Aquaforest

Kingfisher
Reference Guide

Aquaforest

Making Content Findable

Kingfisher
Reference Guide

Version 2.5
April 2024

© Aquaforest Ltd 2001 - 2023

Content

1 INTRODUCTION 5

1.1 System Requirements 5

1.2 Installation 5
1.2.1 User Permissions 5
1.2.2 Service User 6

1.3 Licensing and Purchasing 7

1.4 Version Breakdown 8

2 KINGFISHER USER INTERFACE 9

2.1 Dashboard 9

2.2 Job Settings 10
2.2.1 Job Definition 11
2.2.1.1 Rename PDF File 12
2.2.1.1.1 Rename based on Text Match 12
2.2.1.1.2 Rename based on Barcode 13
2.2.1.1.3 Rename based on Text Match and Barcode 13
2.2.1.2 Split PDF Files 14
2.2.1.2.1 Split File based on Barcode 14
2.2.1.2.2 Split File based on Text Match 15
2.2.1.2.3 Split File based on Text Change 16
2.2.1.2.4 Split File based on Page Ranges 17
2.2.1.3 Extract Pages from PDF Files 18
2.2.1.3.1 Extract Pages from file based on Barcode 18
2.2.1.3.2 Extract Pages from file based on Text Match 18
2.2.1.4 Export PDF Content 19
2.2.1.4.1 PDF to Text 19
2.2.1.4.2 PDF to CSV/XLSX 19
2.2.1.4.3 Advanced Export to CSV/XLSX 20
2.2.2 Location Settings 21
2.2.2.1 Location Types 22
2.2.2.1.1 File System 22
2.2.2.1.1.1 Local, Mapped and UNC paths 22
2.2.2.1.2 SharePoint (On-Premises) 23
2.2.2.1.3 Office365 (SharePoint OnLine) 25
2.2.2.1.3.1 Username and Password 25
2.2.2.1.3.2 Modern Authentication 26
2.2.2.2 Apply Input Filters 27
2.2.2.2.1 Filter Locations by Regular Expression 27
2.2.2.2.2 Exclude Locations 28
2.2.3 Select Variables 29
2.2.3.1 Template file 29
2.2.3.2 Variables 29
2.2.3.3 Zones 29
2.2.3.4 Text Match Option 31
2.2.3.5 Barcode 33
2.2.3.6 Text Pattern Option 34
2.2.3.7 Refine Extracted Text 35
2.2.4 Scheduler 36

2.2.5 Alerts 37
2.2.6 File Naming 38
2.2.6.1 Name Template 39
2.2.6.1.1 File naming example 40
2.2.7 Advanced Settings 41
2.2.7.1 Security 41
2.2.7.1.1 Source Document Security 41
2.2.7.1.2 Target Document Security 41
2.2.7.2 Properties 42
2.2.7.2.1 Display Modes 42
2.2.7.2.2 Settings 42
2.2.7.2.3 Bookmark Settings 42
2.2.7.3 PDF Metadata 43
2.2.7.4 OCR Settings 44
2.2.7.5 Barcode Settings 45
2.2.7.6 Other Settings 46
2.2.8 Running Job 47

2.3 Options 48
2.3.1 License Keys 48
2.3.2 Email 49
2.3.2.1 Username and password 49
2.3.2.2 Modern (OAuth2) Authentication 50
2.3.3 Theme 51
2.3.4 Advanced Settings 51

2.4 Help 52

3 KINGFISHER OPERATIONS 53

3.1 Rename PDF File 53
3.1.1 Rename Based on Barcode 54
3.1.2 Rename Based on Text Match 54
3.1.3 Rename Based on Barcode and Text Match 54

3.2 Split PDF File 55
3.2.1 Split File based on Text Match 55
3.2.2 Split File based on Text Change 57
3.2.3 Split File based on Barcode 58
3.2.4 Split File based on Page Ranges 59

3.3 Extract Pages from PDF File 60
3.3.1 Extract Pages from File based on Text Match 60
3.3.2 Extract Pages from File Name based on Barcode 61

3.4 Extract Content from PDF File 62
3.4.1 Extract Text from PDF File 62
3.4.2 PDF to CSV/XLSX 62
3.4.3 Advanced Export to CSV/XLSX 62

4 CREATING AND RUNNING JOBS WITH THE GUI 63

4.1 Creating a new Job 63
4.1.1 Error Codes 70

4.2 Editing a job 71

5 EXTRACT TABULAR DATA FROM PDF 72

5.1 Export to CSV/XLSX **72**
5.1.1 The Document View 73
5.1.2 Table View 75
5.1.3 Tables 75
5.1.4 Grouped Tables 76
5.1.4.1 Grouping Criteria 76
5.1.4.1.1 Group by Column Count 76
5.1.4.1.2 Where the rows below are the same 76
5.1.4.1.3 Where the cells below are the same 77
5.1.4.1.4 Where cells below match expression 77
5.1.4.1.5 Where Columns below match expression 77
5.1.4.2 Exclusion criteria 78
5.1.4.2.1 Exclude by Table ID 78
5.1.4.2.2 Exclude cells below match expression 78
5.1.4.2.3 Exclude Where Columns below match expression 79
5.1.5 Data grid 79

5.2 Advanced Export to csv/xlsx **80**

6 CUSTOM SCRIPTS 85

6.1 Custom Script Arguments **85**

6.2 Enabling Custom Scripts **86**

6.3 Custom Scripts Samples **86**
6.3.1 testarguments.wsf 86
6.3.2 archiveinputs.wsf 87
6.3.3 stamp.wsf 87
6.3.4 converttoxml.wsf 88

7 TRIGGER FILES 89

8 KINGFISHER CONFIGURATION FILE (KINGFISHER.CONFIG) 90

9 XML JOB CONDITION FILES 91

10 USING KINGFISHER FROM THE COMMAND LINE 92

10.1 Using Job Definition File **92**

11 **AUTOBAHN DX INTEGRATION** 93

12 KINGFISHER FILES AND DIRECTORIES 93

12.1 Log Files **94**
12.1.1 CSV logs 94

13 SUPPORT 94

14 ACKNOWLEDGEMENTS 95

1 Introduction

Kingfisher provides a set of operations for processing PDF files, this includes splitting, extracting pages, extracting text and renaming of the files based on the content of the files. Kingfisher also provides the ability to set the security, metadata and other attributes of a PDF document.

What makes Kingfisher standout is the ability to perform split, extract and rename operations based on text or barcode found in zones or coordinates in a PDF page. Kingfisher can use Optical Character Recognition (OCR) to extract text data from Image PDF files or pages.

1.1 System Requirements

Supported Operating Systems	Windows 10 (64bit) Windows Server 2012 Windows Server 2016 Windows Server 2019
Supported SharePoint Versions	SharePoint Online (office 365) SharePoint Server 2019 SharePoint Server 2016 SharePoint Server 2013 SharePoint Server 2010
Disk Space	450 Mb
.Net Framework	4.7.2

1.2 Installation

Kingfisher will require installation by a user with Administrator privileges.

The installation MSI will guide you through the process of setting up Kingfisher.

