ITS Product
Guides

DocSlide

INFO TECHNOLOGY SUPPLY LTD.

DocSlide Configuration Guide v2.9.0

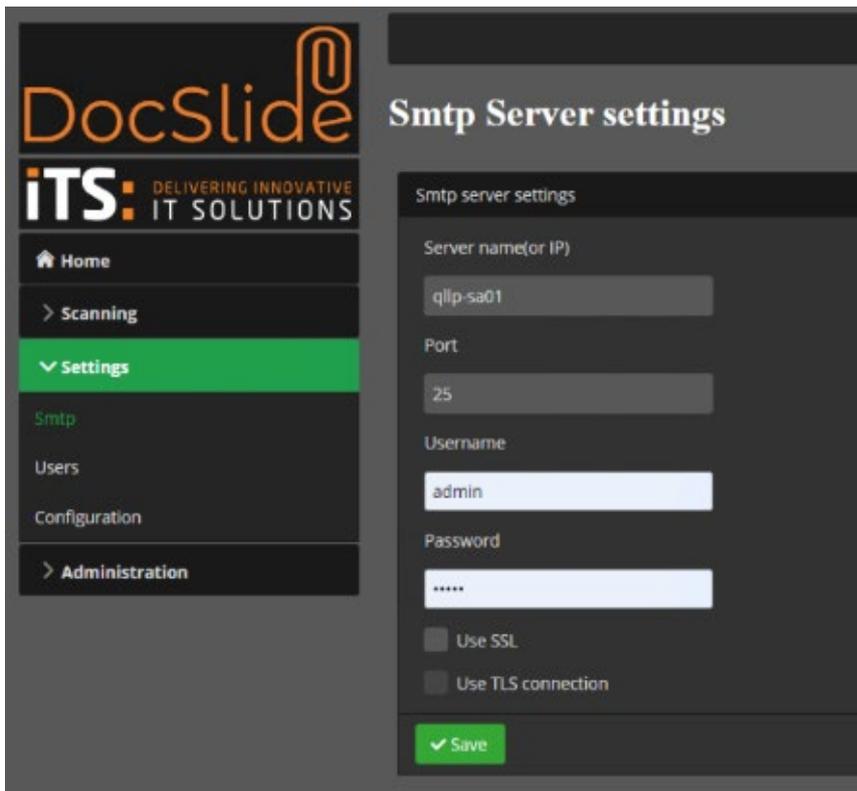iTS: DELIVERING INNOVATIVE
IT SOLUTIONS

Contents

Settings

To begin the configuration of DocSlide, we start with Settings for:

SMTP

Here you will set your SMTP server details. This will be used by DocSlide to send email links and notifications.

Option	Description
Server name	Name of the SMTP Server
Port	25
Username	SMTP Username
Password	SMTP Password
Use SSL	If SSL is required
Use TLS	If TLS is required

Run a test using the SMTP server configuration utility to send a test mail.
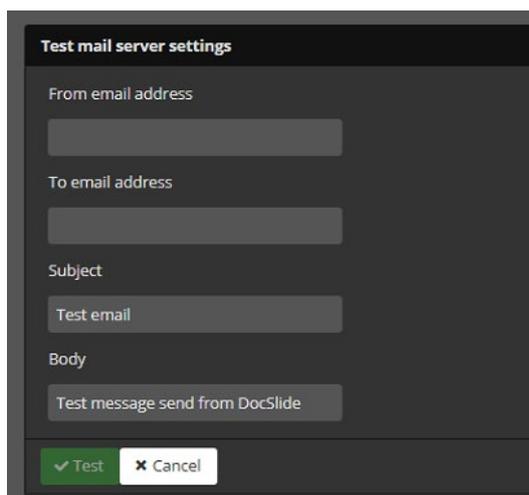
Users

Here you can set where DocSlide is going to search for user attributes.

LDAP

Enter the connection details for the LDAP server: URL, DN, UserName, Password and if necessary, some mapping filters, etc.
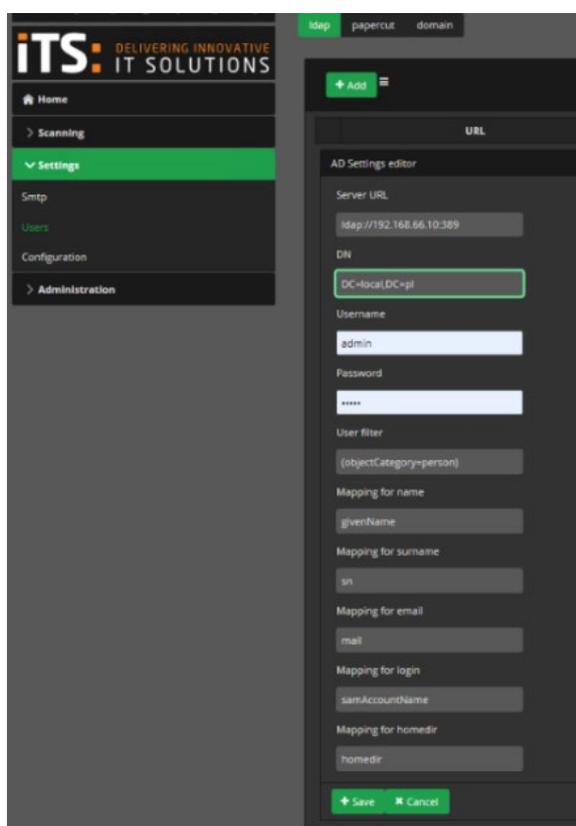
PaperCut

