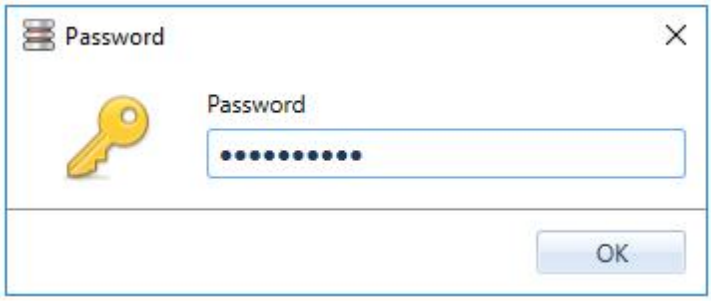


After the backup is complete you will see a status message for each of the two files.

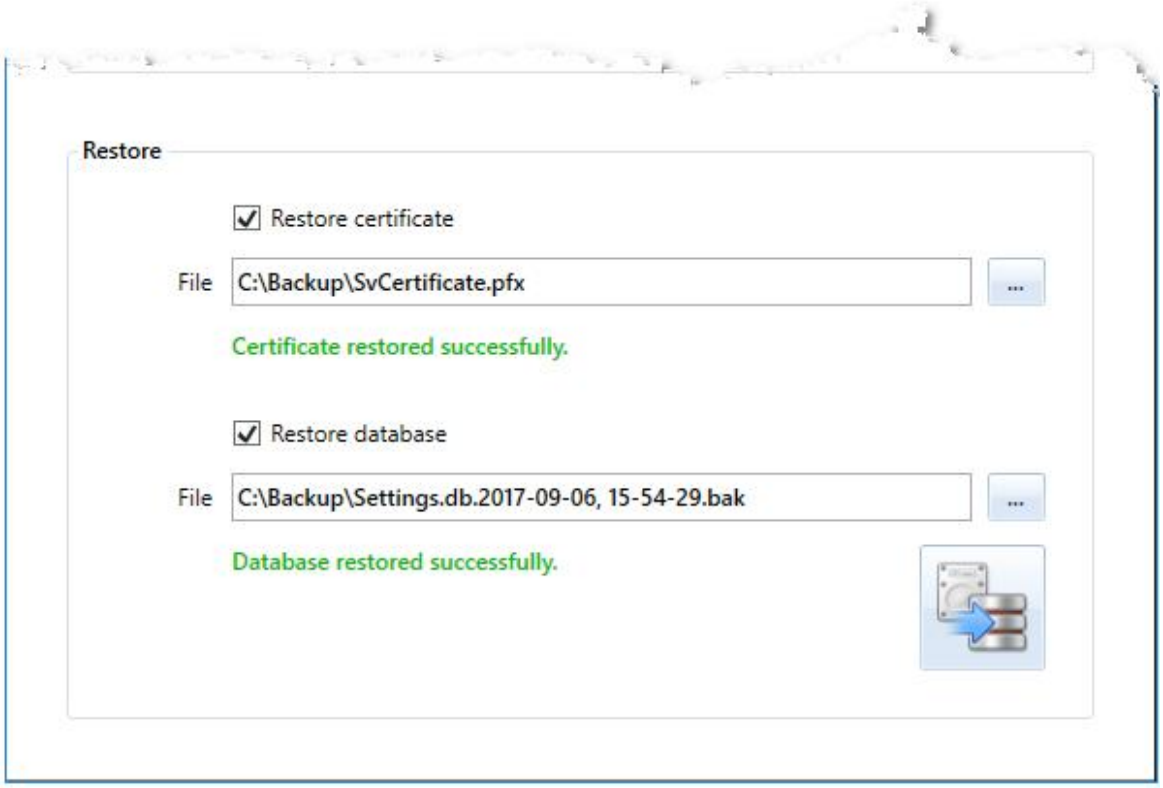
Restore

To restore a certificate and/or database:

1. Select the certificate and/or database backup files you want to restore.
2. Select which files you want to restore.
3. Press the restore button to the right of the check boxes.
4. If you have selected the "Restore certificate" option you will have to provide the password of the certificate export file:



5. Enter a password and click the "OK" button.



After the restoration is complete you will see a status message for each of the two files.

10 Appendices

10.1 Appendix A - Metadata

The term "metadata" refers to data that describes other data. If you take a photograph with a digital camera or cell phone the photo may be captured with the date, time and GPS coordinates of when and where the photo was taken (depending on the specific device capabilities). The photo itself is of course the actual data you are interested in primarily, while the date, time and GPS coordinates are metadata that gives more information about the photo.

ScannerVision metadata contains information about the documents that are processed by ScannerVision and can be:

1. Generated by the ScannerVision itself or the connectors used in the template. This includes data such as the user who captured the document, the date and time of processing and in the case of the Windows File System connector, the output file name of the document.

2. Obtained from the document itself through OCRing or barcode reading.

3. Captured on the client by the user who has to provide this information at the time of the document capture.

ScannerVision metadata is a collection of name value pairs where the name is the identifier of a specific value e.g. YYYY=2013. "YYYY" is the name and "2013" is the value which in this case represents the year when the document was processed. The name of the metadata is also known as the Metadata Tag or just Tag for short. The terms metadata and tag are used interchangeably in the text. Where the use of metadata is allowed in the ScannerVision processing process, you would reference the tag name in square brackets e.g. [YYYY] and the processing engine will replace the name with the actual value when a document is processed.

Here is an example.

Let us say you want to produce a PDF document with a file name that includes the date and time of when the document was processed and you want to copy the file to a network share using the Windows File System connector (WFS). In the WFS connector's "File Name" edit box you would enter the following:

Document [DD] [MMMM], [YYYY] - [HH]_[NN]_[SS].pdf

This will expand to (assuming the date and time of processing was 15 March, 2013 at 10:23:11 AM):

Document 15 March, 2013 - 10_23_11.pdf

You can find more examples here.

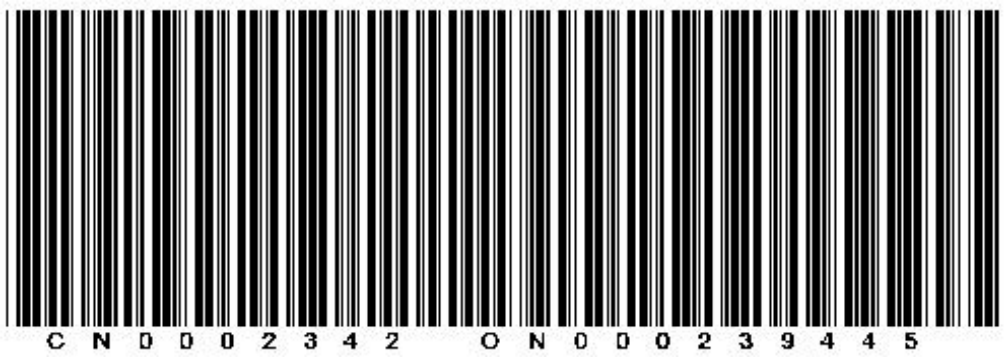
Some metadata is always available like the date and time related tags used in the example above. You don't have to configure anything for them to be available, you just use them. These are called

the Standard tags. All other tags have to be defined by you, the author of the template. You can define metadata on a global level (not worldwide of course, only global across all templates) and/or on a template level. Metadata values can be obtained during document capturing by ScannerVision clients (desktop or MFP) through the answers users provide to template questions or they can be obtained by the processing engine during document processing through for example barcode reading or zone OCRing. If required, the values of metadata you have configured (i.e. all non-standard tags) could be modified during document processing with the use of scripts.

Some connectors offer additional tags that are not necessarily available outside the context of the particular connector. An example of this are the UNIQUESUFFIX tags offered by the Window File System connector. The specifics of these tags are discussed in the File System Connector section.

The value of a tag does not necessarily have to correlate with how it is presented. You could for example have a value of "A-B" and have it presented as "A:B". ScannerVision provides powerful functions with which you can present the data in the format you want without changing the value itself.

For example, let us say you have to read a barcode on the first page of an order form that includes the customer and order numbers separated by a space. So you configure a barcode rule that stores the barcode value in the CUSTOMERORDER tag. Now, let us say you process and order form with the barcode below printed on it:



When ScannerVision processes the document it will read the barcode and put the value "CN0002342 ON000239445" in CUSTOMERORDER tag as instructed. But how will you separate the customer number from the order number? You have two choices. You can either split the two sections on the space character between them or you can pick the characters you need. Either way, you still reference the CUSTOMERORDER tag whenever you need either the customer number or order number. You just need to add some functions that will instruct ScannerVision to take only the part of the data that you need.

You could for example use the **split** function to tell ScannerVision to split the two parts of the barcode on the **" "** (space) character between them and then use the **take** function to select first part for the customer number or the second part for the order number.

To extract the customer number you would use the following expression:

```
[CUSTOMERORDER (split " ")(take 1)]
```

and to extract the order number you would use this expression:

[CUSTOMERORDER (split " ")(take 2)]

Another approach to obtain the customer and order number could be to tell ScannerVision to **take** characters 1-9 for the customer number and 11-21 for the order number.

Where you need the customer number only you would use this:

[CUSTOMERORDER (take 1-9)]

and where you need the order number only you would use this:

[CUSTOMERORDER (take 11-21)]

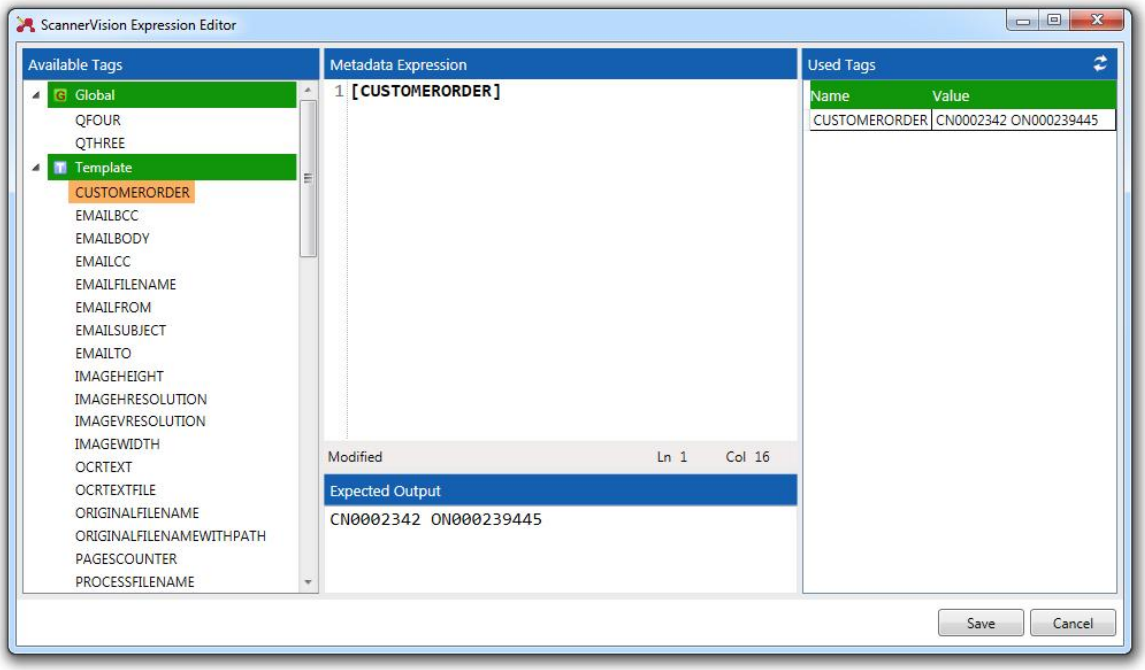
Both techniques would yield the required result. You may ask, "But which one should I use?". The answer is - as is often the case - it depends. If the customer and order numbers are fixed in length i.e. the customer number will always contain 9 characters and the order number 10, then the second technique without the split is the better choice since there is less typing for you and less work for ScannerVision to do. If the numbers are varying in length however you obviously cannot count characters and therefore would have to make use of some sort of delimiter that marks the end of the customer number and the start of the order number. In our example the delimiter is the space character. In this scenario the first method is obviously the one to use.

The ScannerVision metadata functions are explained in detail in the Functions section. Once you understand how they work you can do very powerful things with your metadata!

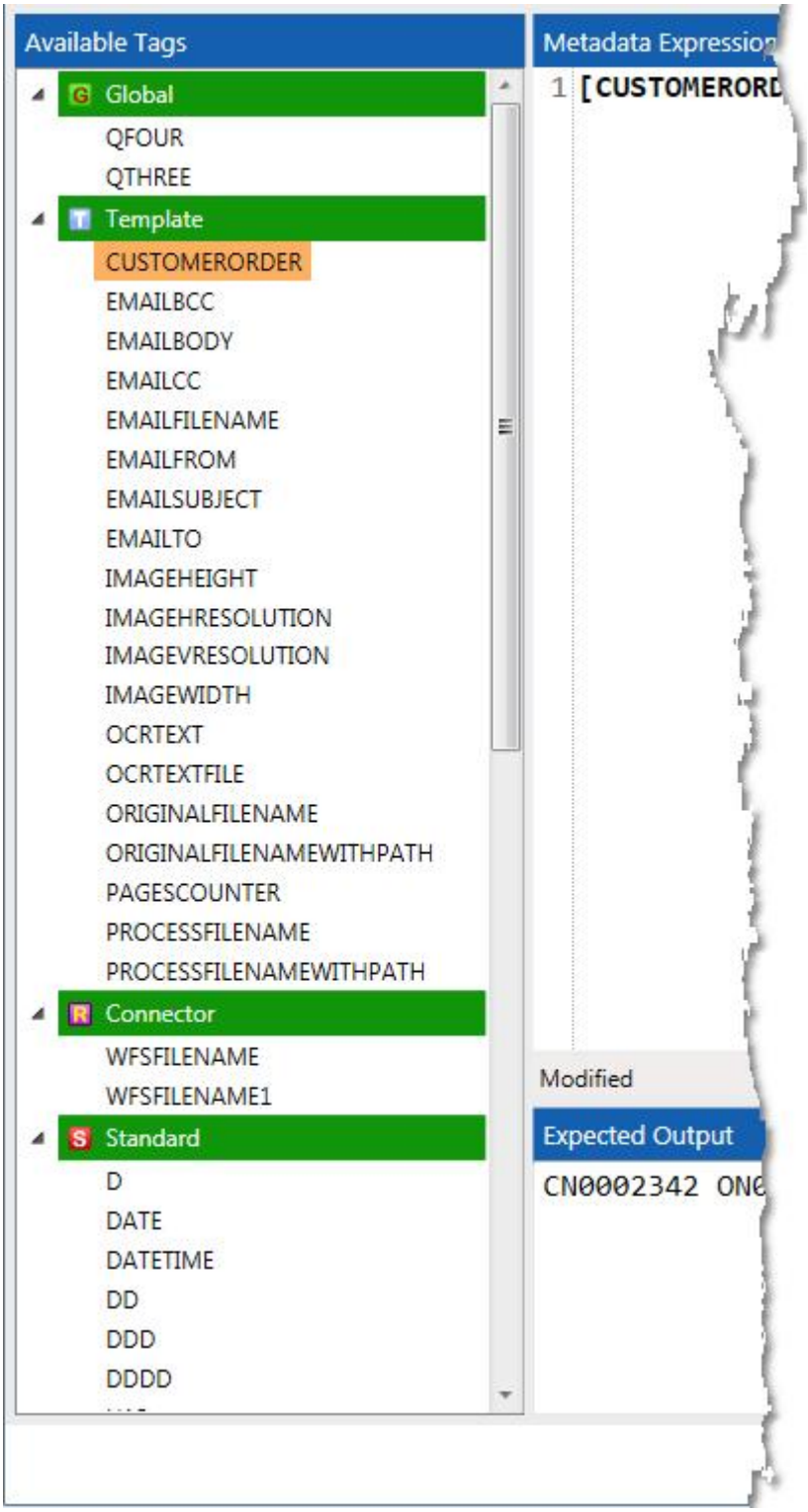
ScannerVision provides you with a powerful ScannerVision Expression Editor with which to create and test your expressions. Using the ScannerVision Expression Editor you can see the result of your expressions using sample data that you can modify in the editor itself.