1.2.1 User Permissions

The logged in user for the User Interface (UI) and the service user require Full Control of the following folders under the Kingfisher installation folder:

- Conditions
- License
- Logfiles
- Status
- Temp

In addition, the user(s) will require Full Control of the following files in the Bin folder:

- Kingfisher.Config
- Kingfisher.Service.exe.config

1.2.2 Service User

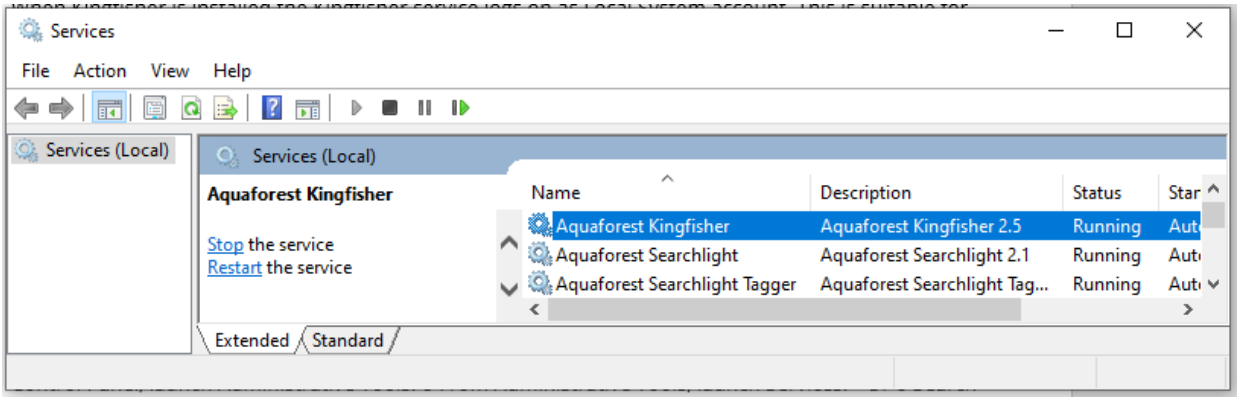
When Kingfisher is installed, the Kingfisher service logs on as the Local System account. This is generally suitable for installations where the files being processed are present on the installation machine, or are stored in SharePoint (online or on-premises).

Windows Server 2019 machines have slightly more restrictive permissions, ensure the service user can access the locations listed in 1.2.1.

If the files are to be referenced via Universal Naming Convention (UNC) paths, an account will need to be created that has the required permissions on those paths.

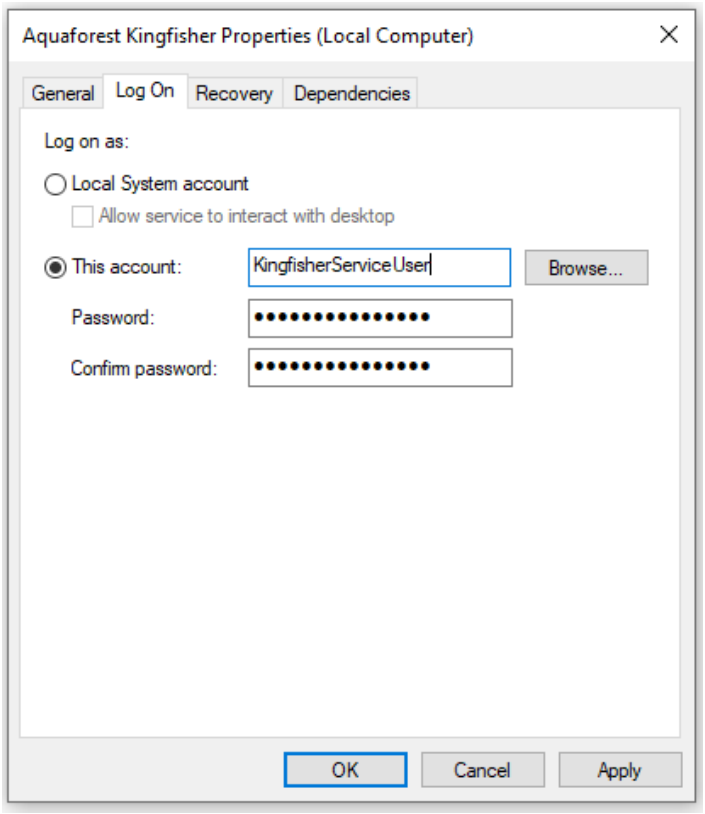
Log-on to the computer as a user with Administrator permissions.

Start the Services app either from the Control Panel or search from the Windows task bar for Services and launch the Services app.



Find the Aquaforest Kingfisher service and right click and select the Properties option.

Select the Log On tab



Enter the username and password and click on OK.

For the service to use the new credentials, it must be restarted. Right click on the Aquaforest Kingfisher service and select Restart.

1.3 Licensing and Purchasing

Please note - from 31 March 2022 Kingfisher is **only** available as a subscription product.

A subscription license will need to be renewed annually and comes with an SMC for the duration of the subscription.

When a license is purchased the relevant subscription license key will be issued for use with the product. This can be entered using the Options tab of the GUI. There is no need to download another version of the software.

The SMC includes support by email, telephone or by remote session, and includes free upgrades for the duration of the SMC.

Existing permanent Kingfisher 2.5 licenses will remain valid (and continue to function as permanent licences).

Additional Support and Maintenance Cover (SMC) for existing permanent licenses will continue to be available for version 2.5.

Trial licenses are fully functional, but all generated PDF files will have the Aquaforest trial stamps on them and text extraction is limited to the first three pages.

1.4 Version Breakdown

	Standard	Professional	Server
Use Cases	*Smart PDF Splitting and Renaming.*	*Smart PDF Page Extraction, Splitting and Renaming.*	*Automated Smart PDF Page Extraction, Splitting and Renaming.*
Split PDF by Content, Barcode or Page Ranges	☑	☑	☑
Rename PDF by Content or Barcode	☑	☑	☑
Extract PDF Pages by Content or Barcode	☑	☑	☑
Includes Optical Recognition for processing Image PDF files	☑	☑	☑
Graphical User Interface	☑	☑	☑
Process Files in SharePoint	☑	☑	☑
Extract PDF Text and Data to CSV, Excel or Text File.		☑	☑
Command Line Interface			☑
Automated Scheduling			☑
Number of Cores Supported	1	1	4

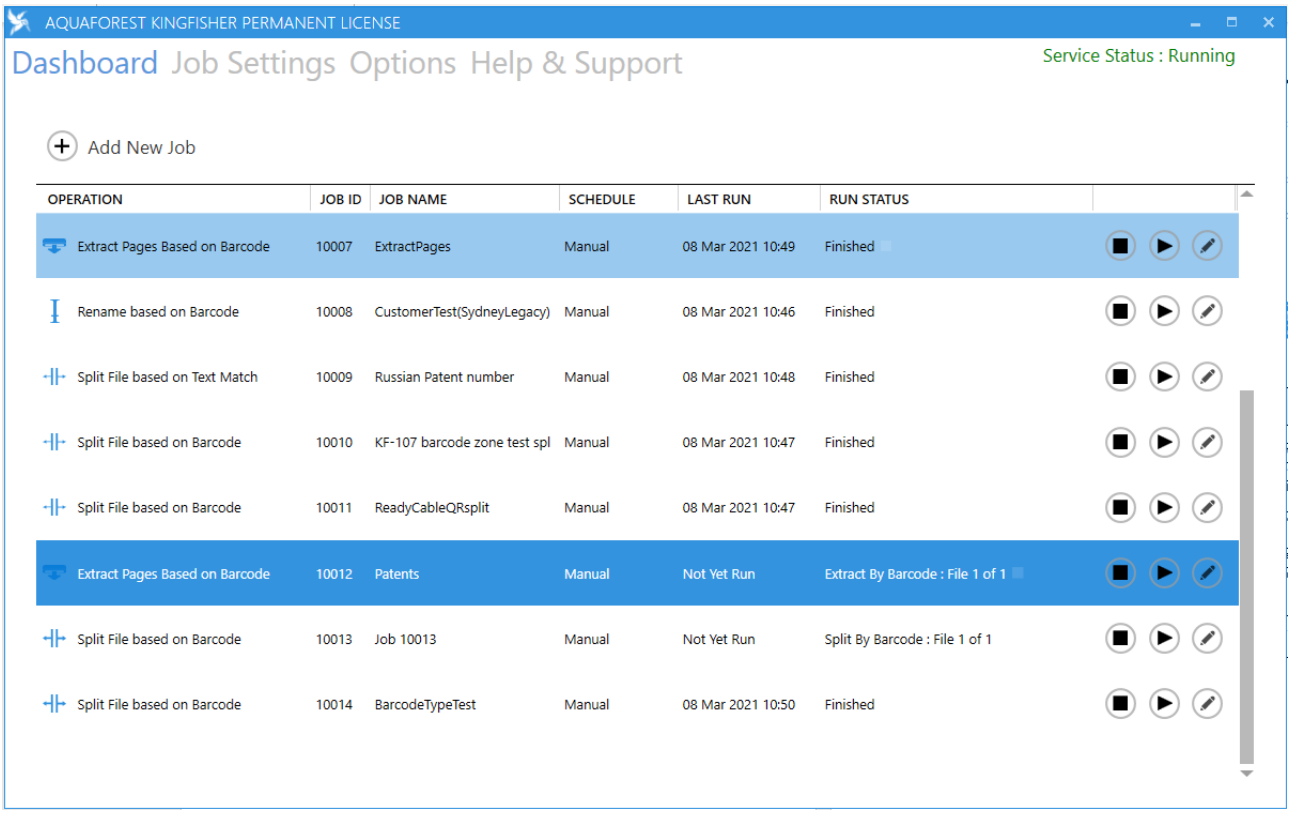
2 Kingfisher User Interface

When Kingfisher is launched, the main window is displayed as shown below. This allows the definition, saving and running of jobs.