To connect to PaperCut, you will need to set the URL of the Papercut Server, and to add the web services auth token. This token is custom defined by yourselves and will need to be entered into the corresponding field in PaperCut as per the below:
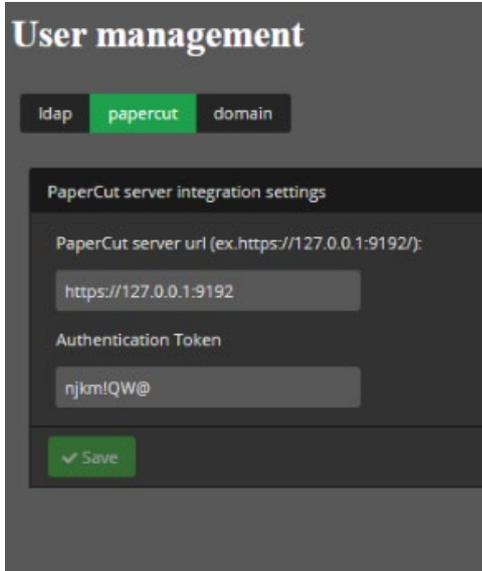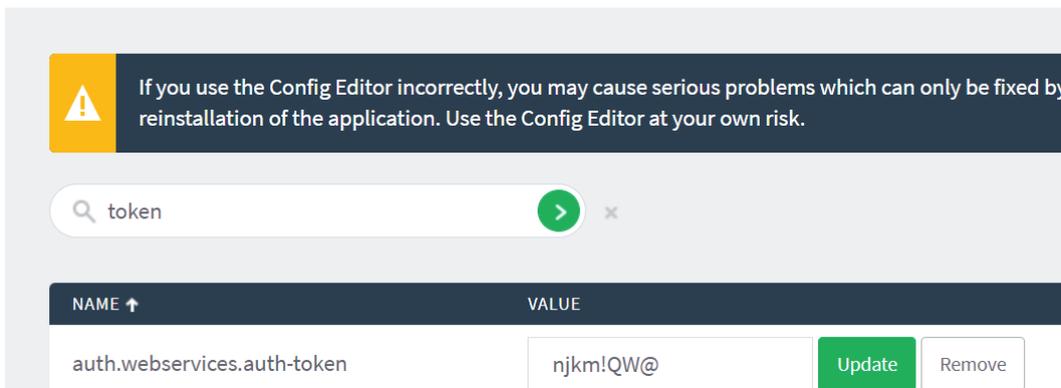
Domain

If you work in a Windows AD environment, DocSlide can bind to the domain natively if the DocSlide server is domain joined.
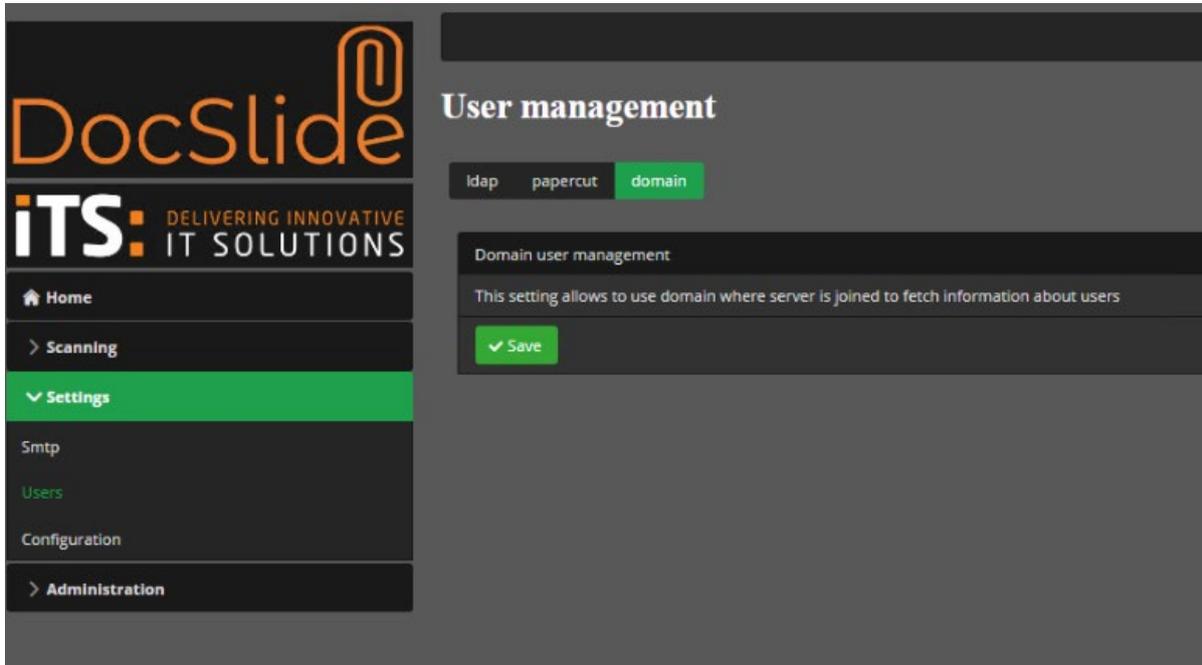
Simple

This setting allows to use simple mapping of user based on login or email address gathered by DocSlide durinh collecting files. No external services are connected but user details are limited to login and email address.

Groups

Here you can add groups to DocSlide from AD or Papercut depending of your user data source

Group list

Groups in the system

Name	
DS Portal	🗑

Search for group

Type at least 3 characters

Groups :

Configuration

Here you can change the system configuration.
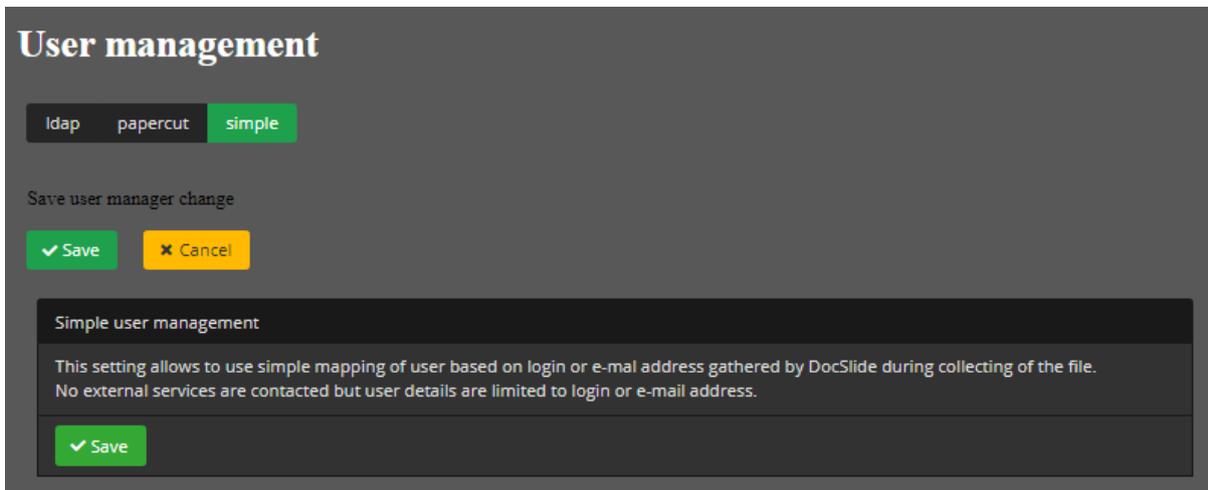
Configuration