10.1.1 ScannerVision Expression Editor

The ScannerVision Expression Editor shown below is a rich environment in which you can build the metadata expressions that you want to use in your template. It offers several aids that are designed to help you to get the results you expect. One of these is the ability to see the output of your expression using sample data that you provide, in real time. You therefore don't have to wonder what the output of the expression would be or to process a test document through the system to see the output of the expression. Provided that your sample data represents real-world data, what you see in the expression editor is what you can expect to see during run time.



Available Tags



The available tags section contains all the metadata tags that are available for you to use in your expression. The tags are presented in a tree view with the tags grouped into between one and four groups. The groups represent the different types of metadata that is available in ScannerVision namely "Global", "Template", "Connector" and "Standard". At a minimum you will get the Standard

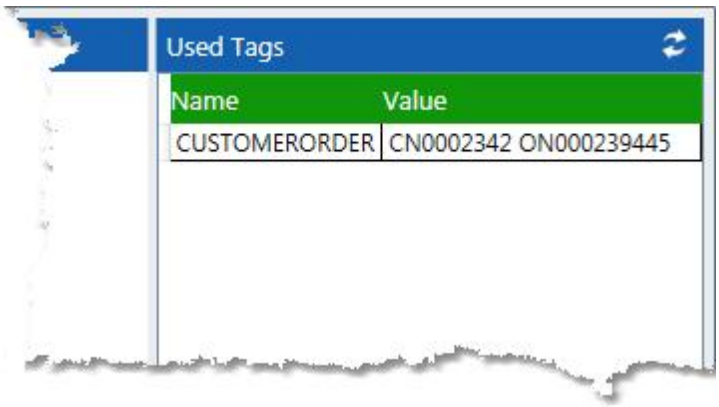
group since these tags are always available. The Global and Template groups are visible when global and/or template metadata have been configured respectively. The connector group will show up when the expression editor is opened from a connector setup screen, but only if the connector supports additional tags such as the Windows File System connector with its UNIQUESUFFIX tags.

The groups in the tree view can be collapsed or expanded by clicking the triangle to the left of the group heading. To collapse or expand all groups simultaneously you can right click with your mouse over the tree view and select the desired option from the context menu.

To insert a tag in your expression place the caret in the desired location of the "Metadata Expression" editor window and double click the tag you want to insert.

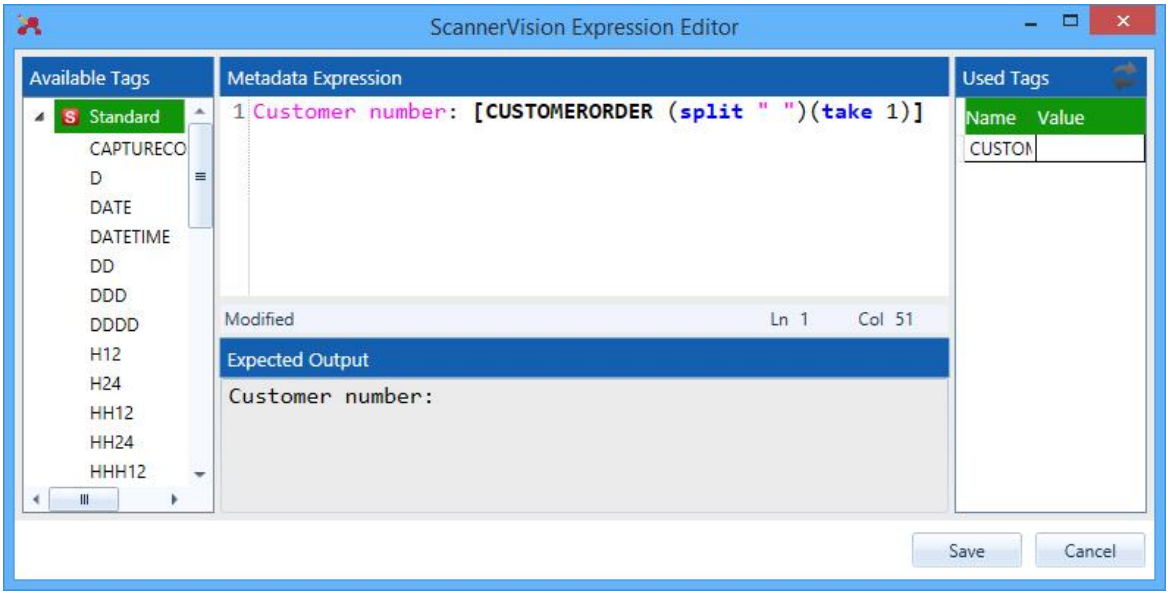
You can modify tags' sample data in the Used Tags grid. To reset the sample data to the original values, click the button in the top right hand corner.

Used Tags



When a valid expression is entered into the expression editor window the Used Tags grid is updated with all the tags that are used in the expression. The grid shows the names of the tags as well as the sample value of each tag. You can edit the sample value in the grid by clicking in the relevant cell. To apply your changes click outside the grid or press Enter. The expression will be parsed using the sample data you entered and the output will appear in the Expected Output window.

Metadata Expression



The metadata expression window is where you enter your expression. The editor offers syntax highlighting which colors the various sections of your expression to make it easier to read. Functions names are printed in bold, blue font; static text is printed in magenta and tag names are printed in a bold, black letters.

There are various ways to enter expressions and you can use them in any combination you wish:

1. You can type the expression using the keyboard.

2. You can double click a tag in the Available Tags section to insert it at the current caret position.

3. You can use "Code Completion" by pressing Ctrl-Space while the caret is between the square brackets of a metadata tag or the round brackets of a function. A context menu will appear with either tag names or function names - depending on the context of the caret. You can then navigate through the menus using the cursor keys on your keyboard or using the mouse. To insert the selected item in the menu, press the Enter key or left click on it with your mouse. Existing values are replaced with the newly selected ones.

4. You can right click in the expression editor area with your mouse to show the context menu from where you can select the desired action. Depending on the context of the caret in the expression, certain menu options may be disabled.

To use options 3 and 4 above, the caret in the editor has to be in between either square brackets for metadata tags or round brackets for functions. If this is not the case, pressing Ctrl-Space will have no effect and the menu items in the right click context menu will be disabled. Tags or functions names that appear outside the scope of their respective style of bracket is treated as plain text.

In the screen shot above there are three sections in the expression with peculiar highlighting:

1. Plain text. The plain text "Order number: " is printed in a magenta colored font. Text parameters of functions such as "match", "join" and "split" are also printed in a magenta colored font.

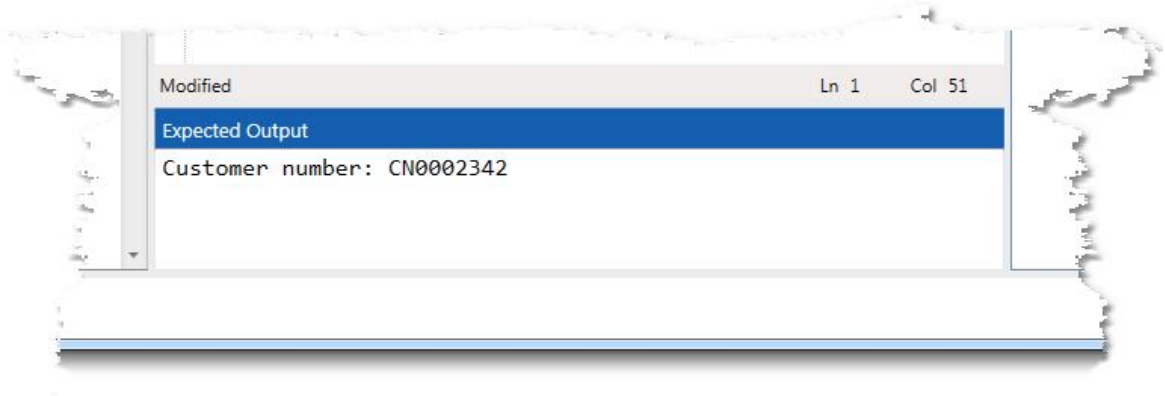
2. Metadata tag. The CUSTOMERORDER tag appears between opening and closing square brackets "[" and "]" and is printed in bold, black font.

3. Functions. The "split" and "take" functions appear between opening and closing round brackets "("

and ")" and are printed in bold, blue font.

At the bottom of the expression editor is a status bar with three sections which indicate the state of the editor window. The first section would contain the word "Modified" when the expression has been modified, "Error" when there is a syntax or parsing error in the expression or blank when there is no error in the expression and it has not been modified. To the right of the status bar are two sections that indicate the position of the caret in the editor. The "Ln" value indicates the current line number and the "Col" value indicates offset into the current line.

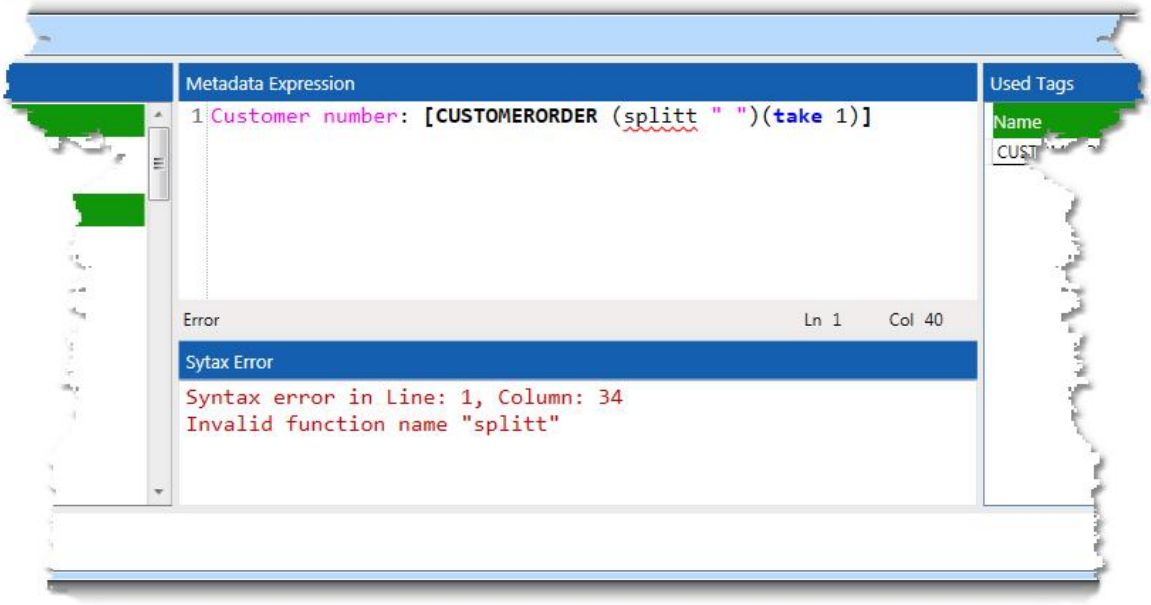
Expected Output



Whenever a change is made to the expression in the expression editor it is parsed. The parsing of the expression does several things:

1. It validates the syntax of the expression. If there are any errors in the expression the parsing will fail. When this happens the caption of the output window changes from "Expected Output" to "Syntax Error" and the output window will indicate what the error is.
2. If the expression does not contain any errors, the output of the expression is shown in the output window using the given sample data.
3. If the expression does not contain any errors, the Used Tags grid is updated with the names (if any) of the tags used in the expression. You can then modify the sample data if you desire.

In the screen shot below you can see what the expression editor would look like when there is an error in the expression.



There are several aids to indicate that there is a problem with the expression and to help you to figure out what the problem is. They are:

1. The word "Error" appears in the status bar below the expression editor window.
2. A red squiggly line appears in the location where the problem is. In the screen shot above the line appears under the function name "splitt".
3. The caption of the output window below the expression editor changes from "Expected Output" to "Syntax Error".
4. An error message appears in the output window. In the example above the message indicates that the "splitt" function name is invalid.
5. The Save button disables.

After you have entered an expression and you have satisfied with the result you can click the save button. This will copy the expression into the edit box from where the ScannerVision Expression Editor was launched.

10.1.2 Tags

In this section we present all the Standard metadata tags that are available in every template.

Standard Tags

Tag Name	Description	Example value
Date and Time		
DOCUMENTPROCESSORID	The ID of the Document Processor that is processing the document	3
DATETIME	Date and time, formatted as YYYY-MM-YY HH24:NN:SS	2013-04-02 08:03:07
DATE	Culture specific short date format. This date format conforms to the regional settings of the operating system on which ScannerVision is installed.	4/2/2013
TIME	Culture specific short time format. This date format conforms to the regional settings of the operating system on which ScannerVision is installed.	8:03 AM
UTC	Coordinated Universal Time, formatted as 'yyyy'-'MM'-'dd'T'HH':'mm':'ss'.'fffffffZ'	2013-04-02T08:03:07.1234567Z
UCT	Coordinated Universal Time, formatted as 'yyyy'-'MM'-'dd'T'HH':'mm':'ss'Z'	2013-04-02T08:03:07Z
TZD	Time zone offset, formatted as +/-HH12:mm or "Z" if offset is zero.	+02:00 or Z
YYYY	Four digit year	2013
YY	Two digit year without century	13
MMMM	The full name of the month	April
MMM	The abbreviated name of the month	Apr
MM	The month with leading zeros from 01 through 12	04
M	The month without leading zeros from 1 through 12	4
DDDD	The full name of the day of the week	Tuesday
DDD	The abbreviated name of the day of the week	Tue

DD	The day of the month with leading zeros from 01 through 31	02
D	The day of the month without leading zeros from 1 through 31	2
WW	Week of the year with leading zeros from 01 through 52	15
W	Week of the year without leading zeros from 1 through 52	15
HH24	The hour with leading zeros, using a 24-hour clock from 00 to 23	08
H24	The hour without leading zeros, using a 24-hour clock from 0 to 23	8
HHH12	AM/PM designator	AM
HH12	The hour with leading zeros, using a 12-hour clock from 00 to 12	08
H12	The hour without leading zeros, using a 12-hour clock from 0 to 12	8
NN	The minute with leading zeros, from 00 through 59	03
N	The minute without leading zeros, from 0 through 59	3
SS	The second with leading zeros, from 00 through 59	07
S	The second without leading zeros, from 0 through 59	7
ZZZ	Thousandths of a second	000
ZZ	Hundredths of a second	00
Z	Thenths of a second	0
Client		
CLIENTIPADDRESS	The IP address of the client that is connected to the server or from where the document was submitted	192.168.1.103
CLIENTSN	The serial number of the machine from where the document was submitted	SN:0029332022
MFDMODELNR	The model number of the machine from where the document was submitted	Aficio 4502