This section will explain all the components of the Kingfisher Graphical User Interface (UI).

2.1 Dashboard

The Kingfisher Dashboard contains a list of all the jobs defined in Kingfisher, you can start/stop jobs by clicking the "Play" and "Stop" button by the side of the job name, right clicking on a Job also gives some options as shown on the image above.



Screen Field / Button	Description
Add New Job	This button launches a wizard to create a new Kingfisher job.
Stop Job	Stops an already running job.
Run Job	It is represented by the play button, clicking on it runs the job associated with the button.
Edit Job	This can be accessed by either right clicking the job or clicking the button with the pen icon, use this option to Edit a job, or just double click on the job.
Copy Job	Use the copy job in the context menu to make a duplicate of the currently selected job.
Delete Job	Use the delete job in the context menu to delete the currently selected job.
Clear Error	Use this option to clear errors of jobs in error before running them.

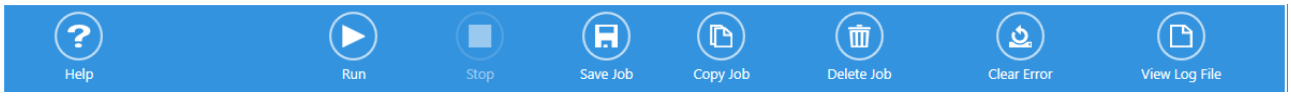
2.2 Job Settings

The Job Settings tab displays the settings for the selected job.

At the top there are sub-tabs that allow you to view and edit the settings for the job.

The sub-tabs and their options are covered in detail below.

At the bottom of the Job Settings tab there is a toolbar with the following options.



Screen Field / Button	Description
View Log File	Click this button to display the log file of the selected Job.
Save	Click this button to save all the changes in the Job.
Run	Click this button to execute the job.
Stop	Click this button to halt the processing of the current batch of files at the earliest convenience.
Delete Job	Click this button to delete the currently selected job.
Copy Job	Click this button to make a copy of the selected job.
Clear Error	Use this option to clear errors of jobs in error before running them.
Explore Destination	Click this button to open the target folder in windows explorer, if the folder does not exist, it will display an error message.

2.2.1 Job Definition

Job Definition contains the common job properties that need to be set for the job to be executed.

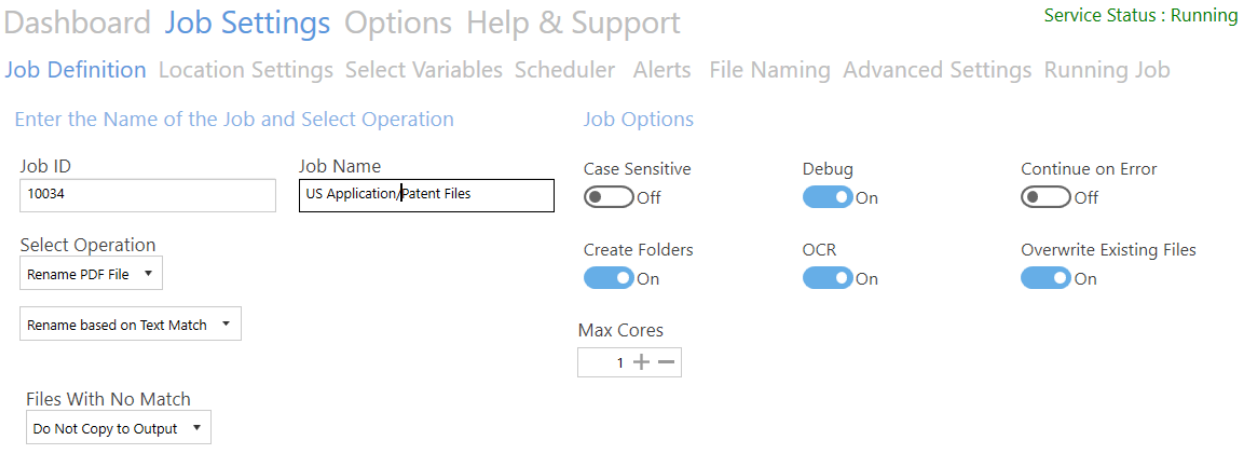
The options displayed depends upon the Operation selected.

The following are common to all operations:

Screen Field / Button	Description
Job ID	Automatically generated Job ID
Job Name	This name is displayed on the Dashboard, and is available as a variable within the job (file names and alerts). Though a default name will be generated when you create a new job, it is always a good practice to give the jobs sensible names when you create them.
Select Operation	Use this box to set the operation you want to carry out, there are four options to choose from. Check section 3 for more details. • Rename PDF File • Split PDF Files • Extract Pages from PDF Files • Extract Content from PDF Files Each option has additional options
Create Folders	If this is checked, Kingfisher will create the target directories if they do not exist.
Overwrite Existing Files	If checked, generated PDF files will overwrite existing files of the same name (but not source files).
Continue on Error	When checked, folder and tree processing jobs will continue when individual files fail.
Debug	If switched on the job log will be more detailed. Additionally, temporary files are not deleted at the end of the run. These can build up if Debug is left on in production
Max Cores	Sets the number of files that should be processed in parallel. Multi-core jobs are only available with the Server license.
Case Sensitive	Setting this to true will make sure all the searches, text operations and pattern matching will be performed in a case sensitive fashion.
OCR	If switched, Kingfisher will use OCR to locate the text on the PDF document, this is good for image PDFs and PDF file with Encoding.