Config key	Config value	
ocr-processor-max-concurrent-tasks	1	✓ 🗑 ↻
portal-authentication	windows	✓ 🗑 ↻
process-queue-max-concurrent-tasks	1	✓ 🗑 ↻
scan-download-audit	false	✓ 🗑 ↻
scan-download-owner-only	False	✓ 🗑 ↻
scan-storage-max-time	48	✓ 🗑 ↻
scan-store-location	store/scans	✓ 🗑 ↻
smtp-cert-checkRevocation	true	✓ 🗑 ↻
smtp-cert-checkValidation	true	✓ 🗑 ↻
tesseract-tessdata-path	./tesseract	✓ 🗑 ↻
web-certificate	Config value	✓ 🗑 ↻
web-limits-max-request	50000000	✓ 🗑 ↻
web-url	http://+:12120/	✓ 🗑 ↻
web-xeroxUpload-rootFolder	store/xerox	✓ 🗑 ↻

Config Key	Config value	Description
ocr-processor-max-concurrent-tasks	1,2,3….	number of simultaneous OCR processes - default 1;
portal-authentication	Windows/Papercut	Which user source will be checked for authentication
process-queue-max-concurrent-tasks	1,2,3….	number of simultaneous processors processing files in the queue, default (number of cpus - 1);
scan-download-audit	True/False	This allows you to monitor who downloads the file in the logs
scan-download-owner-only	True/False	
scan-storage-max-time	48	The number of hours DocSlide will store the scan before the file is deleted in hours
scan-store-location	store/scan	The internal folder for storing scanned jobs
smtp-cert-checkRevocation	True/False	Checking if cert is accepted by local smtp
smtp-cert-checkValidation	True/False	Checking of cert validity
tesseract-tessdata-path	./tesseract	path to the directory with the files above, e.g. ./tessdata or d:/files/tessdata;
web-certificate	Config value	The value of the server certificate if in use
web-limits-max-request	50000000	The maximum scan size DocSlide wil process (approx 50Gb) *do not change
web-url	http://<server-ip or hostname>:12120/	System URL
web-xeroxUpload-rootFolder	store/xerox	Xerox can be set to use https to a custom storage location
web-url	http://<server-ip or hostname>:12120/	System URL
scan-download-audit	false	This allows you to monitor who downloads the file in the logs

Web server

Here you can set Web server configuration (protocol, Listening address, port and ssl cert)

Administration

License info

Here you can find information about your license and M&S.

Password change

Here you can change the admin password.
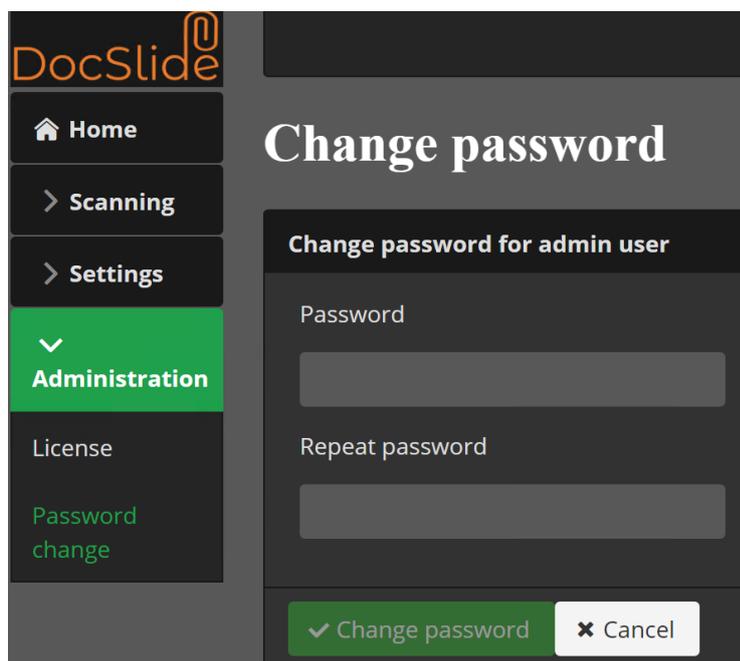
Scanning

Now we are going to configure destination, flows and processors.

Destinations

Adding destination

In the Scanning > Destinations menu select the +Add button
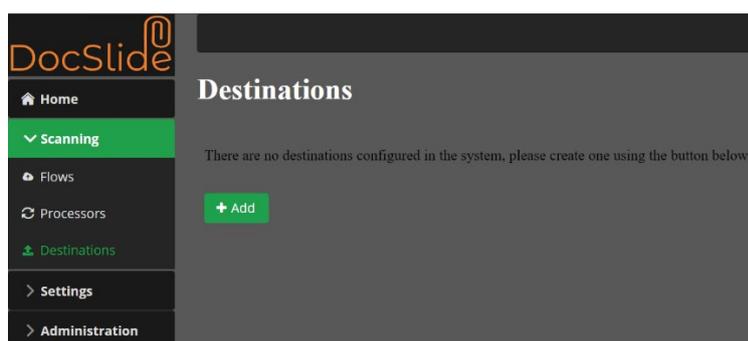
Destination to email

When you want to send scanned files or link to them to user email.

To add email destination, you need to set the following:

Option	Description
Destination name	Name of destination
From email	From mail address, static or variable
To email	To mail address, static or variable
Subject	Subject of the email
Mail Body	Body of mail with attachment

For files sent as an attachment, this destination is now ready to be saved.