User		
USERNAME	The user name of the ScannerVision user who submitted the document	billy
USERMAIL	The email address of the ScannerVision user who submitted the document	billyb@domain.com
USERHOMEFOLDER	The home folder of the ScannerVision user who submitted the document	\some\folder
SMTP Server		
SMTPSERVER	SMTP server address	192.168.1.1
SMTPPORT	SMTP server port	25
SMTPAUTHENTICATION	SMTP server authentication method	No Authentication
SMTPUSER	SMTP server user name	admin
SMTPFROM	FROM address of emails sent by ScannerVision	noreply@domain.com
SMTPTO	Address to which emails are sent to by ScannerVision	administrator@domain.com
ScannerVision		
MAINCAPTURECOUNTER or CAPTURECOUNTER	The total number of documents that have been captured by ScannerVision, excluding split documents. If a document is split into 8 separate documents this counter is incremented by 1.	21325
MAINCOUNTER	The total number of documents that have been processed by ScannerVision, including split documents. If a document is split into 8 separate documents this counter is incremented by 8.	32665
MACHINEID	The product ID of the Windows operating system	55041-049-8010363-86340
SVSERVERVERSION	The ScannerVision version number	4.5.0.0
Document		
ORIGINALFILENAME	The name of the incoming file excluding the file path but including the file extension	Scan_000123.pdf
ORIGINALFILENAMEWITHOUTEXTENSION	The name of the incoming file excluding the file path and extension	Scan_000123
ORIGINALFILENAMEWITHP	The fully qualified name of the incoming file	H:\Hot Folder

ATH		\Scan_000123.pdf
OUTPUTEXTENSION	The extension of the output document	.pdf
PROCESSFILENAME	The name of the file after image processing has been performed, excluding the file path but including the file extension	Scan_000123.tif
PROCESSFILENAMEWITHPATH	The fully qualified name of the file after image processing has been performed	H:\Hot Folder \Scan_000123.tif
PAGESCOUNTER	The number of pages in the original (un-split) document being processed	43
CURRENTPAGECOUNT	The number of pages in the current (split) document being processed. If document splitting is not enabled the value of this tag will be the same as that of the PAGESCOUNTER tag.	15
IMAGEWIDTH	The width of first page in the document being processed	1240
IMAGEHEIGHT	The height of first page in the document being processed	1754
IMAGEHRESOLUTION	The horizontal resolution of the first page in the document being processed	150
IMAGEVRESOLUTION	The vertical resolution of the first page in the document being processed	150
MASTERFORMNAME	The name of the Automatic Forms Recognition master form that recognized the document	Invoice
Template		
TEMPLATECAPTURECOUNTER	The total number of documents that have been captured by the template, excluding split documents. If a document is split into 8 separate documents this counter is incremented by 1.	4587
TEMPLATECOUNTER	The number of documents that have been processed by a template, including split documents. If a document is split into 8 separate documents this counter is incremented by 8.	5412
TEMPLATEERROR	The error message that caused template execution to fail excluding capturing errors.	Unsupported document format.
TEMPLATEID	The GUID of the template	{B35C3F85-0810-40FF-

		AB7C-25DEA97B9E1D}
EMAILID	*An identification number that is reset every time a capture source is processed.	{273740AB-3102-40CA-A441-B6E54A75FF90}
TEMPLATENAME	The name of the template	Invoices
PROCESSINGRESULT	Indicates if a template has executed successfully. This can be used in an end script to perform some action based on this result.	0 = Failure 1 = Success
REJECTEDFOLDERPATH	The path (excluding the trailing backslash) where rejected documents are stored.	c:\templates\rejected
REJECTEDFILENAME	The name of the rejected document.	scan0001 {4EBB7C21-CFF0-4FD4-8426-42878C70D397}.tif
REJECTEDFILENAMEWITHPATH	The fully qualified path of the rejected document.	c:\templates\rejected\scan0001 {4EBB7C21-CFF0-4FD4-8426-42878C70D397}.tif

In the case where a Workflow Template is processing a document that has been split by an Automatic Forms Recognition Template the [ORIGINALFILENAME], [ORIGINALFILENAMEWITHOUTEXTENSION] and [ORIGINALFILENAMEWITHPATH] tags will contain the document that was captured by the Automatic Forms Recognition Template and not that of the split document that is being processed by the Workflow Template.

*EMAILID

The [EMAILID] can be used in the case of capturing from an email address to group documents attached to the email together. If for example you need to store all the documents that were attached to an email in the same, unique folder, you can use this tag.

If for whatever reason the template does not complete the processing of all the documents included in the email and it has to processed again, i.e. next time when the template it scheduled, the EMAILID will be different than the one before.

Email Capture Tags

The following table lists the tags that are available when email capturing has been configured for the template. These tags are only available during document processing.

Tag Name	Description
Email Capture	
EMAILBCC	Complete BCC address as obtained from the email header

EMAILBODY	Email body text
EMAILCC	Complete CC address as obtained from the email header
EMAILFILENAME	Name of attachment
EMAILFROM	Complete FROM address as obtained from the email header
EMAILTO	Complete TO address as obtained from the email header
EMAILSUBJECT	Email subject

The format of email addresses in the TO, FROM, CC and BCC fields is dependent on the email client/server from which the email was sent and could be in any of the following forms (and possibly more)

"Name Surname" <person@domain.com>

or

"'Name Surname'"

or

<person@domain.com>

You can use the ScannerVision metadata functions to extract only the data you want.

OCR Tags

The following table lists the tags that are available when OCRing is enabled in the template. These tags are only available during document processing.

Tag Name	Description
OCR	
OCRTEXT	OCRed text, limited to 64MB*
OCRTEXTFILE	Fully qualified path to a file containing all OCRed text
OCRTEXTXML	XML description of the OCRed words including the position and dimensions of each word's bounding box
OCRTEXTXMLFILE	Fully qualified path to a file containing the XML description of the OCRed words

Since documents can be very large the value of the OCRTEXT and OCRTEXTXML tags are limited to 64MB. If you need to get access to the full body of OCRed text, you can reference the OCRTEXTFILE

or OCRTEXTXMLFILE tags which hold the fully qualified path to text files containing all the OCRed text.

* There is one exception to the 64MB limit. When you configure a data export rule in the Data Export Settings tab and you reference ONLY the OCRTEXT or OCRTEXTXML tags (no metadata functions, custom text etc.) in the "Custom output" field e.g. [OCRTEXT] or [OCRTEXTXML], all OCRed text will be exported to the specified file. When specifying [OCRTEXTXML] in the export rule be sure to select either the Unicode or UTF8 encoding. Failing to do so could result in an output file that may not be parseable by all Xml parsers.

The xml structure produced by the OCRTEXTXML tag is shown below:

```
<Document>
  <Page>
    <word x="37" y="58" w="170" h="29">Pottery</word>
  </Page>
  <Page>
  </Page>
  <Page>
    <word x="665" y="227" w="76" h="21">Product</word>
    <word x="1077" y="1428" w="10" h="17">3</word>
    <word x="1190" y="1427" w="58" h="20">$4.50</word>
    <word x="1306" y="1426" w="56" h="20">$6.00</word>
    <word x="814" y="2231" w="8" h="16">10</word>
  </Page>
</Document>
```

Zone OCR, Zone OMR, Zone ICR, Zone MICR

The following table lists the tags that are available when Zone OCR or OMR has been configured in the template. These tags are only available during document processing.

Tag Name	Description
Zone OCR	
ZONEOCRN	OCRed text of the Nth zone of the last page in the document
ZONEOCRN-N	OCRed text of the Nth zone of the Nth page in the document
Zone OMR	
ZONEOMRN	OMR result of the Nth zone of the last page in the document
ZONEOMRN-N	OMR result of the Nth zone of the Nth page in the document
Zone ICR	
ZONEICRN	ICR result of the Nth zone of the last page in the document
ZONEICRN-N	ICR result of the Nth zone of the Nth page in the document
Zone MICR	

ZONEMICRN	MICR result of the Nth zone of the last page in the document
ZONEMICRN-N	MICR result of the Nth zone of the Nth page in the document

The tags for Zone OCR, Zone OMR, Zone ICR and Zone MICR all work the same and will be explained in the context of Zone OCR only but the conventions apply to all four kinds of tags.

When you want to reference a zone, you would replace the "N" in the tag name above with the particular page and zone number you are interested in. The first "N" represents the zone and the second "-N" represents the page. So, if you want to refer to the 2nd zone of the 3rd page you would use the tag ZONEOCR1-2. Numbering of pages and zones are zero based which is to say page 1 = 0 , page 2 = 1, zone 3 = 2 and so on. Also, page numbers don't necessarily correlate with document pages. If for example you have defined a range of pages on which recognition has to be performed, say 1, 3 and 7 then the tag names of the OCR zones for page one of the document will be ZONEOCRN-0, for page 3 they will be ZONEOCRN-1 and for page 7 ZONEOCRN-2. When you define page ranges you can think of it as if you are OCRing a new document with only the pages you have selected.

The ZONEOCRN tags are special tags in that they refer to the zones of the last page in the document. This is useful when you don't know how many pages there would be in the documents you are processing, but you do know that you want to OCR the last page. With this notation you would just replace the "N" with the particular zone number.

The zone OCR, ICR and MICR tags contain the OCRed text of the particular zone. The zone OMR tags represent the presence or absence of a mark in the defined zone. To indicate that a mark was found the tag will contain the value "1" and if no mark was found it will contain the value "0".

It is recommended that you name your tags explicitly as described in the Zone OCR section instead of using the tags described above.

Barcode 1D & 2D

When automatic barcode recognition is enabled metadata tag names are generated automatically. The convention used for naming 1D barcode tags is "BC1DVALUE" with a number representing the instance of the barcode e.g. "BC1DVALUE0", "BC1DVALUE1" etc. For 2D barcodes the name "BC2DVALUE" is used.

Tag Name	Description
Barcode	
DOCUMENTSPLITTER	The value of the barcode that triggered a document split.

Connectors

When you have configured a connector that executes when a previous one has failed, i.e. by selecting the "Execute on error only" option, the "CONNECTORERROR" tag will be available for use by the connector and it will contain the error message of the failing connector.

Tag Name	Description
Barcode	
CONNECTORERROR	The error message of the failing connector.

Special Tags

You can make use of special tags called "Hash tags" which you won't find in the tag lists that the ScannerVision Expression Editor provides. They are not "Standard", "Global" or "Template" per se, although they are always available. Hash tags are so called because they are comprised of the "#" or "#0x" prefix followed by a number e.g. #13 and #0x00B0. The number is called the "ordinal value". You would use these when you want to include non-printable characters like TAB, CR, LF etc. All characters - whether they are printable or not - have an ordinal value. Here are a few examples:

TAB = #9

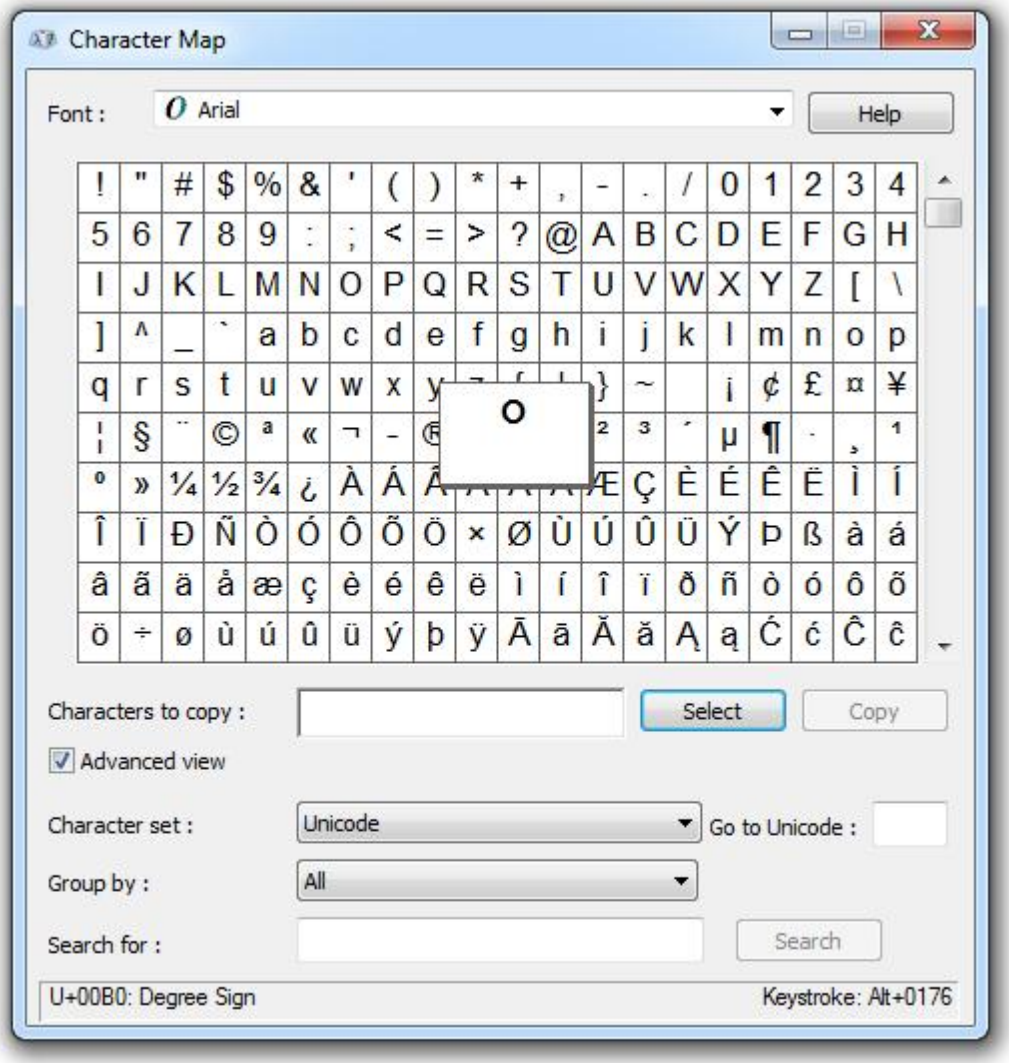
CR = #13

LF = #10

° = #0176 (degree symbol)

TIP

You can find all characters' ordinal values using the built-in character map in Windows. If you have Windows 7 installed, just type "Character Map" in the "Search programs and files" edit box in the Start Menu.



In the screen shot above the degree symbol is highlighted. In the status bar at the bottom of the screen you can see its UNICODE value in both HEX and decimal format. The HEX format "U+00B0" is on the left and the decimal format "0176" on the right. You can use either the decimal or HEX value in expressions. Not all characters have a decimal value though such the Greek alphabet letters e.g. Alpha ("α"). Decimal values have to be prefixed with a "#" symbol and HEX values with "#0x" when they are used in expressions.

The letters in the HEX ordinal values are not case sensitive.

Here are a few examples:

Expression	Output
15[#0x00B0]	15°
15[#0x00b0]	15°

15[#176]	15°
[#0x03b1][#0x03b2][#0x03b3] or [#0x03b1#0x03b2#0x03b3]	αβγ
[#91] or ["["]	[
[#93] or ["]"]	]

10.1.3 Functions

This section of the manual serves as a reference of the available metadata functions that are available in the ScannerVision Expression Editor. The examples given in this section are intended to show the behavior of the relevant function only. For more complete and real world example refer to the Examples section.

The ScannerVision Expression Parser is the engine that reads your expressions and execute the functions you specified. Before explaining the functions and how they work we will explain how the Expression Parser works internally as this will help you to unlock the full power that the functions provide.

Terminology

We will start by defining a few programming terms that we will use in the text that follows.