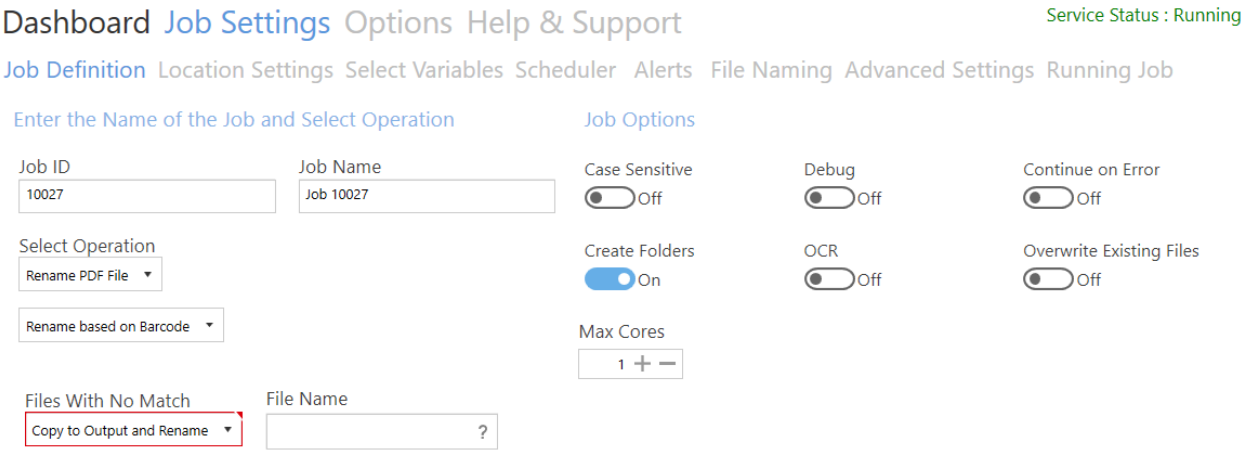
2.2.1.1 Rename PDF File

2.2.1.1.1 Rename based on Text Match



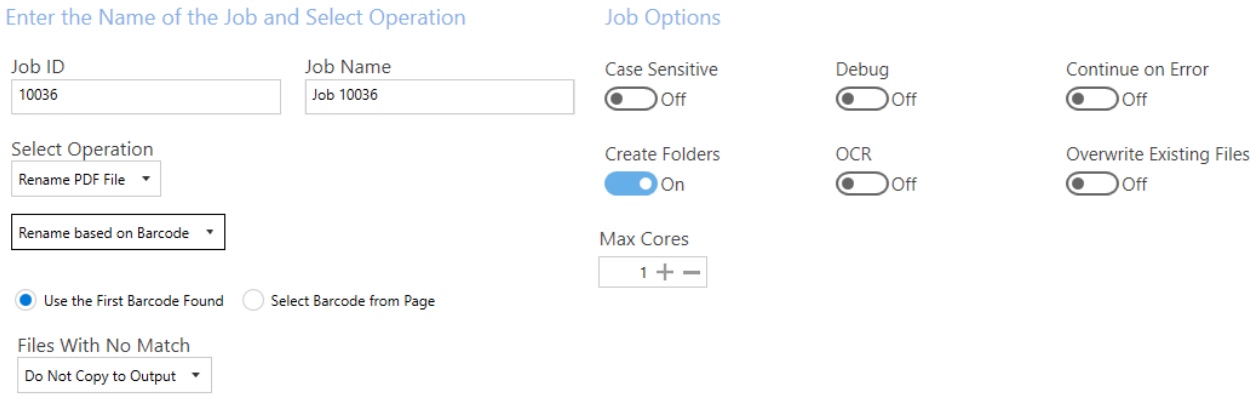
Screen Field / Button	Description
Sub Options	• Rename based on Text Match • Rename based on Barcode • Rename based on Text and Barcode
Files With No Match	Options are: • Do not copy to output • Copy to output • Copy to output and rename (Additional filename template field added to settings)

The “Copy To Output and Rename” requires an additional filename template which can use a number of variables.



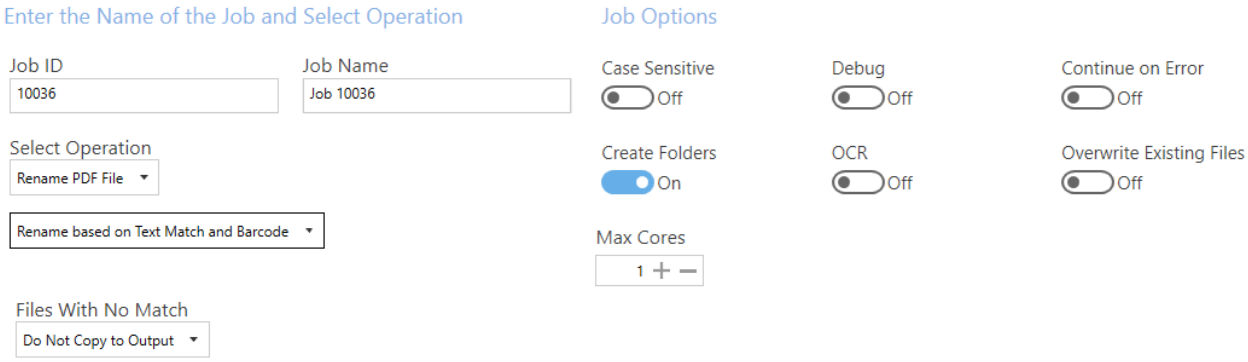
Click on the question mark for details of the available variables.

2.2.1.1.2 Rename based on Barcode



Screen Field / Button	Description
Use the First Barcode found	Use the first barcode located in the document
Select Barcode from Page	Select the barcode in the Select Variable tab
Files With No Match	Options are: • Do not copy to output • Copy to output • Copy to output and rename (Additional filename template field added to settings – see 2.2.1.1.1 for details)

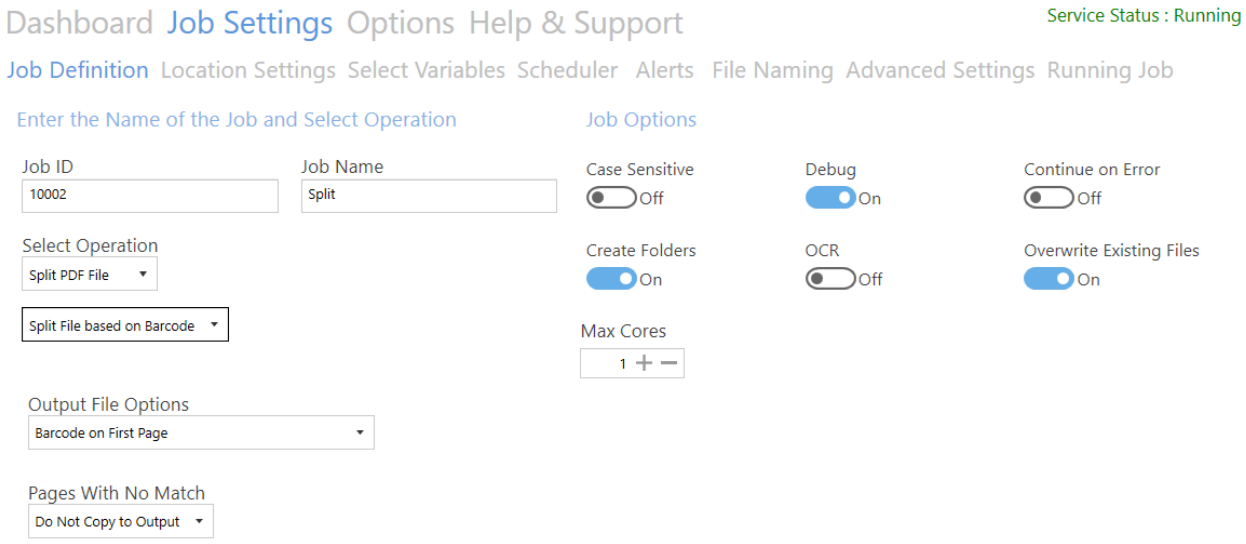
2.2.1.1.3 Rename based on Text Match and Barcode



Screen Field / Button	Description
Files With No Match	Options are: • Do not copy to output • Copy to output • Copy to output and rename (Additional filename template field added to settings – see 2.2.1.1.1 for details)