If you would like to receive emails with encrypted https links you need to select the checkbox *Send as link* and finish the configuration like so:

Option	Description
Size limit	Maximum size of attachment in Kb
Link pattern	Link pattern of the link in the DocSlide Server
Subject	Subject of the email with link
Mail body	Body of the email with the link

* For possible variables just start typing or select from the drop-down menu and you will see the list of available options supported by the DocSlide solution.
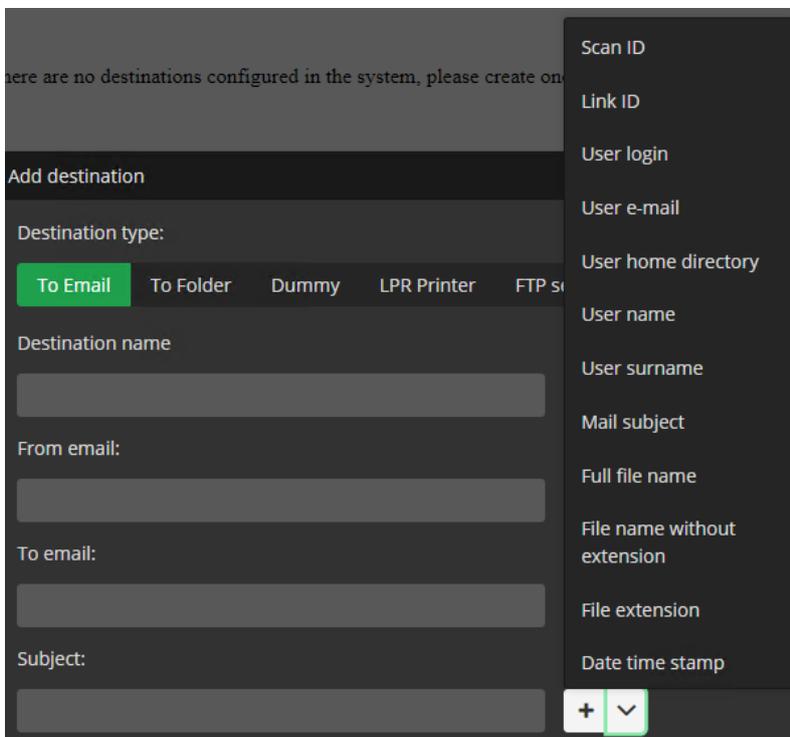
Destination to folder

Destinations can be set up for users to send files to the users home folder or a predefined folder. It can be also used to send files for further processing, e.g to a dedicated Umango Hotfolder which can be used for conversion or extraction of the data. In this case Umango will perform the OCR process and send the file to the predetermined location, or back to DocSlide which in turn would send the file via a HTTPS link in an email.

To add a folder destination, you will need to set the following:

Option	Description
Destination name	Name of the Destination
Folder name pattern	Name of the storage location
File name pattern	Name of the file, static or variable
Use authentication	If the file is to be stored on another server
Substitute characters in file name	File name variable substitions*
Create metadata file	Creates an xml file to accompany your file

* for possible variables start typing [and you will see list of available options supported in the solution.
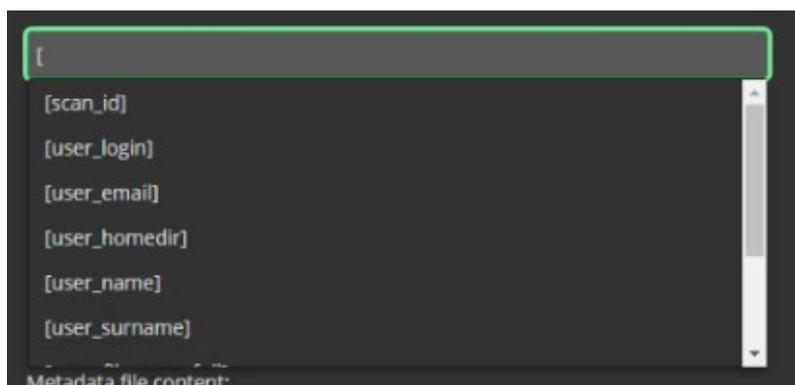
Dummy Destination

Dummy destinations can be used when we require the metadata from a scan but not to retain the fie itself. With dummy destination DocSlide will know everything about files because we will log everything in the log, including the files' metadata, possibly needed for processors.
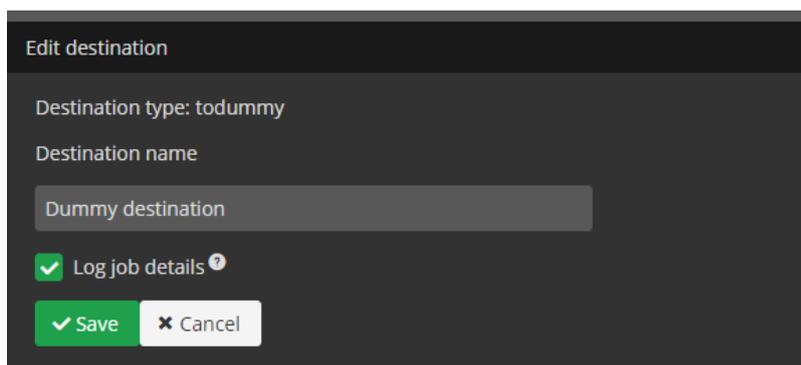
- Destination name   -   *Name of destination*
- Log job details   -   *Log job details*

Destination LPR Printer

DocSlide can use metadata to build the file details and send job to printer using the lpr protocol. The job can be set to a physical printers or shared server print queue. The print queue needs to be shared from a server with the LPR service running.

Option	Description
Destination Name	Name of destination
LPR server or hostname	Server Ipor Hostname for the Print server
LPR queue Name	Name of the LPR Queue
Port	LPR Port Number (default 515)
Job owner	Owner of the job
Document name	Name of the Document
Allowed File extensions	File extensions to be accepted by DocSlide

Destination FTP
DocSlide can send the file to an FTP or SFTP site as required.

Option	Description
Destination Name	Name of destination
Server hostname or IP	FTP Server hostname or IP
Port	FTP Port
Delivery method	FTP or SFTP delivery method
Job owner	Job owner
Send files using standard implementation of FTP protocol/Encryption	Types of encryptions and establishing secure connection
Reject untrusted certificates	Reject untrusted certificates
User name	FTP Username
Password	FTP User password
Server path pattern	FTP Server path
File name pattern	File name

Processors

To create the processor, you need to press *+Add.*

Archive

An archive processor gives you possibility to archive scanned jobs.