Functions and parameters

Functions are instructions that a computer executes. You are telling the computer to DO something. Among the available metadata functions are the "ucase" and "split" functions. With the "ucase" function you are telling the computer, "Make all characters in the text uppercase". With some functions the computer needs more information such as with the "split" function which needs to know WHAT to split on. We will be telling the "split" function what to split on with what is called a function parameter. Some functions may require more parameters than others but all ScannerVision metadata functions have an implied parameter namely the text on which to work. This is mostly the metadata tag in the context of which the function appears but it could also be the result of previous function. You won't have to specify this parameter explicitly.

Metadata tags in ScannerVision are delineated with square brackets. The [DATETIME] tag represents a date and time value such as "2013-04-02 08:03:07". To apply metadata functions to the [DATETIME] tag you would put the functions inside the tag, before the closing square bracket, surrounded with parentheses and with a space between the tag name and the opening parentheses of the first function.

In the following expression:

```
[DATETIME (split "-")(take 1)]
```

we are applying the "split" and "take" functions to the DATETIME tag. In the case of the "split" function we are passing it a parameter "-" which tells the function that we want the date & time to be split on the hyphen. This results in 3 parts namely "2014", "04" and "02 08:03:07" which become the input of the "take" function. In the "take" function we are passing the parameter 1 which tells the function that we want the first part of the split result i.e. "2014".

If a function requires more than one parameter such as the "replace" function they are separated by commas e.g.:

```
[DATETIME (replace "-", "/")]
```

Input & Result

Input is the data on which a function works and *Result* is the outcome of that operation. In the expression:

```
[DATETIME (split "-")(take 1)]
```

the value of the DATETIME tag - let us assume that to be "2013-03-12 14:23:54" - is the ***input*** of the "split" function and the outcome "2013", "03", "12 14:23:54" is its ***result***. The result of the "split" function becomes the input of the "take" function which yields the result "2013".

Characters and Strings

We will explain characters and strings in the context of a text editor like Notepad. Anything you type into Notepad is just text whether it is letters of the alphabet, numbers or symbols. Every keystroke represents a ***character*** e.g. 'A', 'b', '1', '@'. These are all characters and each has a unique number called and ordinal value which we encountered in the discussion of metadata tags and the UNICODE character map. Not all characters are visible such as the Space or Tab characters but they all have an ordinal value.

A sequence of characters is called a ***string***. A string can contain zero or more characters. When a string contains zero characters it is called and empty string. The following are all strings: "Customer", "INVOICE00012345", "$3000", "25º", "A", "".

To distinguish between the ***character*** A and the ***string*** A in the discussion below we use single quotes to indicate the character 'A' and double quotes to indicate the string "A". So,

```
'A' = Character A

"A" = String A

'AB' is not valid because there is no character AB.
```

Integers

Integers are whole numbers which are numbers without a decimal value e.g. 1, 300, -15.

Arrays

An ***array*** is a series of values and you can visualize it as a table with many rows and only one column. The numbers of the rows are called ***indexes*** and the values are called ***elements.***

Index	Elements
1	CN000123
2	ON023456

The table above represents an array with 2 elements. To refer to the elements we use the notation [1], [2] etc. So, [1] = "CN000123" and [2] = "ON023456". Don't confuse the square brackets "[" and "]" with ScannerVision tags. If we want to refer to the whole array, we use the notation: ["CN000123", "ON023456"]. Here the double quotes indicate the elements in the array are *strings*. You could also have ['a', 'b', 'c'] which would be an array of characters.

Arrays in ScannerVision metadata functions will always contain strings or characters. When we want to refer to an array of strings we use the notation string[] and for a character array we use character[].

Any string can be thought of as an array of characters. So the string "Apple" is equivalent to: ['A', 'p', 'p', 'l', 'e']

This is why you are able to use the "take" function on a string. Let us use the DATETIME tag with the value "2013-03-12 14:23:54" as an example. You could define an expression as follows:

```
[DATETIME (take 1-4)]
```

The result of the expression is the array ['2', '0', '1', '3']. If you pasted the expression above in the ScannerVision Expression Editor you won't see the array ['2', '0', '1', '3'] but "2013". The result of the "take" function is an array and whenever the **last** function in an expression produces an array ScannerVision automatically converts it to a string by concatenating all the elements in the array - even when the elements are strings themselves. We discuss this in more detail below. The concatenation of array elements to produce a string is what the "join" function does so we could have written the expression above as follows:

```
[DATETIME (take 1-4)(join)]
```

If we call a function such as "(split "-")" we are instructing ScannerVision to look for all instances of the "-" string and to split the string there. The result is an array of strings. Using the DATETIME example above again, if we had the following expression:

```
[DATETIME (split "-")]
```

the result is an array of strings as follows:

[1] = "2013"

[2] = "03"

[3] = "12 14:23:54"

So, given the expression:

```
[DATETIME (split "-")(take 1)]
```

the string array result of the "split" function becomes the input of the "take" function. We told the "take" function to take array element 1 which is "2013". Not surprisingly, if we had said (take 2) the result would have been "03".

When a string is passed into a function that expects a string[] as input, the function converts the string into an array of strings e.g. "apple" becomes ["a", "p", "p", "l", "e"]. Similarly, when a string[] is passed to a function that expects a character[], the elements in the array are concatenated and then converted into a character[] e.g. ["brown", "dog"] becomes ['b', 'r', 'o', 'w', 'n', 'd', 'o', 'g'].

To see this for yourself, enter the following expression in the ScannerVision Expression Editor:

```
["ABCD" (join "*")]
```

The result of this is: "A*B*C*D". The join function expects an array so the string "ABCD" is converted to ["A", "B", "C", "D"] before the join "*" is performed.

Types

Type is a collective noun for *character*, *string*, *integer* and *array*. Here are the types we have encountered:

Example	Identifier	Description
'A'	character	A single character
"Apple"	string	Zero or more strings
1	integer	A number without decimals
['A', 'b', '$']	character[]	Array of character
["Apple", "Pear"]	string[]	Array of string

Application of the terms

With the information given above you will understand the following statement:

The ***split*** function takes a ***string*** as ***input***, a ***string*** as a ***parameter*** and ***returns*** a ***string[]***.

Implicit "join"

When a metadata function returns an array of string and it is the last function of the tag, ScannerVision does an implicit "join" of the array elements to form a string.

Example:

Let's say the BC4 tag contains the value "2013-03-12" and you perform a split on the "-" character like this:

[BC4 (split "-")]

The result you'll see in the ScannerVision Expression Editor is "20130312" and not ["2013", "04", "12"]. Behind the scenes ScannerVision actually did this:

[BC4 (split "-")(join)]

Metadata concatenation

Anywhere that a metadata tag can be used in a function, multiple tags can be used e.g. [TAG (match ["^"][BC])]. You could for example build up a regular expression using a combination of strings and metadata tags. In the previous example, let's say BC contains the value "INV001" the regular expression will expand to [TAG (match "^INV001")].

Regular Expressions

The "split" and "match" functions take a string parameter which represents the pattern on which to do the split or match. This pattern could be any valid regular expression. An explanation of regular expressions is beyond the scope of this manual. A good understanding of regular expressions is however highly recommended if you want to make full use of the power of ScannerVision metadata functions. We can recommend the Regular-Expressions.info website if you want to brush up on you Regex skills.

White Space

White space characters include the following:

- SPACE (U+0020)
- OGHAM SPACE MARK (U+1680)
- MONGOLIAN VOWEL SEPARATOR (U+180E)
- EN QUAD (U+2000)
- EM QUAD (U+2001)
- EN SPACE (U+2002)
- EM SPACE (U+2003)
- THREE-PER-EM SPACE (U+2004)

- FOUR-PER-EM SPACE (U+2005)
- SIX-PER-EM SPACE (U+2006)
- FIGURE SPACE (U+2007)
- PUNCTUATION SPACE (U+2008)
- THIN SPACE (U+2009)
- HAIR SPACE (U+200A)
- NARROW NO-BREAK SPACE (U+202F)
- MEDIUM MATHEMATICAL SPACE (U+205F)
- IDEOGRAPHIC SPACE (U+3000)
- LINE SEPARATOR character (U+2028)
- PARAGRAPH SEPARATOR character (U+2029)
- CHARACTER TABULATION (U+0009)
- LINE FEED (U+000A)
- LINE TABULATION (U+000B)
- FORM FEED (U+000C)
- CARRIAGE RETURN (U+000D)
- NEXT LINE (U+0085)
- NO-BREAK SPACE (U+00A0).

10.1.3.1 join

Concatenates all elements in a string array inserting the given delimiter in between.

Property	Value
Input	string[]
Parameter	string
Result	string

Examples

Input	Function	Result
"ABCDE"	(join "-")	"A-B-C-D-E"
["2013", "04", "01"]	(join "-")	"2013-04-01"
["ABC", "GHI"]	(join "DEF")	"ABCDEFGHI"
["2013", "04", "01"]	(join "") or (join)	"20130401"

10.1.3.2 lcase

Converts all characters in all elements of a string array to lowercase.

Property	Value
Input	string[]
Parameter	-
Result	string[]

Examples

Input	Function	Result
"Apples"	(lcase)	"apples"
["Apples", "Pears", "ORANGES"]	(lcase)	["apples", "pears", "oranges"]

10.1.3.3 lpad

Left pads all elements in a string array with the given character up to the required length.

Property	Value
Input	string[]
Parameter	character, integer
Result	string[]

Examples

Input	Function	Result
"Apples"	(lpad '*', 10)	****Apples
["Apples", "Pears", "ORANGES"]	(lpad '*', 10)	["****Apples", "*****Pears", "***ORANGES"]

10.1.3.4 ltrim

Removes all leading, white space characters from all elements in a string array.

Property	Value
Input	string[]
Parameter	-
Result	string[]

Examples

Input	Function	Result
" Apples "	(ltrim)	"Apples "
[" Apples ", " Pears ", " Oranges "]	(ltrim)	["Apples ", "Pears ", "Oranges "]

10.1.3.5 match

Finds all matches of the given regular expression in all elements in a string array.

Property	Value
Input	character[]
Parameter	regex string
Result	string[]

Examples

Input	Function	Result
"ab12cd34ef15"	(match "\d\d")	["12", "34", "15"]
"ab12cd34ef15"	(match "1\d")	["12", "15"]
["ab12cd34ef1", "5ab12cd34ef15"]	(match "\d\d")	["12", "34", "15", "12", "34", "15"]

10.1.3.6 remove

Removes all elements in a character or string array that matches a given regular expression. If no regular expression is provided the function removes all empty elements.

Property	Value
Input	character[] / string[]

Parameter	regex string
Result	character[] / string[]

Examples

Input	Function	Result
"2013-07-31 11:40:01"	(remove "[-,:,]")	["20130731114001"]
["ab==ef=gh"]	(split "=")(remove)	["ab", "ef", "gh"]

10.1.3.7 replace

Replaces all instances of a sub string matching a given regular expression with another given string

Property	Value
Input	character[] / string[]
Parameter	regex string, string
Result	character[] / string[]

Examples

Input	Function	Result
"2013-07-31 11:40:01"	(replace "-", "/")	["2013/07/31 11:40:01"]

10.1.3.8 reverse

Reverses all elements in a character or string array.

Property	Value
Input	character[] / string[]

Parameter	-
Result	character[] / string[]

Examples

Input	Function	Result
"abcdef"	(reverse)	["fedcba"]
"ab12cd34ef15gh17"	(match "1\d")(reverse)(take 1,3)	["17", "12"]
["ab=cd=ef=gh"]	(split "=")(reverse)(join "-")	["gh-ef-cd-ab"]

10.1.3.9 rpad

Right pads all elements in a string array with the given character up to the required length.

Property	Value
Input	string[]
Parameter	character, integer
Result	string[]

Examples

Input	Function	Result
"Apples"	(rpad '*', 10)	Apples****
["Apples", "Pears", "ORANGES"]	(rpad '*', 10)	["Apples****", "Pears*****", "ORANGES***"]

10.1.3.10 rtrim

Removes all trailing, non-printable characters from all elements in a string array.

Property	Value
Input	string[]
Parameter	-
Result	string[]

Examples

Input	Function	Result
" Apples "	(rtrim)	" Apples"
[" Apples ", " Pears ", " Oranges "]	(rtrim)	[" Apples", " Pears", " Oranges"]

10.1.3.11 split

Split all elements in a string array on the given regular expression, removing the regular expression.

Property	Value
Input	character[]
Parameter	regex string
Result	string[]

Examples

Input	Function	Result
"ab12cd34ef15"	(split "\d\d")	["ab", "cd", "ef"]
"ab12cd34ef15"	(split "1\d")	["ab", "cd34ef"]
["ABC", "ABC"]	(split "B")	["A", "CA", "C"]

["ABC", "ABC"]	(split "CA")	["AB", "BC"]

10.1.3.12 take

Select a range of elements from string array.

Property	Value
Input	string[]
Parameter	integer, integer, ...
Result	string[]

Notes

You can pass and unlimited number of parameters to the take function.

Individual elements are specified by index e.g. 1, 2, 3

Element ranges are specified as "from" - "to" e.g. 1-5, 4-1. If the range goes from a higher number to a lower number, the elements are inverted.

To select from an index to the end of the array, omit the "to" value in the range e.g. "2-" which select from element 2 to the end of the array.

Indexes and ranges can be mixed and repeated e.g. 1, 3-5, 1, 6, 9-

Examples

Input	Function	Result
"abcde"	(take 1)	"a"
"abcde"	(take 1, 3)	"ac"
"abcde"	(take 1-3)	"abc"
"abcde"	(take 1-3, 5)	"abce"
"abcde"	(take 2-)	"b,c,d,e"
"abcde"	(take 3-1)	"c,b,a"
["AB", "12", "CD", "34"]	(take 1)	["AB"]

["AB", "12", "CD", "34"]	(take 2, 1)	["12", "AB"]
["AB", "12", "CD", "34"]	(take 3-1)	["CD", "12", "AB"]

10.1.3.13 trim

Removes all leading and trailing non-printable characters from all elements in a string array.

Property	Value
Input	string[]
Parameter	-
Result	string[]

Examples

Input	Function	Result
" Apples "	(trim)	"Apples"
[" Apples ", " Pears ", " Oranges "]	(trim)	["Apples", "Pears", "Oranges"]

10.1.3.14 ucase

Converts all characters in all elements of a string array to uppercase.

Property	Value
Input	string[]
Parameter	-
Result	string[]

Examples

Input	Function	Result
"Apples"	(ucase)	"APPLES"
["Apples", "Pears", "oranges"]	(ucase)	["APPLES", "PEARS", "ORANGES"]

10.1.3.15 unique

Removes all duplicates from all elements in a string array.

Property	Value
Input	string[]
Parameter	-
Result	string[]

Examples

Input	Function	Result
"Apples"	(unique)	"Aples"
["Apples", "Apples", "Pears"]	(unique)	["Apples", "Pears"]

10.1.4 Character Escaping

Character escaping in general computing terms is the act of adding a special character(s) to a sequence of characters to give an alternate meaning to a subsequent character(s) or to be a substitute for a specific character(s).

A common example of this is the replacement of the space character ' ' with "%20" in URLs and the substitution of '<' and '>' with "<" and ">" in Xml and Html.

ScannerVision metadata expressions can be used in the context of various domain specific languages (DSL) such as SQL, Xml and VBScript. Since metadata expressions is a DSL in itself there are situations where the syntax of the one DSL conflicts with that of the other.

We'll use database picklists (see Creating Picklists for more information) to explain the conflict in meaning of square brackets in SQL queries and ScannerVision metadata expressions.