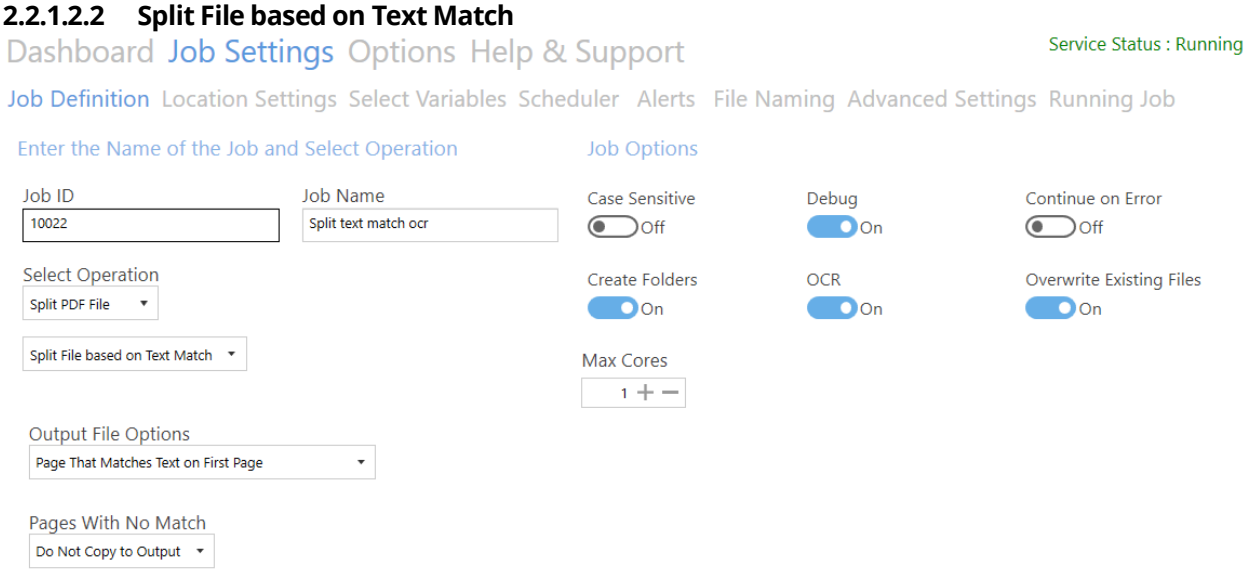
2.2.1.2 Split PDF Files

2.2.1.2.1 Split File based on Barcode



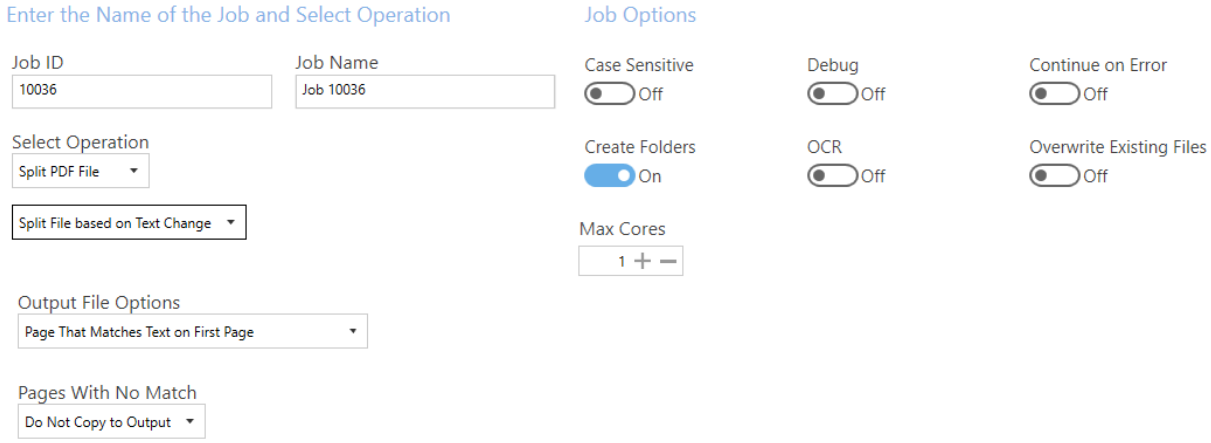
Screen Field / Button	Description
Output File Options	• Barcode on first page • Barcode on last page • Remove Barcode Page
Pages With No Match	Options are: • Do not copy to output • Copy to output • Copy to output and rename (Additional filename template field added to settings – see 2.2.1.1.1 for details)

2.2.1.2.2 Split File based on Text Match



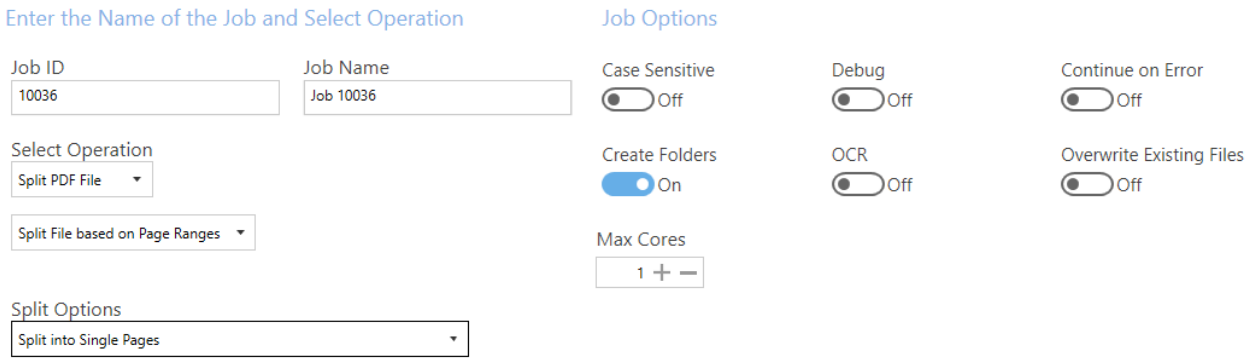
Screen Field / Button	Description
Output File Options	• Barcode on first page • Barcode on last page
Pages With No Match	Options are: • Do not copy to output • Copy to output • Copy to output and rename (Additional filename template field added to settings)

2.2.1.2.3 Split File based on Text Change



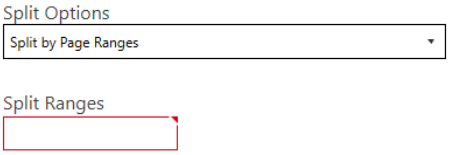
Screen Field / Button	Description
Output File Options	• Page That Matches Text on First Page • Page That Matches Text on Last Page • Remove Page that Matches Text
Pages With No Match	Options are: • Do not copy to output • Copy to output • Copy to output and rename (Additional filename template field added to settings)

2.2.1.2.4 Split File based on Page Ranges



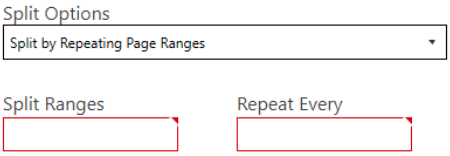
Screen Field / Button	Description
Split Options	• Split into Single Pages • Split by Page Ranges • Split by Repeating Page Ranges

Split by Page Ranges has the one additional field.



Screen Field / Button	Description
Split Ranges	Enter a range

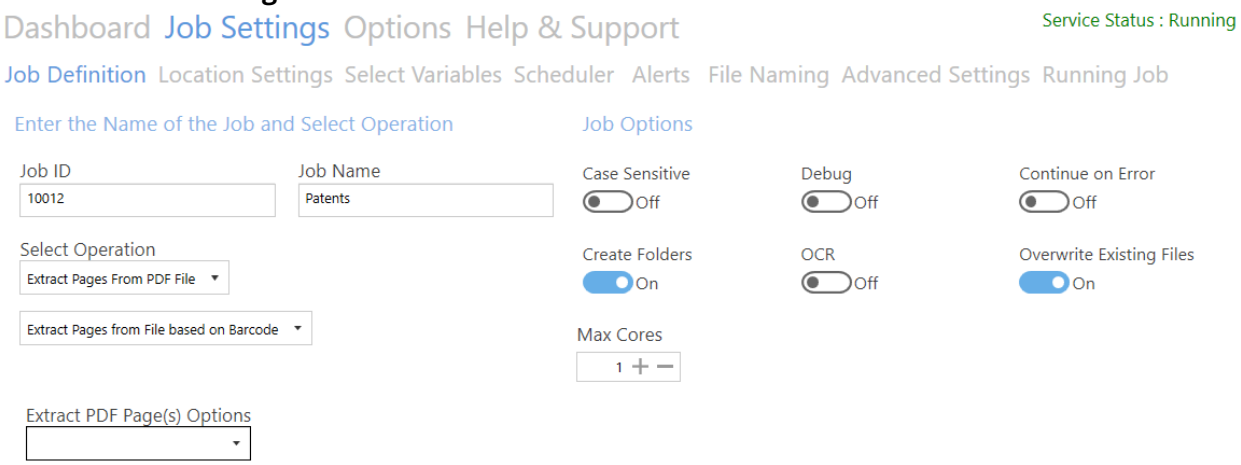
Split by Repeating Page Ranges has two additional fields.