To add archive processor, you need to set:

Option	Description
Processor name	Name of the processor
Type	Archive
Archive folder	Folder path to archive jobs. Can use metadata
Excludes groups	Groups can excluded for personalisation
Excluded users	Users can excluded for personalisation

Merge

The "merge processor" provides the option to merge multiple pdf files into a single pdf.

To add "merge processor", you need to set the following:
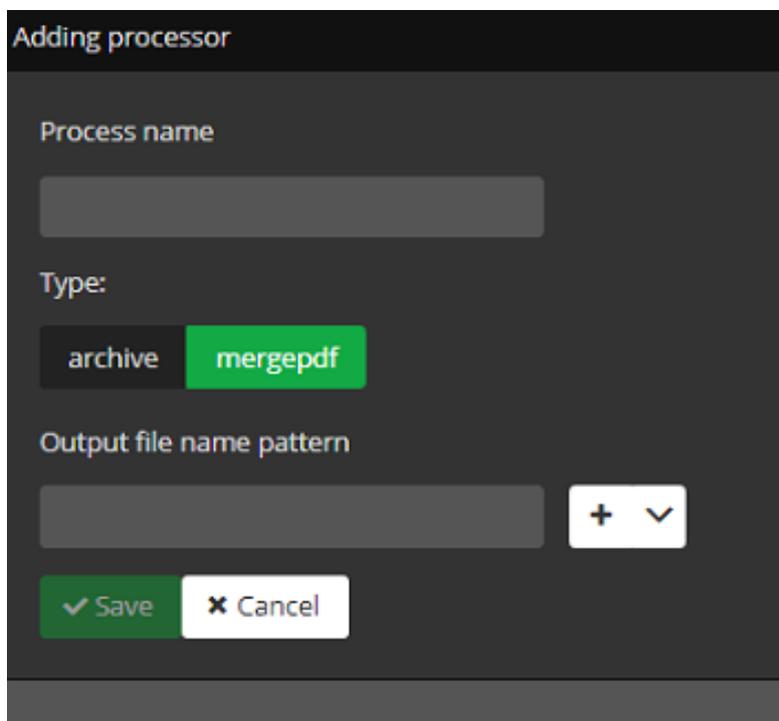
Option	Description
Processor name	Name of the processor
Type	mergepdf
Output file name pattern	Storage file path for the merged job

Digital Signature in PDF Documents via a Certificate

The "sign processor" provides the option to sign pdf documents with a company certificate.

Option	Description
Processor name	Name of the processor
Type	mergepdf
Certificate ised to sign PDF files	Available certificates on the server for signing
Signature name	Name of the signature
Signature contact	Contact for the signature
Signature location	Location of the signature
Signature reason	The reason for using the signature
Use time stamp server	IF the date can't be retrieved from the users' email can use a custom Time Stamp Server

Add PDF File Metadata to PDF documents.
The "metadata processor" provides the option to add metadata to pdf documents.

Option	Description
Processor name	Name of the processor
Type	Audit
Author	Author
Creator	Creator
Keywords	Keywords
Producer	Producer
Subject	Subject
Title	Title

Audit processor

The Audit processor give you the possibility to search inside files processed by DocSlide for sensitive keywords and therefore react accordingly. There are 2 modes available, Silent Monitoring and Blocking. Silent Monitoring allows the user to still receive job, but an admin will receive an email notification with the file and with the highlighted keywords. The alternative is to permanently block the user from receiving the job, in this case the user will receive a notification that the job will not be delivered, however the admin will receive the file with the keywords highlighted.

Option	Description
Processor name	Name of the processor
Type	Audit
Searched keywords in documents	place each new keyword or sentence in a new line and for Regular Expressions, prefix them with "regex"
Mark keyword in audit document	Highlight the specified keywords
Block sending file to user	Choose whether the file will just go to the admin or to the user as well
Notify user that their job was blocked	Notify user that their job was blocked
Notify mail sender	Email to be used to send the user notification
Notify mail recipient	Notify the user that sent the file
Notify mail subject	The mail subject of the user notification
Notify mail body	The mail body of the user notification
PDF files in jobs are only metadata	No copy of the job will be retained and only an image of the file, such as TIFF or JPEG
Audit mail sender address	The address from which to send the admin notification
Audit mail recipient address	The address to send the admin notification
Audit mail subject	The subject of the admin notification
Audit mail body	The body of the admin notification

OCR processor

The OCR processor gives you the possibility to convert tiff files to pdf searchable. DocSlide is using this processor to prepare files for Audit processor to search for keyword inside scan, print and copy jobs.
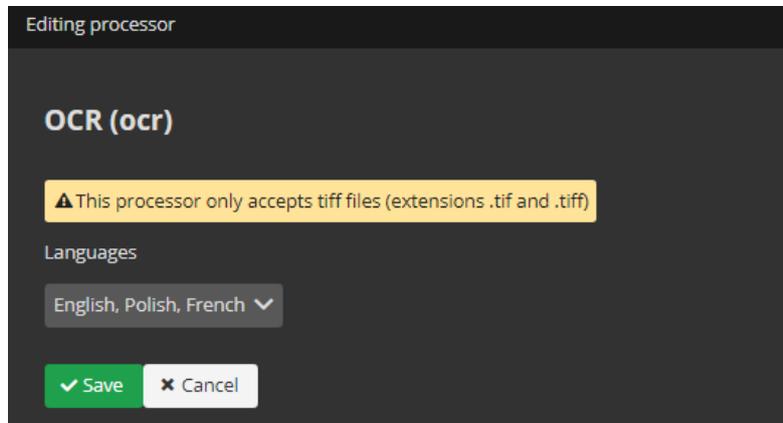
Option	Description
Language	You can choose OCR languages

DocPDF

The DocPDF processor gives you the possibility to convert doc and docx files to pdf searchable. DocSlide uses this processor to prepare files for Audit processor to search for keywords inside scan, print and copy jobs.
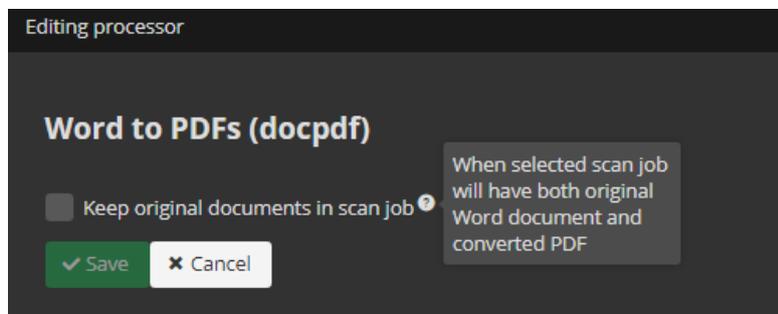
Option	Description
Keep original documents in scan job	When selected scan job will have both original scan job and converted PDF

ImageTiff

The ImageTiff processor gives you the possibility to convert all image files like (jpg, png, bmp) files to pdf searchable. DocSlide uses this processor to prepare files for Audit processor to search for keywords inside scan, print and copy jobs.