Let's say we are tasked with creating a ScannerVision template that must ask the user for his/her department and employee number. We want the user to select a department from a list of departments and then to selected his/her name from a list of employees which has been filtered to

show only those employees belonging to the selected department.

The query below is the Microsoft TSQL query we will use to build the list of departments:

```
select [DepartmentName], [DepartmentId]
from [dbo].[Departments]
```

The metadata tag name we assign to the departments picklist is [DEPTNO]. When the user selects a department, the department id is put in the metadata tag named [DEPTNO]. This information is sent to the ScannerVision server so that it is available in subsequent queries like the one we'll use to create the list of employees which is shown below:

```
select [EmployeeName], [EmployeeId]
from [dbo].[Employees]
where [DepartmentId] = "[DEPTNO]"
```

The Problem

Before ScannerVision can execute the SQL query it must first parse the query to find all the ScannerVision metadata tags and replace them with actual values. Both Microsoft TSQL and metadata expressions make use of square brackets to indicate special entities such as tables and column names in the case of TSQL and metadata expressions in the case of ScannerVision. So given the queries we presented above the ScannerVision Expression Parser would not know which values between square brackets are database entities and which are metadata tags. In the form above the parser will treat them all as metadata tags and you will and up with a TSQL query after parsing that looks like this:

`select , from .` and `select , from .where = ""` respectively

The Solution

To solve this problem we have to **escape** all TSQL related square brackets so that the ScannerVision Expression Parser does not try and replace them with metadata values. To escape a character in a ScannerVision metadata expression you would use the sequence ["x"] where x is the character(s) you are escaping. The queries above in their escaped form would look like this:

```
select ["["]DepartmentName["]"], ["["]DepartmentId["]"]

from ["["]dbo["]"].["["]Departments["]"]
```

or

```
select ["[DepartmentName]"], ["[DepartmentId]"]

from ["[dbo].[Departments]"]
```

or

```
["select [EmployeeName], [EmployeeId]

from [dbo].[Employees]

where [DepartmentId]"] = "[DEPTNO]"
```

and

```
select ["[]EmployeeName[]"], ["[]EmployeeId[]"]
from ["[]dbo[]"].["[]Employees[]"]
where ["[]DepartmentId[]"] = "[DEPTNO]"
```

or

```
select ["[EmployeeName], [EmployeeId]"]
from ["[dbo].[Employees]"]
where ["[DepartmentId]"] = "[DEPTNO]"
```

or

```
["select [EmployeeName], [EmployeeId]
from [dbo].[Employees]
where [DepartmentId]"] = "[DEPTNO]"
```

Now after parsing the queries would look like the ones we had originally except that the [DEPTNO] metadata tag would have been replaced with whatever the value for [DEPTNO] was. As you can see, there are several ways to escape the queries. They are all valid. It is up to you to decide which is more readable.

We suggest that you copy and paste the various queries above into the ScannerVision Expression Editor to see the results for yourself.

In situations where your queries don't make use of the square brackets to delineate database entities, i.e. they are only used to reference ScannerVision metadata tags, you are welcome to use a regular "unescaped" query such as:

```
select EmployeeName, EmployeeId
from Employees
where DepartmentId = [DEPTNO]
```

Be ware though that [DEPTO] refers to a ScannerVision metadata tag.

10.1.5 Examples

In the discussion so far you have seen the ScannerVision metadata functions used in fairly trivial expressions. In this section we present several more complex examples of how you can use these functions.

Throughout the examples we'll use the fictitious BC tag which holds our sample data. If you want to try out these examples for yourself, copy and paste the expression into the ScannerVision Expression Editor. The first time you do this, the Expected Output window will show the un-parsed expression and you will see the "BC" tags appear in the Used Tags grid. Once you have entered some sample data into the grid the expression will parse as usual. Alternatively, you can put the sample data in double quotes in the place of the BC tag e.g.

```
["2013/05/25" (split "/")(join "-")]
```

We present the sample data in the examples in double quotes so that you can see if there are leading or trailing non-printable characters included in the data. When you copy the sample data, make sure to copy everything BETWEEN the quotes and not to include the quotes themselves - unless of course you use the shorthand shown above in which case you must include the quotes as well.

Reformat Date and Time

Requirement	Replace "/" with "-" in date.
Sample data	"2013/05/25"
Expression	[BC (replace "/" "-")]
Result	2013-05-25

Requirement	Convert date and time to UTC format.
Sample data	"05/25/2013 23:07:48"
Expression	[BC (split "/| ")(take 3, 1, 2)(join "-")]T[BC (split " ")(take 2)]Z
Result	2013-05-25T23:07:48Z

Requirement	Ensure that month and day have a leading zero and replace "/" with "-".
Sample data	"2013/3/4"
Expression	[DATETIME (split "/")(lpad '0', 2)(join "-")]
Result	2013-03-04

Character case

Requirement	Ensure that the name of day start with a capital letter and the remaining letters are lowercase.
Sample data	"monday", "MONDAY", "moNDaY"
Expression	[BC (ucase)(take 1)][BC (lcase)(take 2-)]
Result	Monday

Email address

Requirement	Create a string of semi-colon delimited email addresses given the TO field from an email message header.
Sample data	""Some User1", <someuser1@domain.com>, "Some User2", <someuser2@domain.com>"
Expression	[BC (match "[_a-z0-9-]+(\.[_a-z0-9-]+)*@[a-z0-9-]+(\.[a-z0-9-]+)*(\.[a-z]{2,4})")(join "; ")]
Result	someuser1@domain.com; someuser2@domain.com

Note

To create a Regex that will match every possible form of email address is no trivial task. The one presented here will match most common forms but you will have to adapt it for specific situations if this one is not sufficient.

Matching

Requirement	Extract all 2 digit numbers from a string.
Sample Data	"ab12cd34ef56qw1we3"
Expression	[BC (match "\d\d")]
Result	123456

Find nth last instance

Requirement	Find last sub folder name in a folder path with unknown number of sub folders
Sample Data	"C:\Users\Public\Documents\microsoft\"
Expression	(split "\\")(remove)(reverse)(take 1)
Result	microsoft

Note

Since the split function takes a regular expression the backslash needs to be escaped.

Since the sample data contains a trailing backslash the split function will split on the last backslash which results in an empty last element in the string array. The remove function without a parameter removes all empty elements in the string array. If you omitted the remove function you would have needed to use "take 2" instead.

This technique above can be used to get any last nth element in the array by just specifying the relevant instance in the take function call.

Extract directory

Requirement	Find the directory only of a fully qualified path.

Sample Data	"C:\Users\Public\Documents\microsoft\filename.tiff"
Expression	(match "(.+\\)*")
Result	C:\Users\Public\Documents\microsoft\

Note

Since the match function takes a regular expression the backslash needs to be escaped.

For this expression to work reliably the path:

- Has to contain a file name

  OR

- Has to end with a trailing backslash

Otherwise the last sub directory will be stripped off e.g. "C:\Users\Public\Documents\microsoft" will end up as
"C:\Users\Public\Documents\"

10.1.6 Conditional Expressions

ScannerVision allows you to make use of conditional logic in your expressions. Conditional expressions allow you to use different values as metadata depending on the value of other metadata.

As an example, let's use the following scenario: We want to split a document on a QR barcode called BCSPLITTER and the newly split document must be renamed to the value of the barcode. In order to do this you would use the DOCUMENTSPLITTER tag for the file name in the <%WFS%>. However, if the original document contains only one page then this technique will not work. Since there is only one page in the document it cannot be split and therefore the DOCUMENTSPLITTER tag would be empty. Actually, the DOCUMENTSPLITTER tag would contain the text "[DOCUMENTSPLITTER]" which by ScannerVision convention represents a tag that has no value.

It would be very useful if you could actually test if the DOCUMENTSPLITTER tag was empty and if so use a different tag value e.g. BCSPLITTER. This is exactly what a conditional expression allows you to do.

The syntax of conditional expressions is:

```
{a} ? b : c
```

where

a = Condition
b = Value to return if condition is TRUE
c = Value to return if condition is FALSE

Condition

The condition part of a conditional expression has to be enclosed in curly braces. You would typically compare the value of a metadata data tag with a known value or the value of another tag. In our example above we want to test whether or not the DOCUMENTSPLITTER tag is "empty" - which is to say is contains the value "[DOCUMENTSPLITTER]". The condition part of the expressions would then look like this:

{[DOCUMENTSPLITTER] == "[DOCUMENTSPLITTER]"}

The double equals sign above notes comparison and the double quotes around the second [DOCUMENTSPLITTER] denotes a string literal. "[DOCUMENTSPLITTER]" is not the same as [DOCUMENTSPLITTER]. As stated "[DOCUMENTSPLITTER]" is treated as TEXT while [DOCUMENTSPLITTER] denotes the DOCUMENTSPLITTER TAG and more specifically the value of the DOCUMENTSPLITTER tag.

Arguments

The "b" and "c" in the syntax definition above are called arguments. Arguments could be text (string literals) or tags.

Full Expression

The full conditional expression of our example looks like this:

[{[DOCUMENTSPLITTER] == "[DOCUMENTSPLITTER]"} ? [BCSPLITTER] : [DOCUMENTSPLI

In English the above expression reads like this: If the DOCUMENTSPLITTER tag is empty (i.e. contains the text value "[DOCUMENTSPLITTER]") then use the value of the BCSPLITTER tag otherwise use the value of the DOCUMENTSPLITTER tag.

So why would this expression do what we want? We stated in our example that single page documents must also be renamed to the value of the BCSPLITTER tag (i.e. the barcode value), the same as for multi-page documents containing multiple instances of the particular QR code we want to split on. By convention ScannerVision populates the BCSPLITTER tag with the value of the last instance of the QR code found the original document, we can't just use the BCSPLITTER tag as the file name for the split documents since they would then all have the same name. We can however (and should) use the DOCUMENTSPLITTER tag since it would contain the value of the barcode that caused a documents split. In the single page document case no split occurs and therefore the DOCUMENTSPLITTER tag is empty, but the BCSPLITTER tag is not - it contains the value of the one and only QR code in the document.

Most robust solution

There is still one problem with the full expression we created in the previous section. What happens when we have a single page document does NOT contain a QR code? Now both the DOCUMENTSPLITTER and BCSPLITTER tags are empty! To solve this we would use a second conditional expression in place of the first argument "b". This expression would do exactly the same as the first one, except it would test if the BCSPLITTER tag is empty. It is up to you how you would handle this situation. You could for example use the ORIGINALFILENAME if no QR code is found.

Here is the final expression that will handle all situations:

```
[{[DOCUMENTSPLITTER] == "[DOCUMENTSPLITTER]"} ? [{[BCSPLITTER] == "[BCSPLITTER
```

10.2 Appendix B - XPath Expressions

XPath is a querying language for XML data. An XPath expression resembles a UNIX file system path e.g. "/node/node/node".

Consider this the following xml document:

```
<document>
  <name>Invoice 89363</name>
  <company>ABC Ltd</company>
  <id>473276256</id>
  <ref>BOL8817</ref>
  <date>2010/04/26</date>
</document>
```

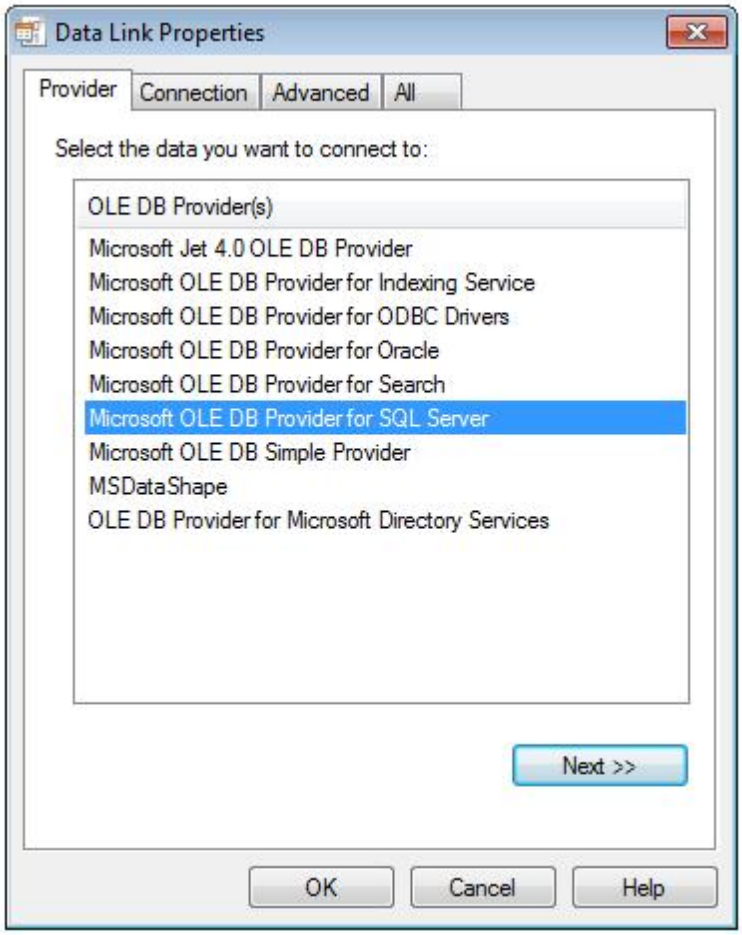
To find the "id" node you would use the following Xpath expression:

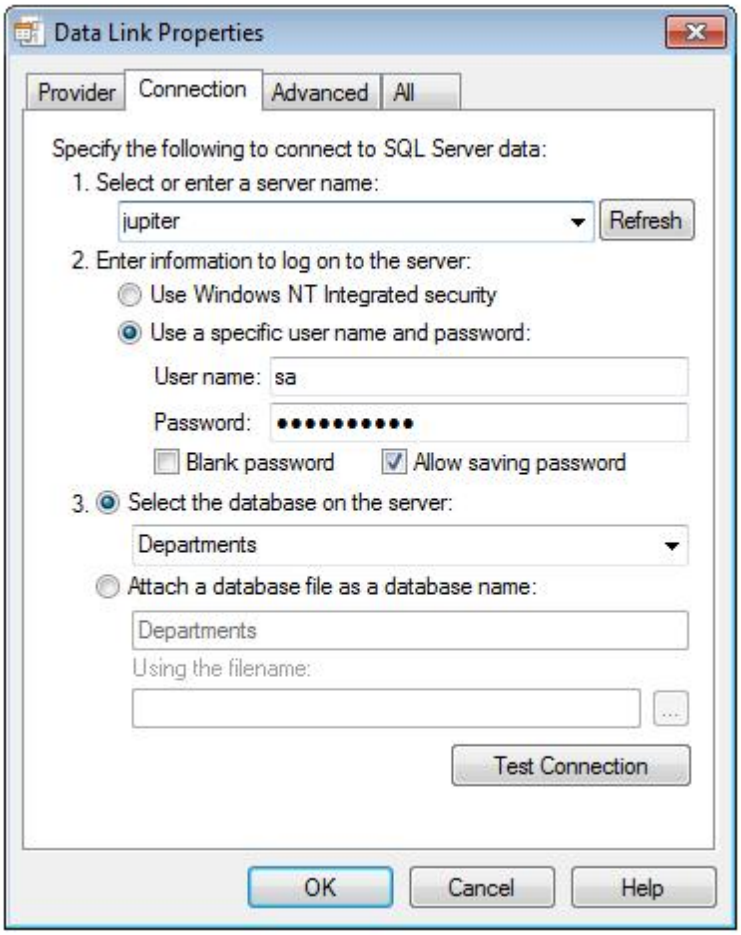
```
/document/id
```

A complete explanation of the XPath language please refer to: XPath Tutorial

10.3 Appendix C - Database Connection Strings

Although a database connection string can be typed manually it is easier to make use of the "Data Link Properties" dialog shown below to configure the connection string.