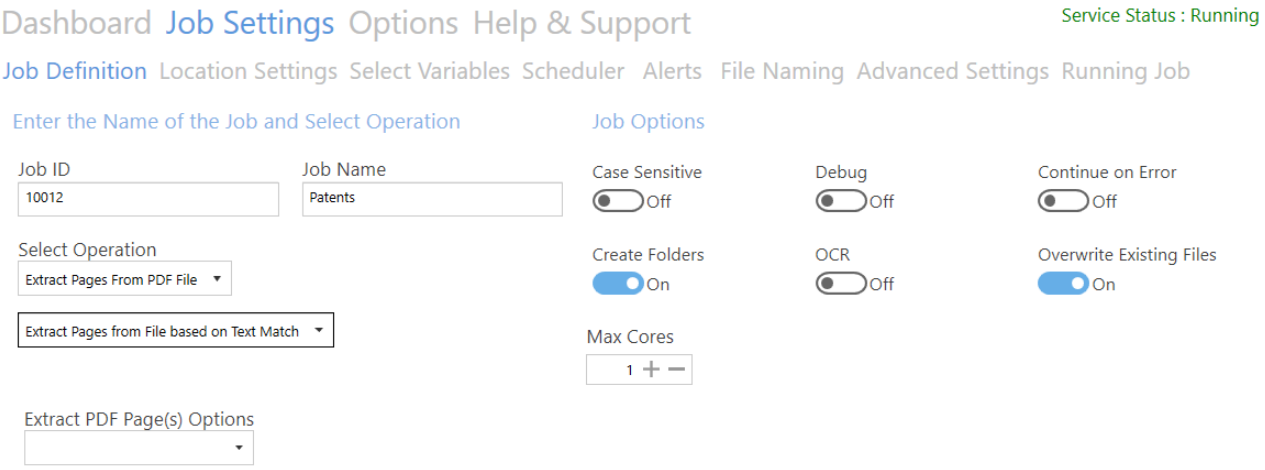
Screen Field / Button	Description
Split Ranges	Enter a range string. See section 3.2.4
Repeat Every	Repeat the range after the number of pages entered

2.2.1.3 Extract Pages from PDF Files

2.2.1.3.1 Extract Pages from file based on Barcode



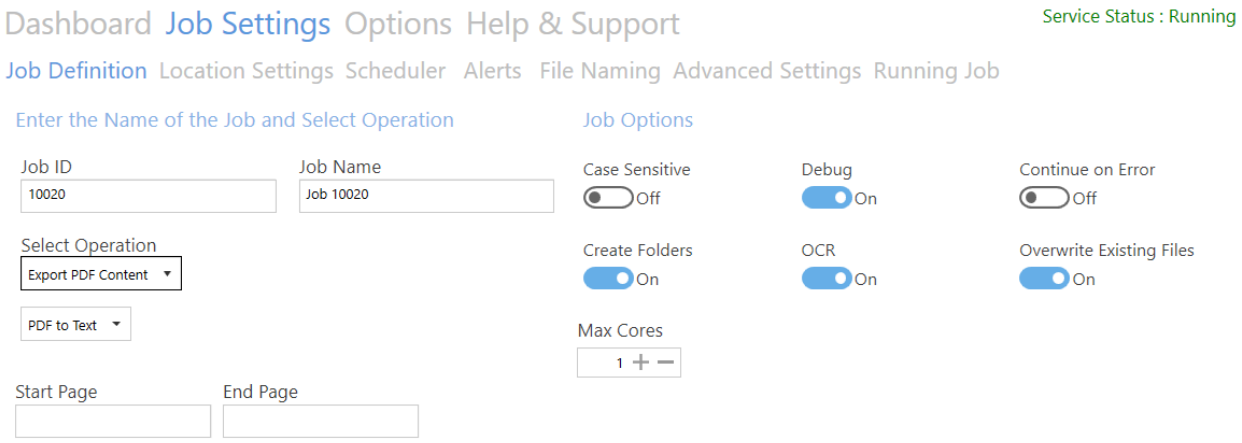
2.2.1.3.2 Extract Pages from file based on Text Match



Screen Field / Button	Description
Extract as Single Pages	Extract each page and save separately. Ensure that each page will be named differently or add the %INDEX% variable.
Merge All extracted pages	Extract each page and merge into a single document
Merge All extracted pages on unique expressions that match	Last page to extract (blank is last page)

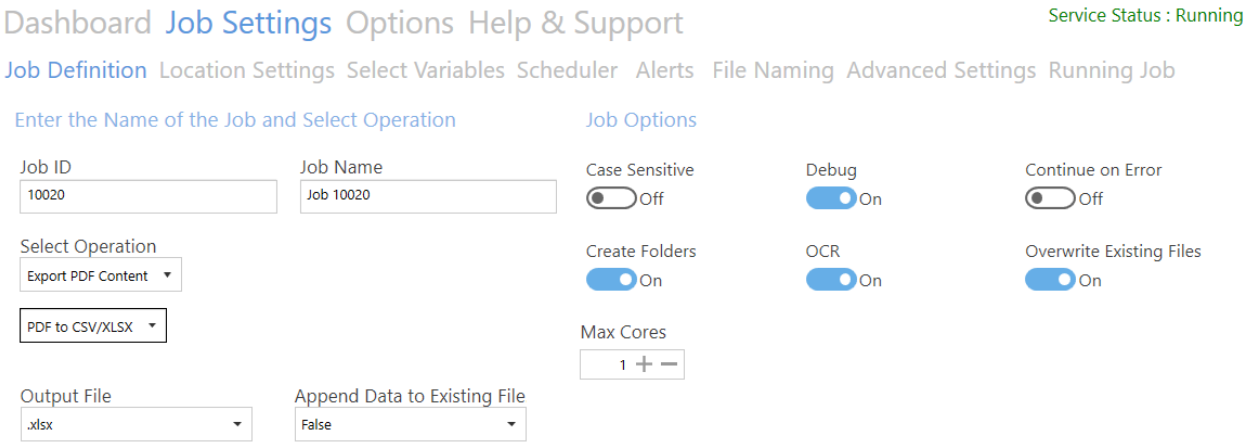
2.2.1.4 Export PDF Content

2.2.1.4.1 PDF to Text



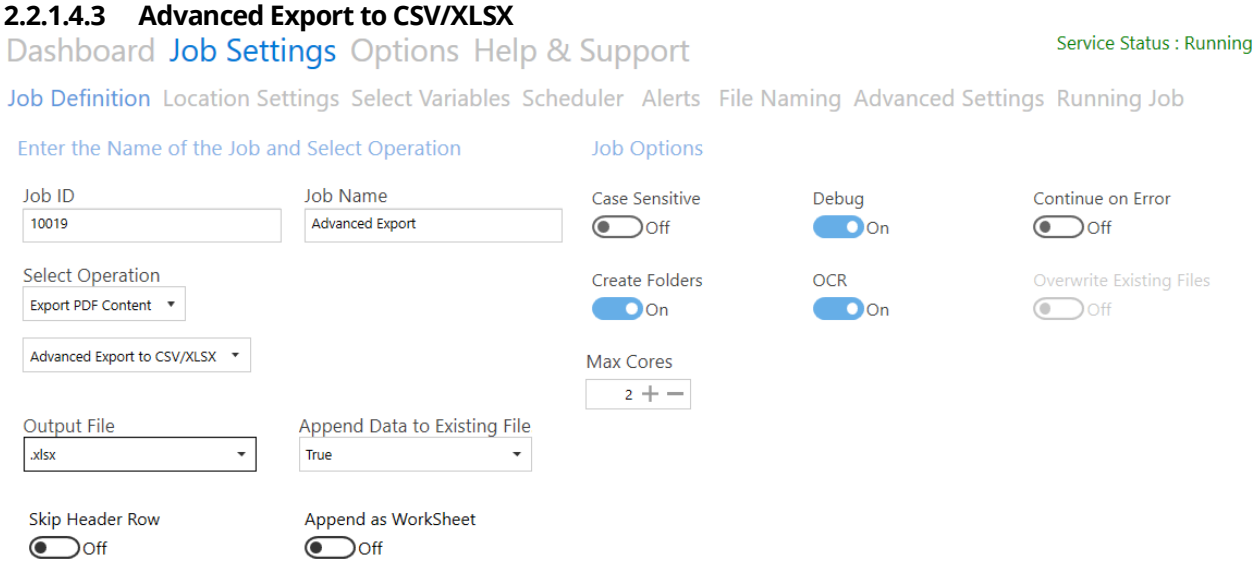
Screen Field / Button	Description
Start Page	First page to extract (blank is first page)
End Page	Last page to extract (blank is last page)