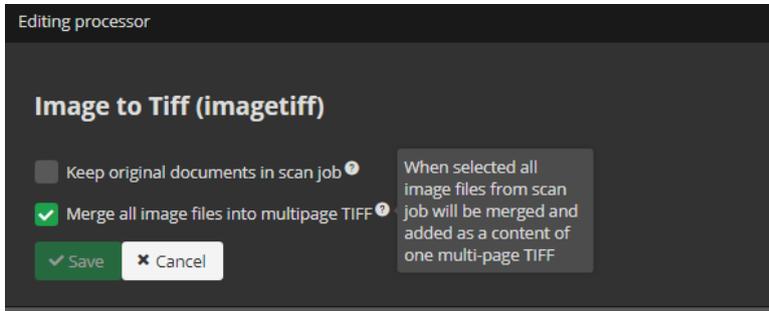| Option	Description
Keep original documents in scan job	When selected scan job will have also original scan job
Merge all image files into multiple TIFF	When selected all image files will be merge to one tiff file

PdfTiff

The PdfTiff processor gives you the possibility to convert all pdf non searchable files to pdf searchable. DocSlide uses this processor to prepare files for Audit processor to search for keywords inside scan, print and copy jobs.

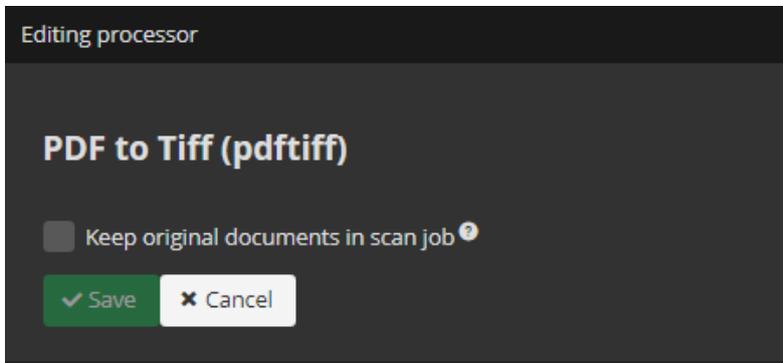| Option	Description
Keep original documents in scan job	When selected scan job will have both original scan job and converted PDF

Flows

Flows are the full lifecycle of the job, these are a combination of the input locations, Processors and Destinations to configure what should happen with the different files.

To create flows, press *+Add.*
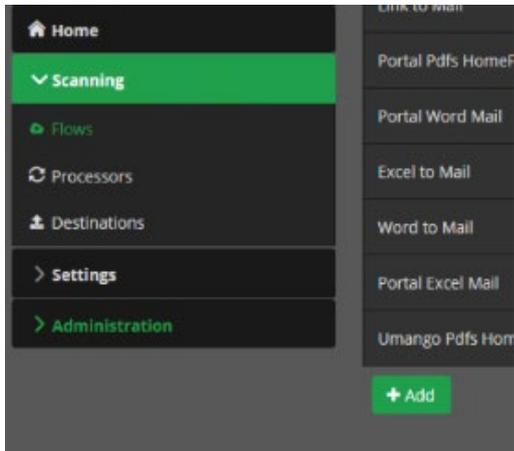
There are a few flow types possible:
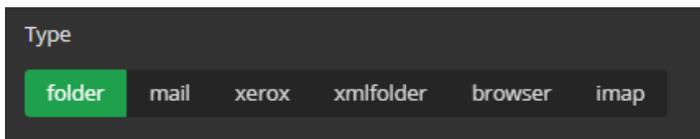
- Folder
- Mail
- Xerox
- XmlFolder
- Browser
- Imap

Folder Flow

Folder Flows take the file from a nominated Hot Folder to be processed. The file name could be set to include metadata, for example user login. Based on this we can use DocSlide to perform actions such as to send this scanned file to the user's email or homefolder, etc.

To create folder flow, you need to set:

Adding flow

Flow name

Type

| folder | mail | xerox | xmlfolder | browser | imap |

Folder path

☐ Include subfolders

File mask

Separator in the filename

Metadata pattern

+ ∨

Editing flow Umango Pdfs HomeFolder

Folder path

C:\DocSlide\Input\Folder

File mask

.

Sepataror in the filename

_

Metadata pattern

{login}

Processors

Archive

Deliver to

☑ Simple mode

| Mail | Portal_Umango_HomeFolder | Umango Mail Excel | Umango Word | Home Folder |