There are many options available when configuring a database connection string. Please press the "Help" button in the lower right hand corner of the dialog for additional help.

The list of "OLE DB Providers" shown in the screen shot to the left is dependent on the number of OLE DB drivers installed on the system. The list you see may differ.

Word of warning

The "SQL Server Native Client" provider does not store the password in the connection string - even thought the "Allow saving password" option is selected. This is a known bug/limitation of the provider and **not of** ScannerVision.

10.4 Appendix D - Pdf Input Documents

A PDF document can be a vector or a raster format or a combination of the two. Vectors are lines, shapes, symbols and text using specified fonts that represent the content of the document. When this document is loaded by ScannerVision it needs to be converted into a 2 dimensional array of pixels

called a bitmap. This process is called rasterization. The value of the "PDF Load Resolution" setting on the Template General Settings screen determines the resolution of this bitmap. The higher the resolution the better the quality of the bitmap would be but the bigger it would be also. Any operation like OCRing is done on the bitmap and not the original vector document so you have to ensure the bitmap is of sufficient quality to ensure good results. 300dpi is a good number for OCRing.

When the PDF document contains raster images (such as .bmp, .png, .jpeg which are already arrays of pixels and not lines, symbols or fonts) the image does not need to be rasterized like in the case of vector PDFs, but it still needs to be loaded as a 2 dimensional array of pixels (bitmap) like the vector. A different process called resampling could now possibly some into play. When the original raster image is for example 600dpi but the "PDF Load Resolution" option is set to 300pdi it means that the original image would not fit into the 300dpi bitmap. The original image therefore needs to be scaled down or re-sampled to fit into the 300dpi bitmap. The same process happens when the original image is 200dpi. In this case the image is scaled up (blown up) to 300dpi.

In both of the cases described above the "PDF Load Resolution" determines how big the bitmap is into which ScannerVision will load the PDF. As has been mentioned already, this value plays an important role in the accuracy of OCRing. It also determines the size of output documents. For example, if you need to convert low resolution images like fax documents which are usually 1 bit (monochrome) and between 150dpi and 200dpi to PDF documents, there would be little point in producing 300dpi images to go into the PDF. It would be more sensible to set the "PDF Load Resolution" to the same value as the incoming documents' resolution. Conversely, say you need to reduce the size of PDF documents for backup purposes and the incoming documents are 600dpi or higher then setting the "PDF Load Resolution" to 300dpi would result in smaller PDFs with little loss in perceived quality at normal zoom levels.

The chances of image quality loss as a result of the rasterization and resampling process cannot be eliminated completely. If you have the option to use Tiff documents over PDF as input to ScannerVision we recommend Tiff.

10.5 Appendix E - Document Size

Various factors influence the size of output documents produced by ScannerVision including:

1. Size (A4, A3 etc.), resolution and color/bit depth of the input document.

2. PDF load resolution in the case where the input document is a PDF.

3. Color/bit depth of the output document.

4. Compression algorithm of the output document.

To reduce the size of output documents you can start by making sure the input documents are not larger than what they should be to unsure that your results are good enough. For example, if any processing you perform on the document requires any form of OCRing or barcode reading then the incoming document should have a minimum resolution of 300 dpi. If it is higher the improvement in OCR accuracy may not justify the increase in output document size. 300 dpi would be "good enough".

Secondly, higher bit depth documents such as 24bpp would normally be bigger than 8bpp or 1bpp ones. Select the lowest output bit depth that would still yield adequate color fidelity.

Lastly, the compression algorithm you choose for Tiff and PDF documents will have a big impact on the size of the output document. The default compression algorithm for Tiff and raster PDF

documents is LZW which is a good all round choice for most documents. If you only work with 1bit, black and white input documents CCITT or G4 would yield much smaller documents.

When the output document format is PDF/A or searchable PDF the default compression is algorithm is LZW for 1bpp documents. For all other bit depths it is JPEG for PDF/A and JPEG2000 for searchable PDF. If the output document size is too big when using the defaults you can experiment with other algorithms but the results may vary depending on the type of content in the incoming documents. If for example you are converting fax documents (i.e. 1bpp) with little graphic content to PDF/A you will probably get better results when using G4 compression or even CCITT since these algorithms were specifically designed for fax documents. In the case of documents with higher bit depths JPEG 4:1:1 will yield better results since a measure of quality loss is introduced when using this option.

10.6 Appendix F - Equitrac Authentication

ScannerVision allows for single sign-on with Equitrac when both the Equitrac and ScannerVision MFP clients are installed on an MFP. What this means is that when a user has been authenticated by the Equitrac client the user will also be authenticated in the ScannerVision client. This allows the user to sign in with the Equitrac client and then to switch to the ScannerVision client without the need to sign in again on the ScannerVision client.

This single sign on is made possible with the use of a facility offered by Equitrac server where it would call a user defined application for every user that signs in using the Equitrac MFP client application. When the Equitrac server calls the external application it passes it the path to an Xml file that is generated by the Equitrac server. The Xml file contains information about the user as well as the action that was perform i.e. sign in or sign out.

When ScannerVision is installed an application called EquitracNotifier.exe is installed in the root of the ScannerVision installation directory. This is the application that must be called by the Equitrac server to authenticate users in ScannerVision. The Equitrac Notifier application does not need to be installed on the same server as ScannerVision and it offers a facility that makes it easy to port it to other machines.

Equitrac configuration

A discussion of the configuration of the Equitrac system is beyond the scope of this manual. Please refer to the Equitrac documentation for details. Pay particular attention to how command line arguments are passed to external applications.

Command line arguments

When the Equitrac Notifier application is started without command line arguments the graphical user interface (GUI) is shown. The application accepts two command line arguments that have to be provided in the correct sequence. These arguments are "--file" followed by the fully qualified path to the Xml file generated by the Equitrac server. Below is an example of how the Equitrac Notifier application is called with command line arguments:

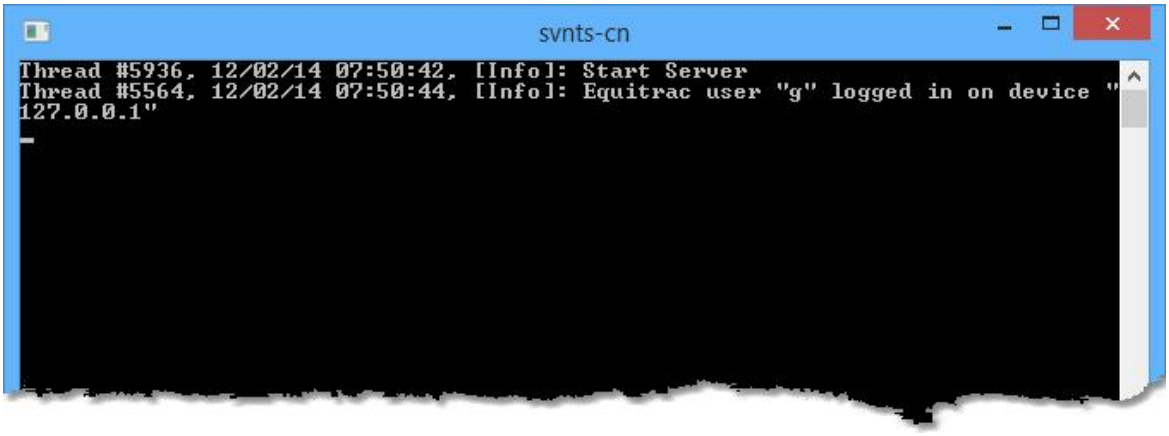
```
$Path$\EquitracNotifier.exe --file "C:\SomePath\EquitracActivateDevice.xml"
```

Note

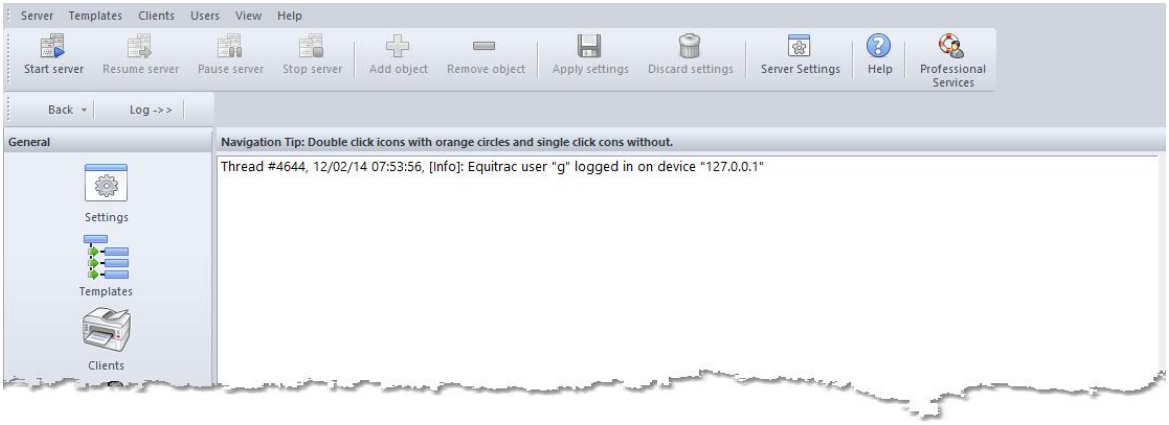
The example above only shows the **pattern** for calling the Equitrac Notifier application with command line arguments. The actual path to the application and xml file will be different for actual installations.

Operation

When the EquitracNotifier.exe is launched with the command line arguments shown above no graphical user interface is shown. In this mode the application reads the xml file, passes the relevant information to the ScannerVision Networking Server and if configured to do so deletes the xml file and then shuts down. If you are running the ScannerVision Networking Server in console mode you will see log output in the console window of all Equitrac activity, as shown below:



The Processing Engine User Interface's log window also shows Equitrac sign in/out activity:

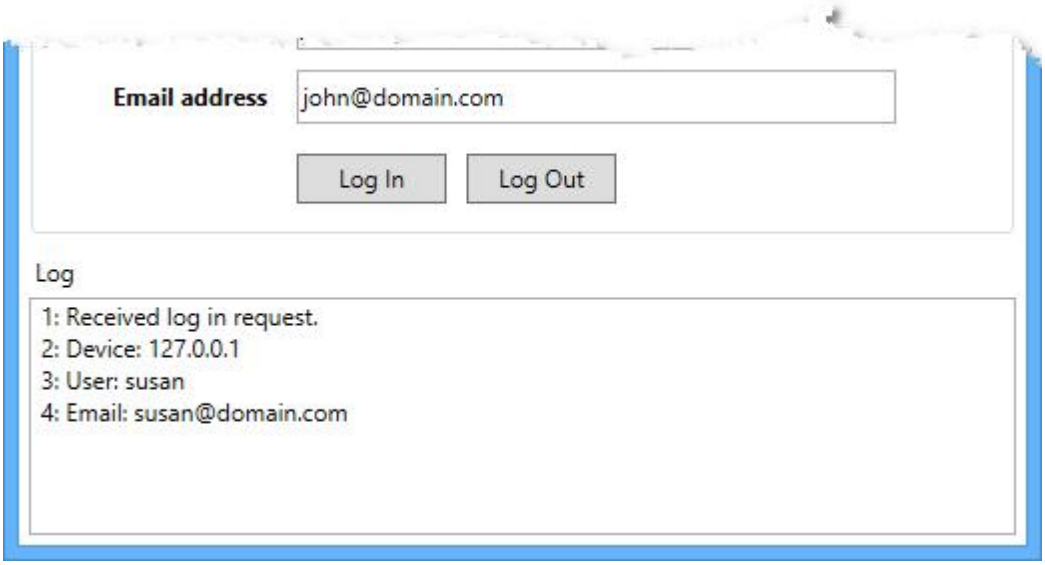


When the EquitracNotifier.exe is launched without command line arguments only the GUI is shown. With the GUI you can configure various settings which are discussed in the Equitrac Notifier Setup

section as well as perform test log ins and log outs.

Running instances

When the Equitrac Notifier GUI is open and a second instance of the EquitracNotifier.exe is launched **on the same computer** with command line arguments (e.g. a user signs in to Equitrac) the details for the user signing in or out is displayed in the log window of the GUI as shown below:

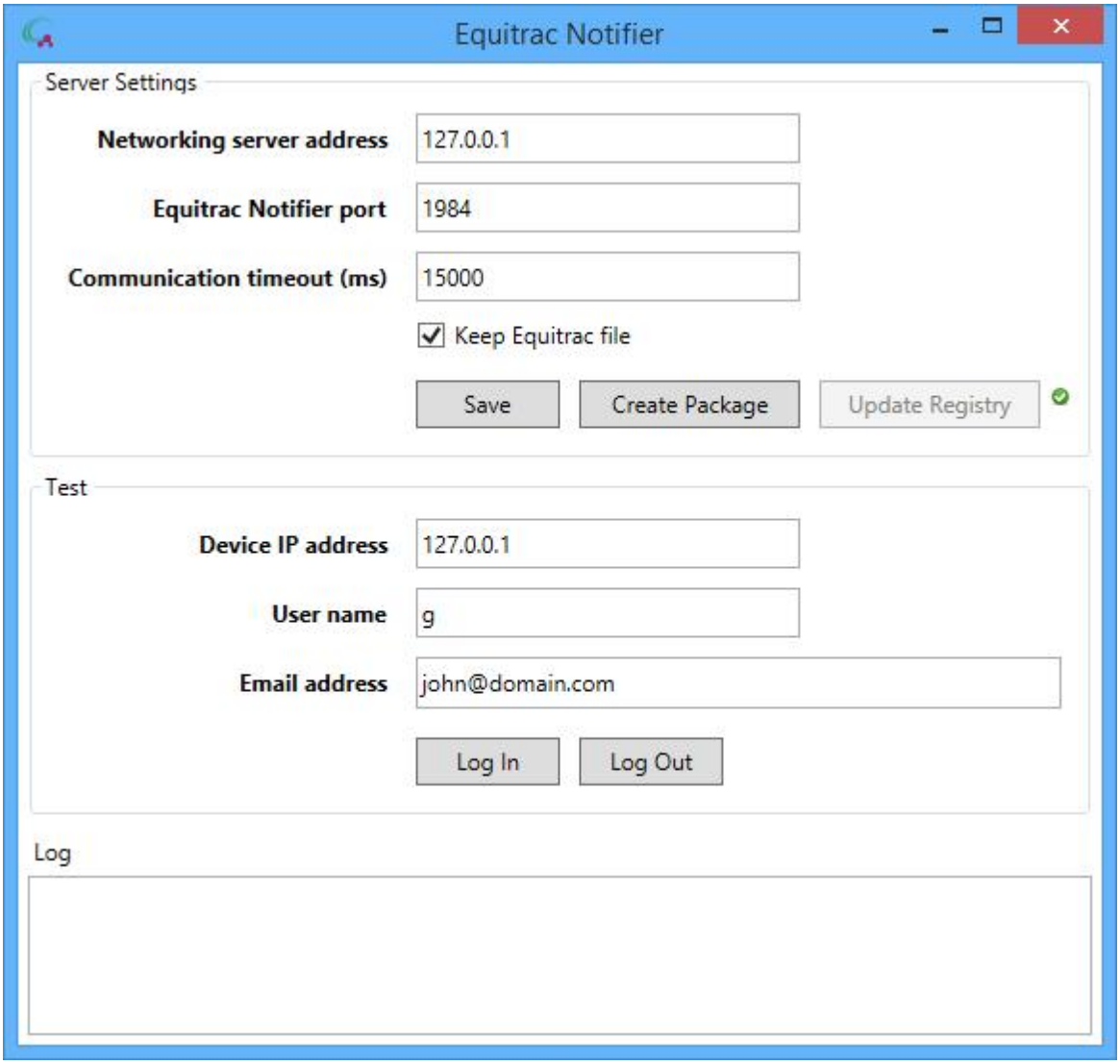


Note

Both the GUI instance and the command line instance must be run on the same computer. You cannot therefore view log activity of remote instances.