2.2.1.4.2 PDF to CSV/XLSX



Screen Field / Button	Description
Output File	• CSV – Comma Separated Values • XLSX – Excel sheet
Append Data to Existing File	Set to True to append to an existing file, False to overwrite an existing file

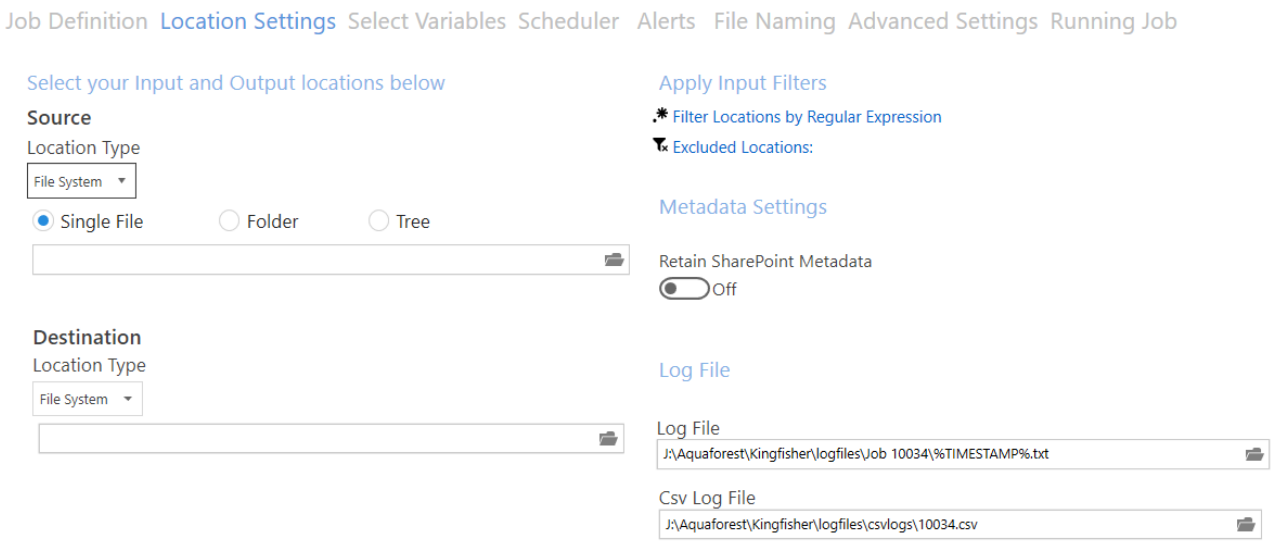
2.2.1.4.3 Advanced Export to CSV/XLSX



Screen Field / Button	Description
Output File	• CSV • XLSX
Append Data to Existing File	• True – add lines to an existing file • False – Create a new file (see Overwrite Existing Files option below)
Skip Header Row	Skip header row in sheet (Append Data to Existing File = True)
Append as WorkSheet	Add new data as a new worksheet in the file (.xlsx output and Append Data to Existing File = True only)

2.2.2 Location Settings

Kingfisher allows users to process files in SharePoint Online (Office 365), SharePoint On-Premises and on the Windows file system. This section will show you the details of the "Location Settings" tab.



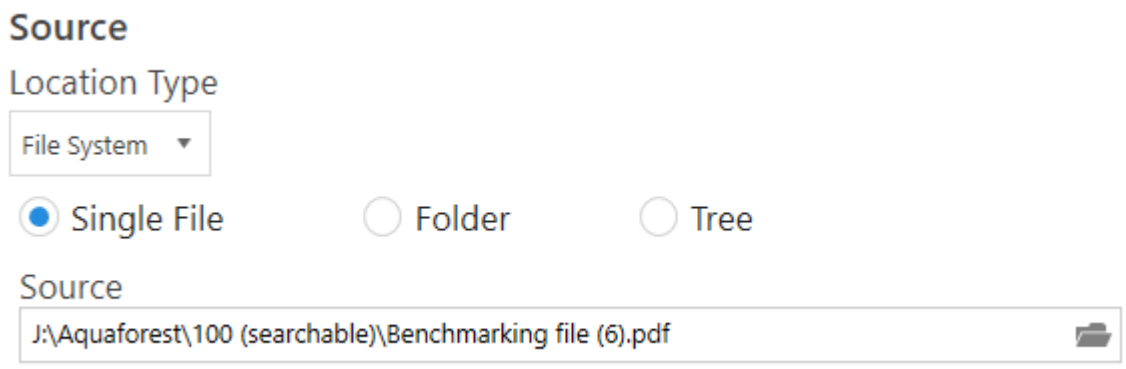
Screen Field / Button	Description
Source	Use this Combo Box to select the Source location type, the options are given below. • File System • SharePoint • Office365
Destination	Use this Combo Box to select the Source location type, the options are given below. • Same as Source: This option puts the file back in the source location • File System • SharePoint • Office365
Filter Locations by Regular Expression	Regular expressions can be used to filter the input files that are processed by Kingfisher.
Exclude Locations	Add paths that you want Kingfisher to skip.
Retain PDF Metadata	Switch this on to retain the PDF metadata of the source PDF file.
Retain SharePoint Metadata	Switch this on to retain the SharePoint metadata of the source PDF file, note that the destination library must have the same column types and names as the source library for this to work.
Log File	The path to the log file for the current rule, you can use the %TIMESTAMP% and %DATESTAMP% variables to create log files based on dates or time.
Csv Log File	The csv log files contain a summary of the new file names generated by Kingfisher. , You can use the %TIMESTAMP% and %DATESTAMP% variables to create log files based on dates or time. You can manually set the See section 12.1.1 for details of the CSV logs contents.

2.2.2.1 Location Types

Kingfisher gives you an option of 3 different location types, this section explains the different ways there are added or edited in a job.

2.2.2.1.1 File System

This is straight forward, choose between processing a “Single File”, “Folder” or “Tree”. After chose the file system type, you can use the browse button to select the location.



The Browse button allows you to browse to the location of the file or folder.

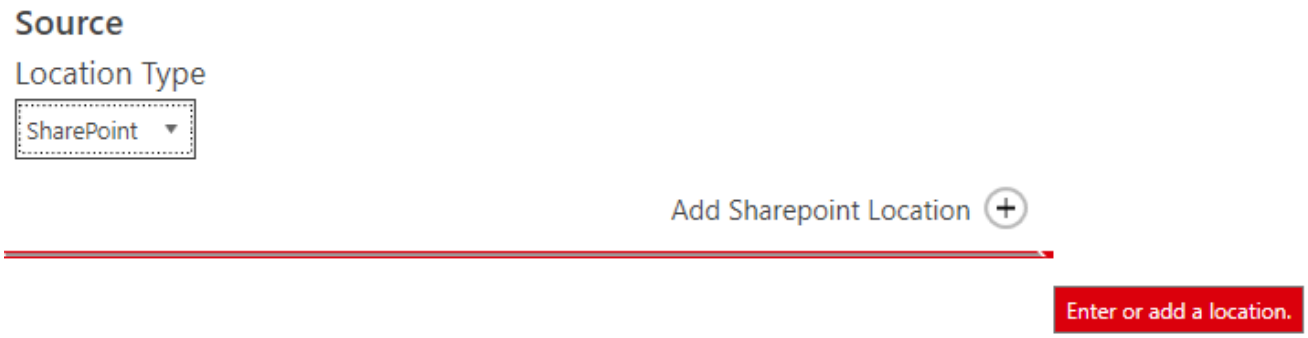
2.2.2.1.1.1 Local, Mapped and UNC paths

Local paths are those that are on drives local to the machine running Kingfisher. These are (subject to suitable permissions for the Kingfisher service user) usable as paths for locations for use by Kingfisher.