Option	Description
Flow name	Name of the flow
Type	Folder
Folder path	The path or the folder the flow will monitor
File mask	*.* will permit all fileyes to be used in this Flow
Separator in the filename	the separator to take the data up to for the filename
Metadata pattern	the metatadata of the username or email
Processor	the name of the processor if required
Destination	select from the list of preconfigured destinations
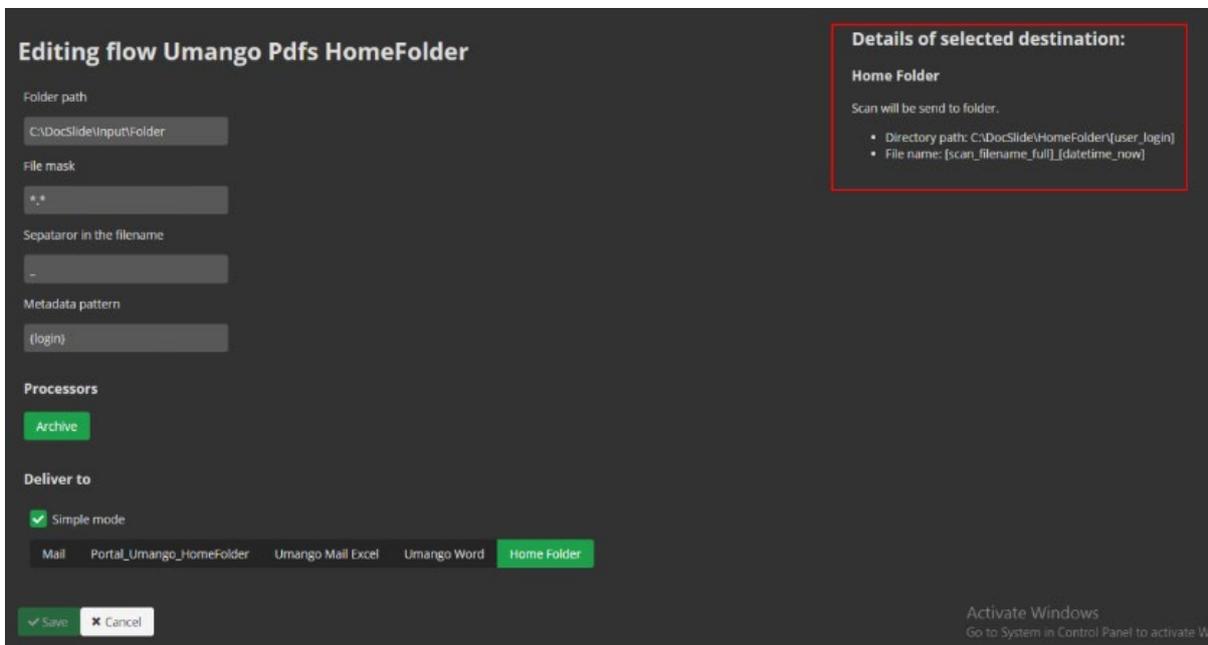
Mail flow

To receive the processed files from an email, DocSlide has an embedded SMTP server and can receive emails from MFDs. Useful in cases where the MFD cannot support PaperCut's Integrated Scanning feature.

To create a mail flow, you will need to set:

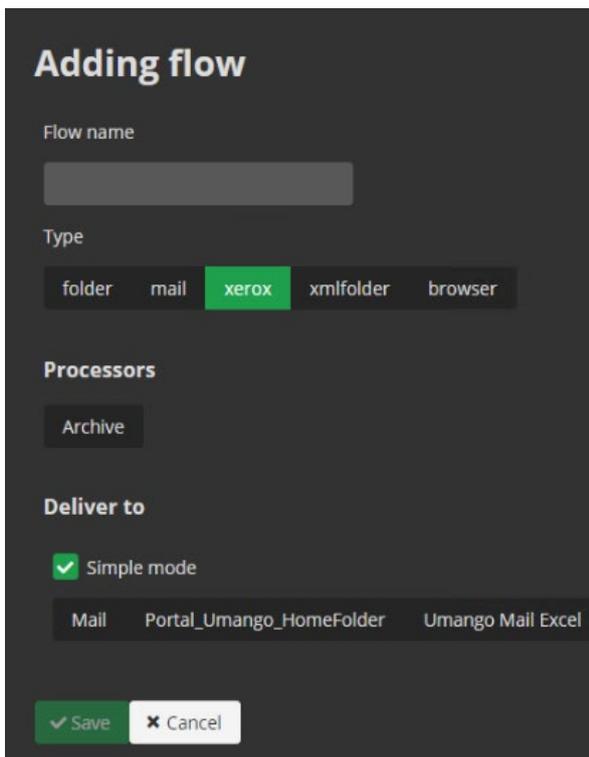| Option	Description
Flow name	Name of the flow
Type	Folder
Server Name	DocSlide server name or IP Address
Port	25
SSL Protocols	The form or SSL required
Certificate	The server certificate (if in use)

Processors	Which preconfigured Processor to use
Destination	Which preconfigured Destination to use

Xerox flow

Xerox MFDs have the native ability to send scanned files via HTTPS. DocSlide serves as the scanning endpoint with the address of http(s)://*<server_ip>*:(*port*)/xerox/scan. DocSlide will receive the file and parse an XST metadata file including user and other variables. DocSlide will then process this as required and output to destinations e.g., user email or homefolder, etc.

To create a Xerox flow, you need to set:

Option	Description
Flow name	Name of the flow
Type	Xerox
Processor	Processor if required
Destination	Output destination

How to configure Xerox device to integrate via https with DocSlide

We need to set https destination which will be DocSlide server:
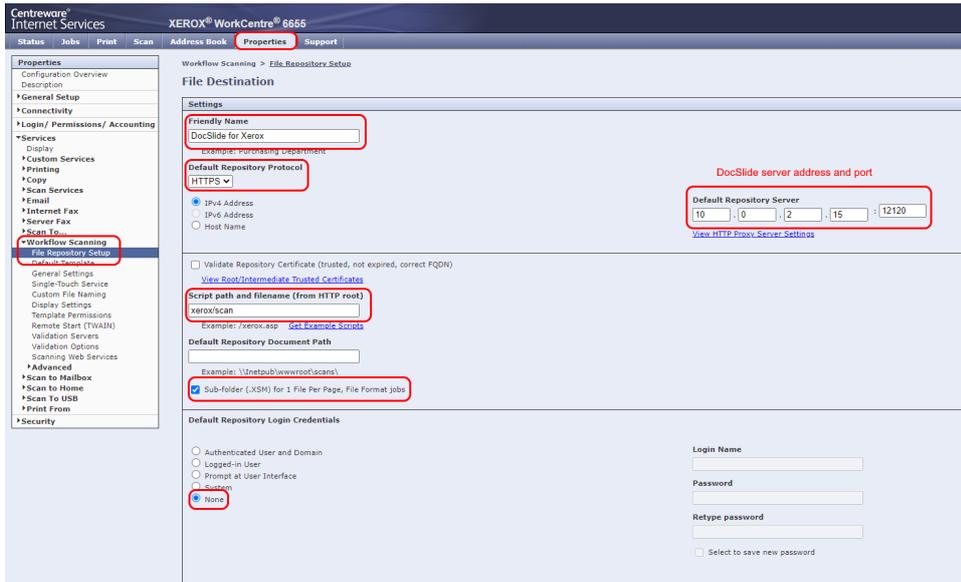
We need to create Workflow Scan Template.

We need to select destination created previously.