10.6.1 Equitrac Notifier Setup

The Equitrac Notifier application screen is shown below.



Server Settings

ScannerVision Networking Server address

The IP address of the ScannerVision Networking Server. If this value is left blank it is assumed that the ScannerVision Networking Server is installed on the local machine.

Equitrac Notifier port

The HTTP port on which the Equitrac Notifier communicates with the Networking Server.

Note

If you change this port you have to change it on the Networking Server also and you have to make sure that the port is allowed through the server's firewall.

Communication timeout

The timeout that is allowed for successful communication to the ScannerVision Networking Server. The default is 5000 milliseconds.

Keep Equitrac file

When this option is selected the Equitrac Xml file is not deleted after processing.

Save

Saves the current settings in the "C:\ProgramData\ScannerVision\Config\EquitracSettings.xml" file.

Create package

Creates a Zip archive of all the files that are necessary to run the Equitrac Notifier application. When you click this button you are prompted to select a folder where the package must be created. Make sure that the folder you select is completely empty otherwise you will get an error.

The "EquitracNotifier.zip" package that is created can be extracted on any machine into any desired directory.

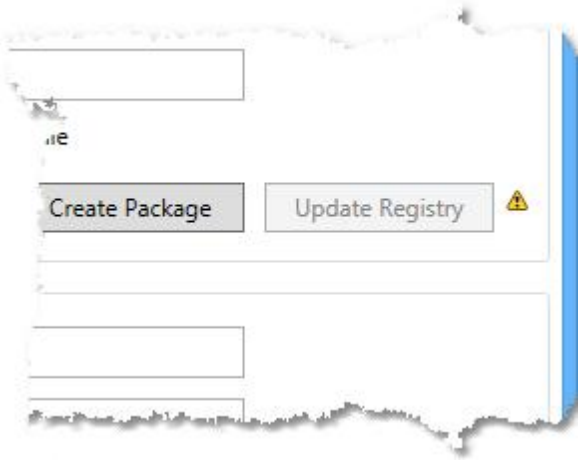
Update Registry

As of this writing the Equitrac server looks in a specific Windows Registry key for the path of the external application that is to be called when a user signs in or out of Equitrac. To simplify this process press the "Update Registry" button to update the registry key with the path of the Equitrac Notifier application.

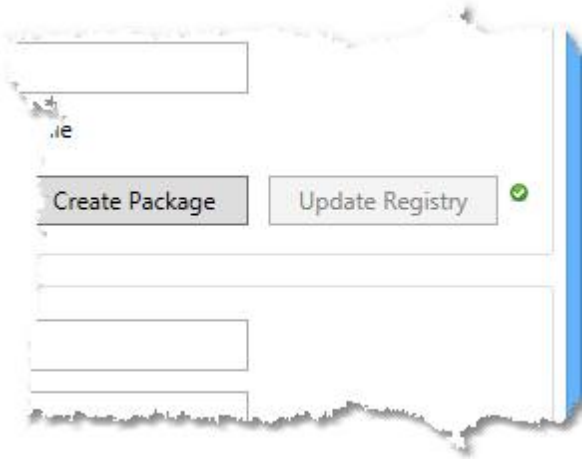
> **Note**
>
> You have to run the Equitrac Notifier application with administrator privileges before you will be able to use this facility.

If the "Update Registry" button is disabled there could be one of 2 reasons:

1. You are not running the application with administrator privileges **and** the registry key does not exist or does not point to the current path. If this is the case then a yellow exclamation mark will be shown next to the button like this:



2. The registry is up to date. In this scenario a green check mark will be shown next to the button like this:



Test

You can test your settings by simulating an Equitrac server sign in or sign out event. Provide the IP address of the simulated MFP client, the user name and email address of the user you want to sign in or out and then press the respective button of the action you want to simulate. The ScannerVision log window will show the activity if everything is configure correctly. Make sure that the ScannerVision Networking Server is running and that the log level is set to "Detailed".

Device IP address

The IP address of the MFP client on which the user is to be signed in or out.

User name

The user name of the user you want to sign in or out.

Email address

The email address of the user you want to sign in or out.

Log In / Out

Press the respective button to sign the user in or out.

10.7 Appendix G - Kyocera MFP Client Installation

In this section you will find a brief overview of the process to get the installed and configured on a compatible Kyocera MFP. More detailed information is available in the ScannerVision MFP Client user manual.

Here are the steps you should follow:

1. Download the latest version of the client from the Partner Portal onto a USB memory stick.
2. Insert the memory stick with the .pkg file into the USB port of the Kyocera MFP and proceed to install the application from the Admin tools section of the MFP screen. This is the standard installation procedure.
3. After installation is complete start the ScannerVision MFP Client. When the application is started for the first time the administration screen is shown automatically.
4. There are several categories of settings that you can configure. So do so select the respective category and when prompted for a password enter the default administrator password "0000" (four zeros).
5. At a minimum you have to configure the IP address of the ScannerVision Server. To do so select the "Server Settings" category and enter the administrator password.
6. Enter the ScannerVision Server IP address by tapping the "Server IP" button. Enter the IP address on the keypad that appears.
7. Click the "Save" button.

10.8 Appendix H - Unique Suffix

The File System and Ftp connectors provide several options to dynamically generate file names. This is useful in situations where you expect files to exist in the destination folder and you want to avoid file name conflicts. The file naming option are tabulated below:

Tag Name	Description
UNIQUESUFFIX	Append to or insert into the file name the first available number starting at 0
UNIQUESUFFIX+1	Append to or insert into the file name the first available number starting at 1
UNIQUESUFFIX2+1	Append to or insert into the file name the first available number starting at 01
UNIQUESUFFIX3+1	Append to or insert into the file name the first available number starting at 001
UNIQUESUFFIX4+1	Append to or insert into the file name the first available number starting at 0001

Windows File System Connector Only

When a "?" is added to the end of the unique suffix tag e.g. [UNIQUESUFFIX?], [UNIQUESUFFIX3+1?], the unique suffix is applied only if there is an existing file(s) in the destination.

The uniqueness of a file name is guaranteed by putting a number in the position where the tag is specified. Here are a couple of examples:

Document-[UNIQUESUFFIX]	=	Document-0, Document-1, Document-2 etc.
Document-[UNIQUESUFFIX2+50]	=	Document-50, Document-51, Document-52 etc.
Invoice[UNIQUESUFFIX4]	=	Invoice0001, Invoice0002, Invoice0003 etc.

The number after the word UNIQUESUFFIX e.g. UNIQUESUFFIX2 represents the number of zeros with which to pad the number. So if you specify UNIQUESUFFIX2 the number below 10 will be padded with 1 zero e.g. 01, 02 etc. If you specify UNIQUESUFFIX3 the numbers below 100 will be padded with 2 zeros e.g. 001, 002 and the numbers from 10 up to 99 will be padded with 1 zero e.g. 010, 011...098, 099. The same pattern holds for UNIQUESUFFIX4 except that 3 zeros are added to numbers below 1000, 2 zeros for numbers 10 to 99 e.g. 0010 and 0011.

If you don't include a "plus number" in the tag name the numbering starts at 0. The plus number indicates where numbering should start. When specifying "+1" the numbering starts at 1 e.g.:

1 UNIQUESUFFIX+1

01 UNIQUESUFFIX2+1

001 UNIQUESUFFIX3+1

0001 UNIQUESUFFIX4+1

For "+25" the numbering would start at 25 e.g.:

25 UNIQUESUFFIX+25

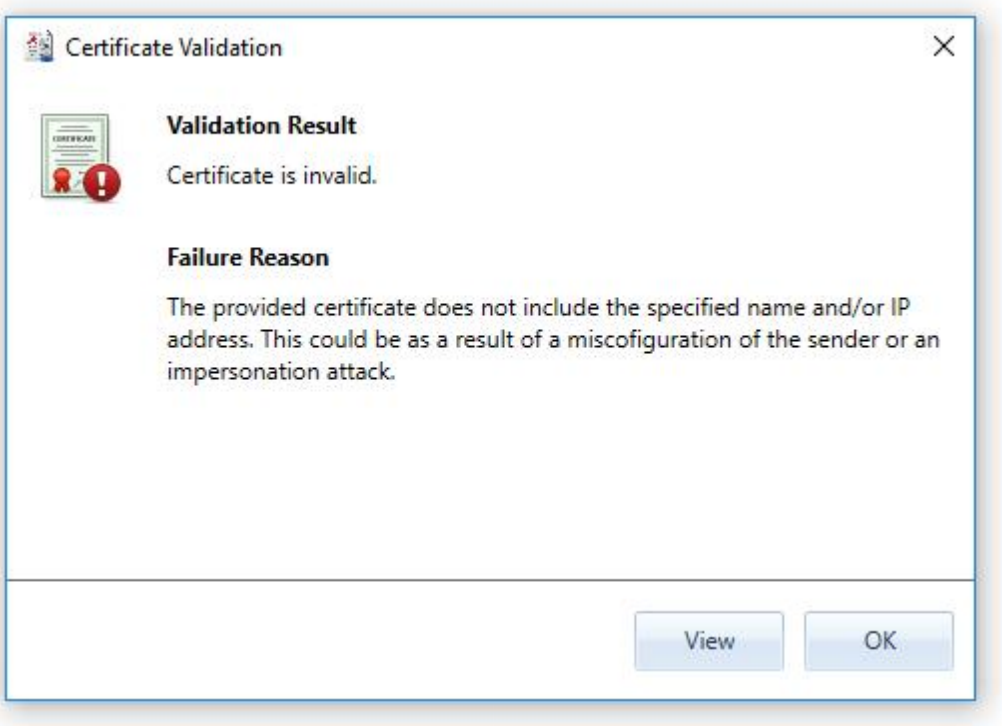
25 UNIQUESUFFIX2+25

025 UNIQUESUFFIX3+25

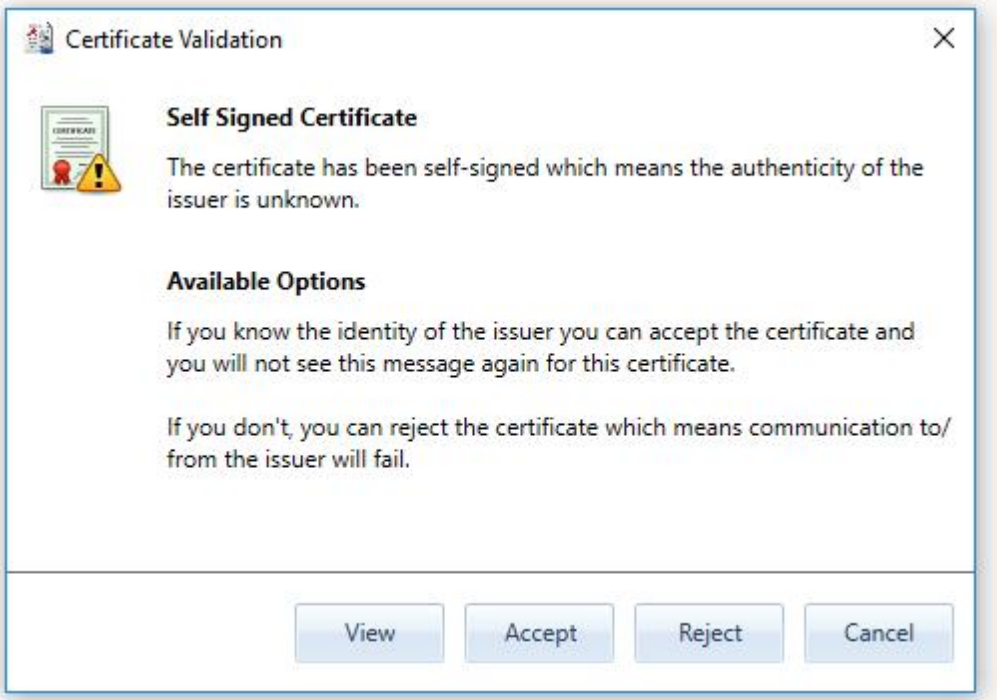
0025 UNIQUESUFFIX4+25

10.9 Appendix I - SSL Certificates

When connecting to a server over a secured (SSL) connection with for example, the WebDAV connector, you may be presented with one of the following popup dialogs:



You will see the above message when there is something wrong with the certificate. Read the reason for failure carefully and make sure that you are connecting to the correct server. Contact the server's administrator for further instructions.



This is a common scenario wneh SSL is enabled on a server that is not using a certificate that was issued by a trusted Certificate Authority (CA). Instead, the certificate was generated on the server itself rendering it as untrusted. However, if you DO trust the server and wish to allow the connection you can instruct ScannerVision to trust the certificate in one of 2 ways:

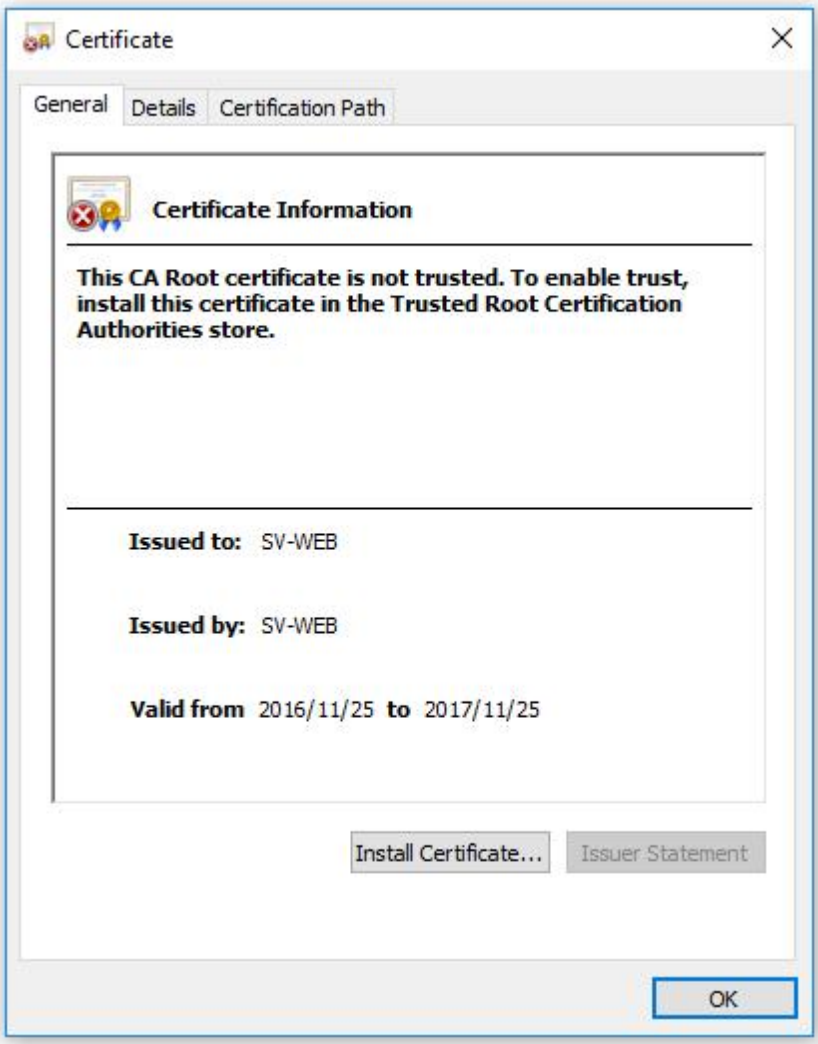
Accept

If you click the "Accept" button, ScannerVision generates a fingerprint of the the certificate and stores it in the "C:\ProgramData\ScannerVision\Config\SvCertificateStore.xml" file in the "acceptedcertificates" node. If you want to permanently reject the certificate, click the "Reject" button. The fingerprint of the certificate will be stored in the "rejectedcertificates" node of the file.