Also, we will use Workflow tags to create metadata aligned to proper DocSlide workflow.
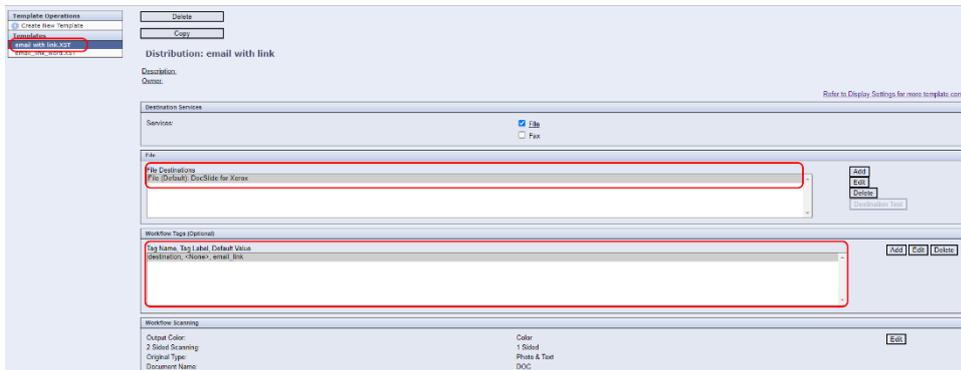
We need to set file extensions.

We need to create Workflow tab - "destination" with value "email_link" to align this to standard workflow to send encrypted link to user email.

We need to create Workflow tab - "destination" with value "email_link_word" to align this to workflow which will send file to do OCR in Umango and then send encrypted link to user email. Same process we can user for other extension.

In DocSlide we need to user Advance destination to align previously created Workflow Tags in Xerox device with proper destinations

XmlFolder

DocSlide can receive files with a corresponding XML file to gather the required metadata. DocSlide will receive the file and parse xml metadata file for user and other variables and then process this to preconfigured destinations such as: user mail, homefolder, etc.

To create Xerox flow, you need to set:

Option	Description
Flow name	Name of the flow
Type	xmlfolder
Folder path	The Input folder
Xpath to owner login	Path in the xml for the username
Xpath to Filename	Path in the xml for the filename
Add metadata to be extracted from XML	Any additional metadata in the xml to be extracted
Processors	Processors if required
Destination	Preconfigured Destinations

Browser

The Browser Flow allows you to upload files for the DocSlide process from a User web portal. *(Please ensure you have purchased the DocSlide Upload Module).* DocSlide will receive the files and based on the preconfigured Flow, send the processed scanned files to the users email, homefolder or Umango for further features etc.

To create Browser flow, you need to set:

Option	Description
Flow name	Name of the flow
Type	Browser
Flow label	Name of the flow - this will be visible to the users
Description of the flow	Description of the Flow - this will be visible to users
Allow file extensions	File extensions permitted
Processors	Select from preconfigured processor is required
Destination	Select from the list of preconfigured Destinations

IMAP

The Browser Flow allows you to upload files for the DocSlide process from a User web portal. *(Please ensure you have purchased the DocSlide Upload Module).* DocSlide will receive the files and based on the preconfigured Flow, send the processed scanned files to the users email, homefolder or Umango for further features etc.

Option	Description
Flow name	Name of the flow
Type	Browser
Server	IMAP server IP or DNS Name
Port	Imap server port
Use SSL	SSL or not
Use TLS connection	TLS or not
User name	Imap user name
Password	Imap user password
Test settings	Test option
Interval in minutes	Interval in minutes to check mailbox
Allowed file extensions	Setting allowed files extensions
Allow any sender	In case of not finding sender email in user source, this email will be processed

Appendix

When using Folder as a Flow type, in metadata pattern you can use {login} or {email} to extract information about file's owner from the scan file name. Based on this information DocSlide will fetch the owner of the job from the defined source of users (domain, PaperCut, or LDAP) based on this information. Any additional options will be treated as metadata bound to the current scan file, e.g:
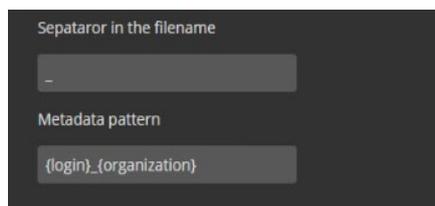
For example, if a file named jsmith_its_scan.pdf is submitted, DocSlide will get details of the user based on the login *jsmith*. Also the metadata can include [organization] so it will be available to be used with the processors and destination, in this case it will have value of *its.*

You can use predefined metadata values in a form of [metadata_name], DocSlide will substitute it with proper value assigned to a scan file being processed:

Glossary

Metadata name	Availability condition	Metadata replacement description
[date(*dateFormat*)]	Always	Current date and time in specified format, ex: yyyy – year, ex. 2020, MM – month, ex. 05 for May, dd – day of month, ex. 07 for 7th May, HH – hour, 24 hours based, mm – minute, ss – second, fff – miliseconds (more described at https://docs.microsoft.com/en-us/dotnet/standard/base-types/custom-date-and-time-format-strings)
[datetime_now]	Always	Current date and time in format *yyyyMMddHHmmssfff*
[user_login]	Always	Scan owner's login
[user_name]	Always	Scan owner's given name
[user_surname]	Always	Scan owner's surname
[user_email]	Always	Scan owner's email address
[user_homedir]	Always	Scan owner's home directory
[from]	Source is mail	Sender of mail message received by DocSlide
[to0]-[ton]	Source is mail	Recipient of mail message received by DocSlide (first recipient's mail is [to0], nth recipient is [ton]
[cc0]-[ccn]	Source is mail	Copy address of mail message received by DocSlide (first mail address is [cc0], nth address is [ccn]
[bcc0]-[bccn]	Source is mail	Hidden copy address of mail message received by DocSlide (first mail address is [bcc0], nth address is [bccn]
org_file_full	Source is browser	Full file name uploaded by DocSlide Portal
org_file_ext	Source is browser	Extension of file name uploaded by DocSlide Portal
org_file_wo_ext	Source is browser	File name without extension of file uploaded by DocSlide Portal
scan_filename_full	Destination is folder	Full file name imported by DocSlide
scan_filename_raw	Destination is folder	File name without extension of file imported by DocSlide
scan_filename_ext	Destination is folder	Extension of file name imported by DocSlide