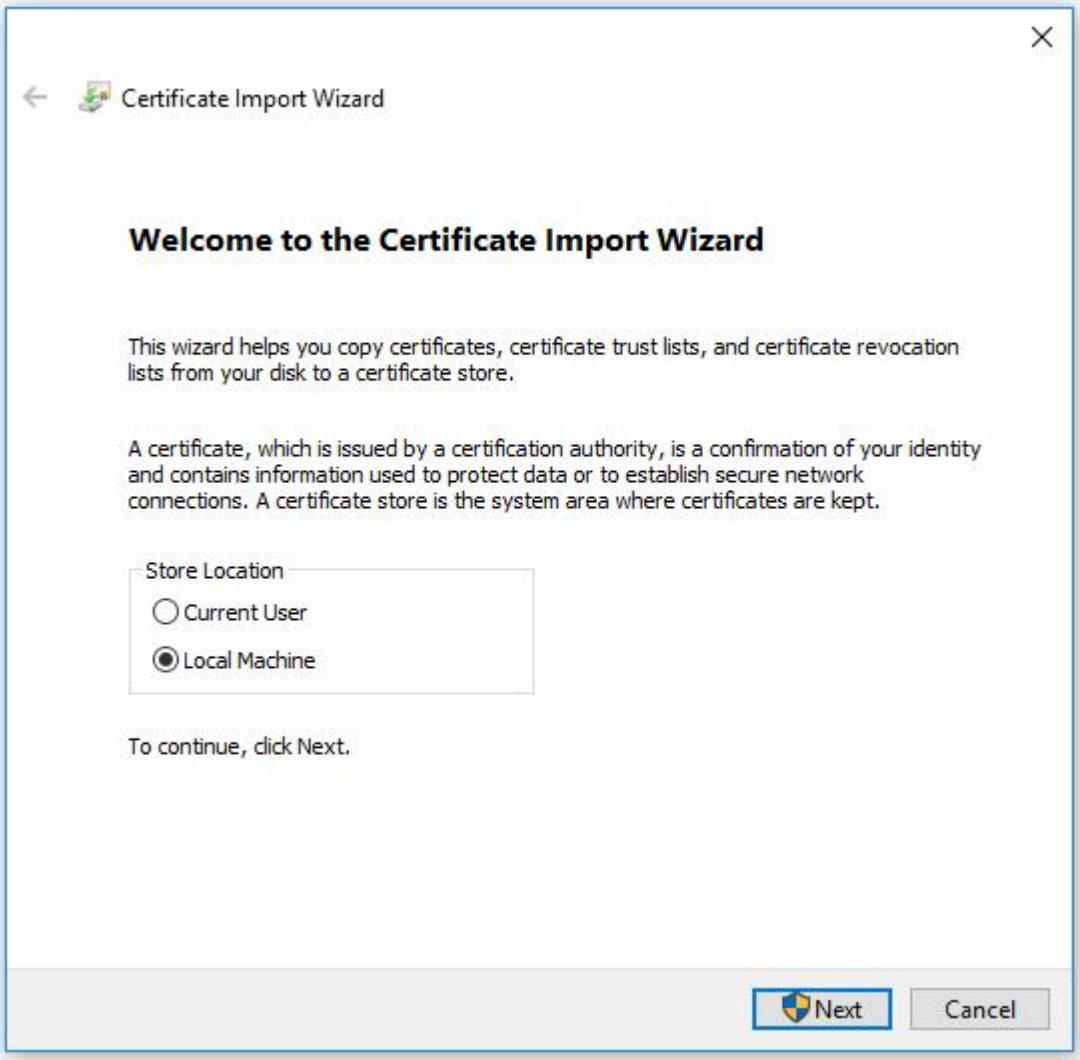
By accepting or rejecting the certificate this way, only ScannerVision is aware of your decision and it will not affect any other application.

Installing the certificate

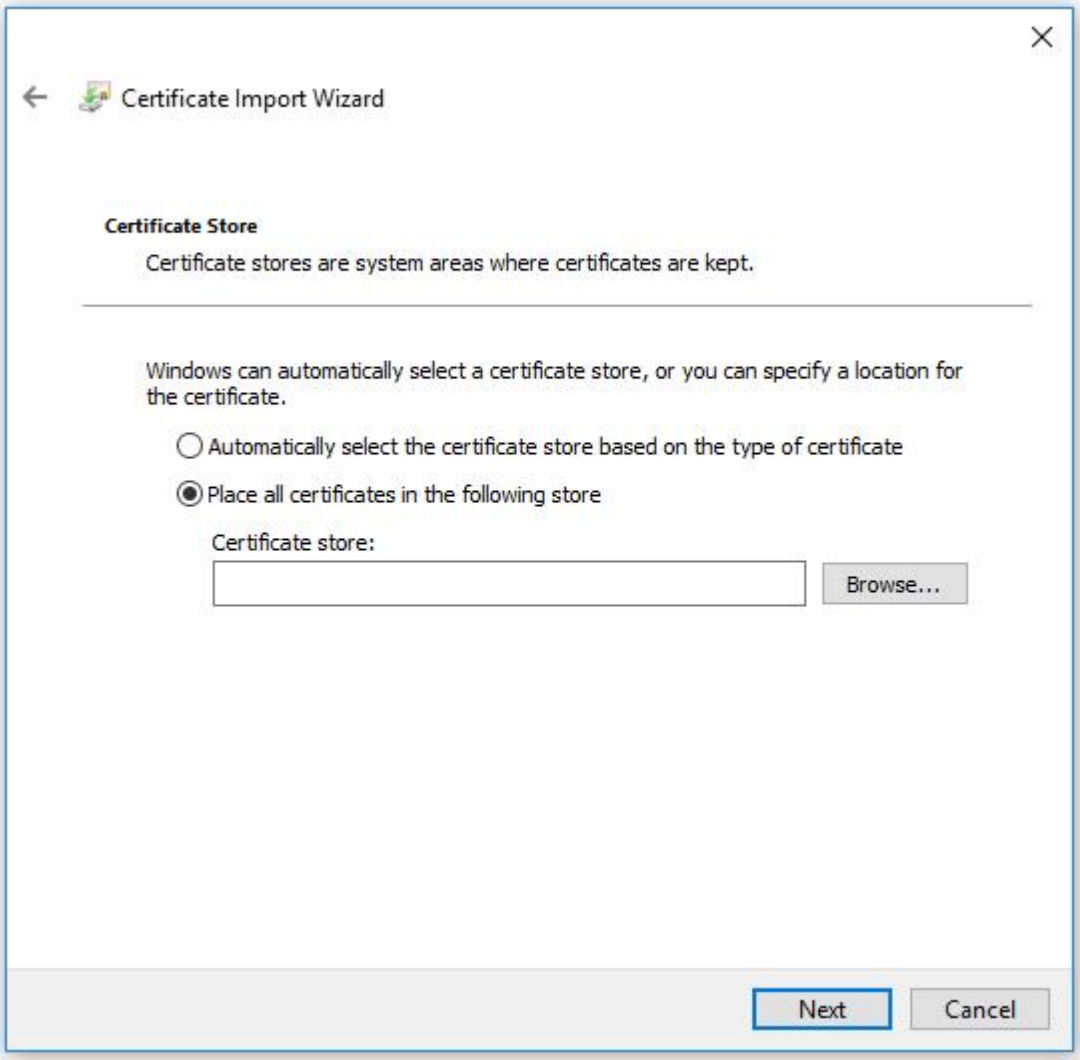
You may wish to install the certificate to your local machine. If you do this your operating system will trust the certificate and consequently all applications will trust it. To do this, click the "View" button which will display the following dialog:



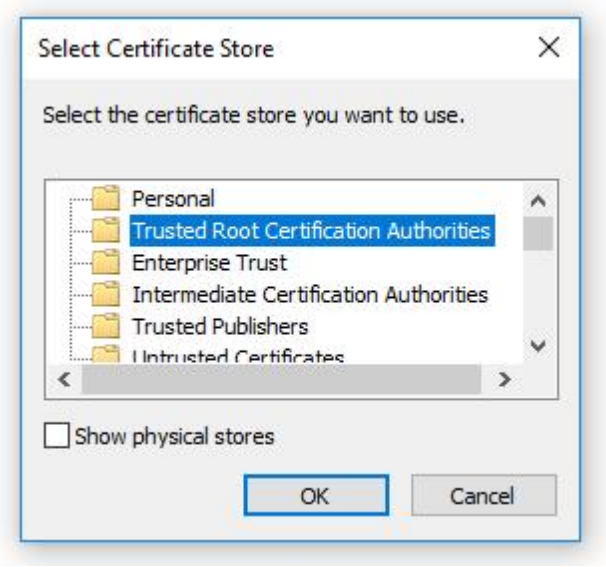
Verify that the certificate information is correct and that you can trust the certificate. If you are satisfied click the "Install Certificate" button. You will see the following dialog:



Select the "Local Machine" option and click the "Next" button. You will see the following dialog:



Select the "Place all certificates in the following store" radio button and click the "Browse" button. You will see the following dialog:



Select the "Trusted Root Certificate Authorities" option and click "OK" and then follow the prompts until completion.

11 Tutorials

11.1 ScannerVision Expression Editor

In the tutorials that follow we will be using the stand-alone ScannerVision Expression Editor application that is installed in your ScannerVision installation directory. You are welcome to work through the tutorials using the ScannerVision Expression Editor that you have launched from the ScannerVision application, just be aware that the last expression you enter will become part of your template if you save the expression. The stand-alone ScannerVision Expression Editor does not save your expressions anywhere.

The tutorials follow on each other so we strongly recommend that you follow along accordingly. The later the tutorial the less instruction is given since you would have covered that material in earlier tutorials.

Launching the ScannerVision Expression Editor

You are looking for and application called "ExpressionEditor.exe" which is installed in your ScannerVision installation directory. Its icon looks like this:

The default ScannerVision installation directory is:

32bit Windows: "C:\Program Files\ScannerVision\ScannerVision Server N"

64bit Windows: "C:\Program Files (x86)\ScannerVision\ScannerVision Server N"

N = ScannerVision major version.

11.1.1 Entering expressions with keyboard only

In this tutorial you will learn how to enter metadata expressions using your keyboard only.

1. Launch the ScannerVision Expression Editor.
2. If you don't see a blinking caret in the Metadata Expression window press the Tab key on your keyboard until the caret appears (or you could just click with your mouse in the window to give it focus).
3. Press the "[" key on your keyboard. A matching closing bracket "]" is inserted automatically and your expression is parsed. You will notice that you get a Syntax Error immediately. DON'T WORRY! This will go away once you have entered a valid expression.
4. With the caret between the square brackets like this: "[|]", type the word "DATETIME" or press Ctrl-Space and select the DATETIME menu option from the "Standard" group and press Enter. You now have a valid expression with parses successfully. You will see a date and time string appear in the Expected Output window and the word "DATETIME" will appear in the Name column of the Used Tags grid and the sample data in the Value column.
5. Position the caret in front of the closing square bracket using the arrow keys like this: "[DATETIME|]"

   We are now going to add a function

6. Press the Space bar and enter a opening round bracket "(" (Shift-9 on most keyboards). As with the square brackets, the editor automatically adds the matching closing bracket ")" for you. You will again see the Syntax Error as you did when you entered the square bracket which will go away when you've entered a valid function.
7. The caret should now be between the two round brackets like this: "[DATETIME (|)]". Type the word "split" or press Ctrl-Space, select the split menu option and press Enter. Notice that as you type (if you type slowly) the text between the round brackets are black with a red squiggly line underneath. These are aids that tell you that there is a syntax error in your expression and where it is.
   Once you have typed the word "split" completely the color of the font changes to blue. However, there is still a red squiggly line under the closing ")" bracket. If you read the error message in the Syntax Error window below you will see why. The split function takes a string parameter that tells it what to split on, and there needs to be white space before the parameter.
8. With the caret in front of the closing ")" bracket like this: "[DATETIME (split|)]" press the Space bar and enter a double quote (not two single quotes!), a space and then another double quote.

   Your expression should now look like this: "[DATETIME (split " ")]"

If you look at the Expected Output window you will see that the space between the date and time portions of the original string has disappeared! The reason for that is that the split function has separated the date and time portions into a string array and it has removed the space character. And because ScannerVision does an implicit "join" on functions that return string arrays as a result when they are the last function in the metadata tag, the two strings in the array are concatenated. Let's fix that.

9. With the caret between the closing ")" and "]" brackets enter another opening round bracket and enter the function name "take" with a integer parameter of 1. Your expression should look like this:

   "[DATETIME (split " ")(take 1)]"

   Now the output looks better. We have selected the 1 element in the 2 element string array that was produced by the split function.

You can now play around with different characters to split on and selecting different elements from the string array.

11.1.2 Entering expressions with keyboard and mouse

In this tutorial you will learn how to enter metadata expressions using your keyboard and mouse. We are carrying on from the first tutorial called Entering expressions with keyboard only so if you have not completed it yet please do so before you continue here.

1. Select all text in the Metadata Expression window (use your mouse or press Ctrl-A) and delete it.
2. Double click the DATETIME tag in the Available Tags tree view. The expression "[DATETIME]" appears in the expression window, a date time string appears in the Expected Output window and the Used Tags grid is updated with the DATETIME tag and its sample data.
3. Position the caret anywhere between the "[" and "]" brackets.
4. Double click the DATE tag in the Available Tags tree view. The DATETIME tag in the expression window is replaced with the DATE tag and the Expected Output window and Used Tags grid are updated accordingly.
5. Replace the DATE tag with the DATETIME tag.
6. Insert a space between the DATETIME tag and the closing "]" bracket.
7. With the caret in front of the closing "]" bracket, right click with your mouse and select the "split" function from the "Insert Function" menu. The caret automatically positions itself between the double quotes.
8. Enter a space between the double quotes.
9. Position the caret between the closing ")" and "]".
10. Using the same procedure as in step 7. insert the "take" function and enter 1 as the parameter.

To replace a tag or function name with the mouse, position the caret anywhere over the tag or function name, right click and select the new value from the menu.

11.1.3 Providing sample data

In this tutorial you will learn how to supply sample data for your tags.

1. Enter the following expression in the Metadata Expression window:

   "Document processed up until [DATETIME]: [MAINCOUNTER]"

   The Used Tags grid is updated with the DATETIME and MAINCOUNTER tag names and sample data.

2. Select the Value cell of the MAINCOUNTER row in the Used Tags grid and enter a new counter value.

3. Press Enter. The Expected Output window is updated with the new counter value you entered.