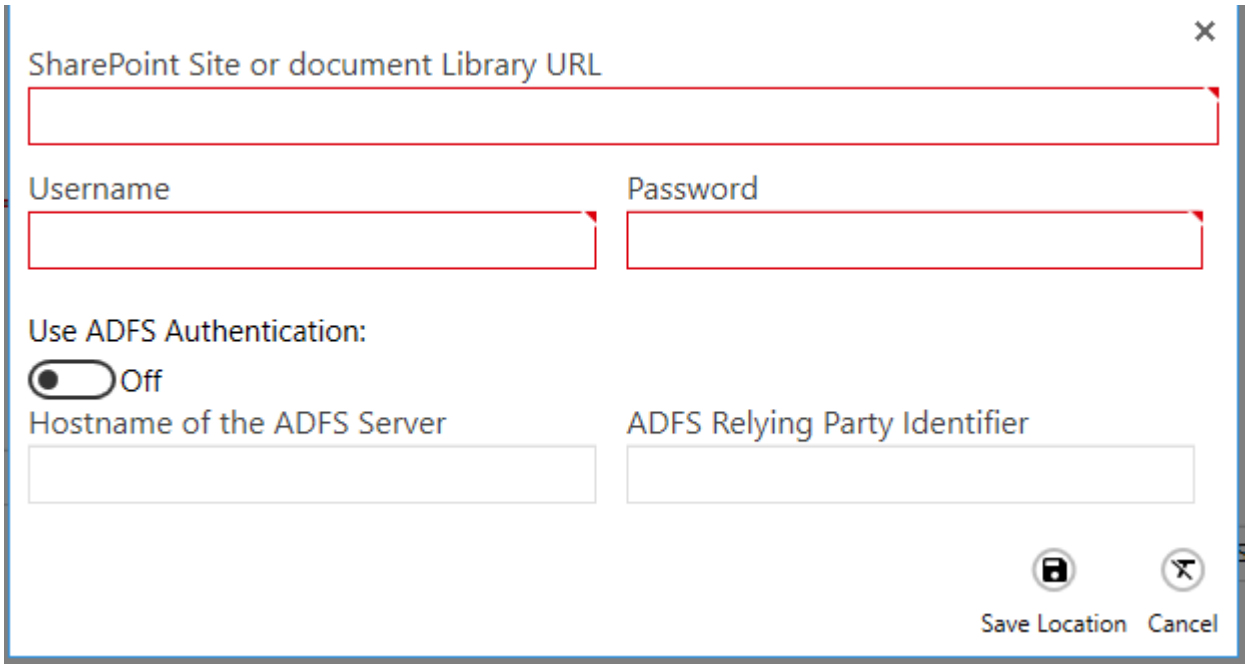
Mapped drives (where a device has a network path mapped to a drive letter) must not be used in paths in Kingfisher. This is due to how drive mapping is applied to service users. If you have a network path, please use the Uniform Naming Convention (UNC) path. If you select a path using the file or folder browser, it will convert it to the UNC path.

2.2.2.1.2 SharePoint (On-Premises)

Adding a SharePoint location is a bit different, after select the "SharePoint" option, you will see the button below.



Click on the "Add SharePoint Location" button to open the SharePoint location dialog box shown below.



The SharePoint locations can be accessed using either Username and Password, or by Active Directory Federation Services (ADFS).

Screen Field / Button	Description
SharePoint URL	Provide the URL of the SharePoint site or document library you want Kingfisher to process.
UserName	Provide the username of a user that has access to the URL you provided above.
Password	Provide a password for the User above.
Use ADFS Authentication	Switch this on if you use Active Directory for your SharePoint User Management.
Hostname of the ADFS Server	Provide the name of the Active Directory server.
ADFS Relying Party Identifier	Provide the Relying Party Trust identifier for your SharePoint.

After you enter your SharePoint details, click the “Save Location” button and the locations will be validated as shown below.

After saving the location, you can always edit by clicking the “Edit SharePoint Location” shown below.

The key thing to be aware of is that Kingfisher will traverse everything under the URL, so if the output location is in the same location, the files will be picked up on subsequent runs.

To specify what folders will and will not be processed, use the Input Filters:

- Filter Locations by Regular Expression
- Excluded Locations

Additionally if the output location is supposed to be a sub folder, ensure that the URl has EXACTLY the same format (Shared%20Documents and Shared Documents are treated as different locations and generates a nested folder structure.

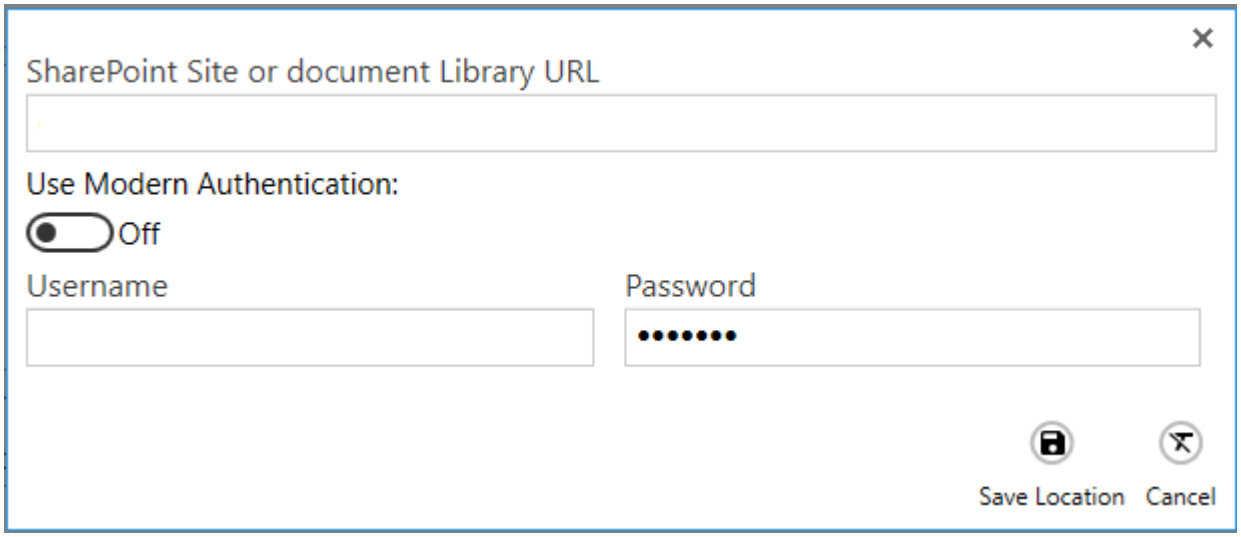
2.2.2.1.3 Office365 (SharePoint OnLine)

Adding a new Office 365 location is similar to adding a new SharePoint location.

There are two options for authentication:

- Username and Password
- Modern Authentication/OAuth2

2.2.2.1.3.1 Username and Password



Screen Field / Button	Description
SharePoint URL	Provide the URL of the SharePoint site or document library you want Kingfisher to process.
UserName	Provide the username of a user that has access to the URL you provided above.
Password	Provide a password for the user above.

The key thing to be aware of is that Kingfisher will traverse everything under the URL, so if the output location is in the same location, the files will be picked up on subsequent runs.

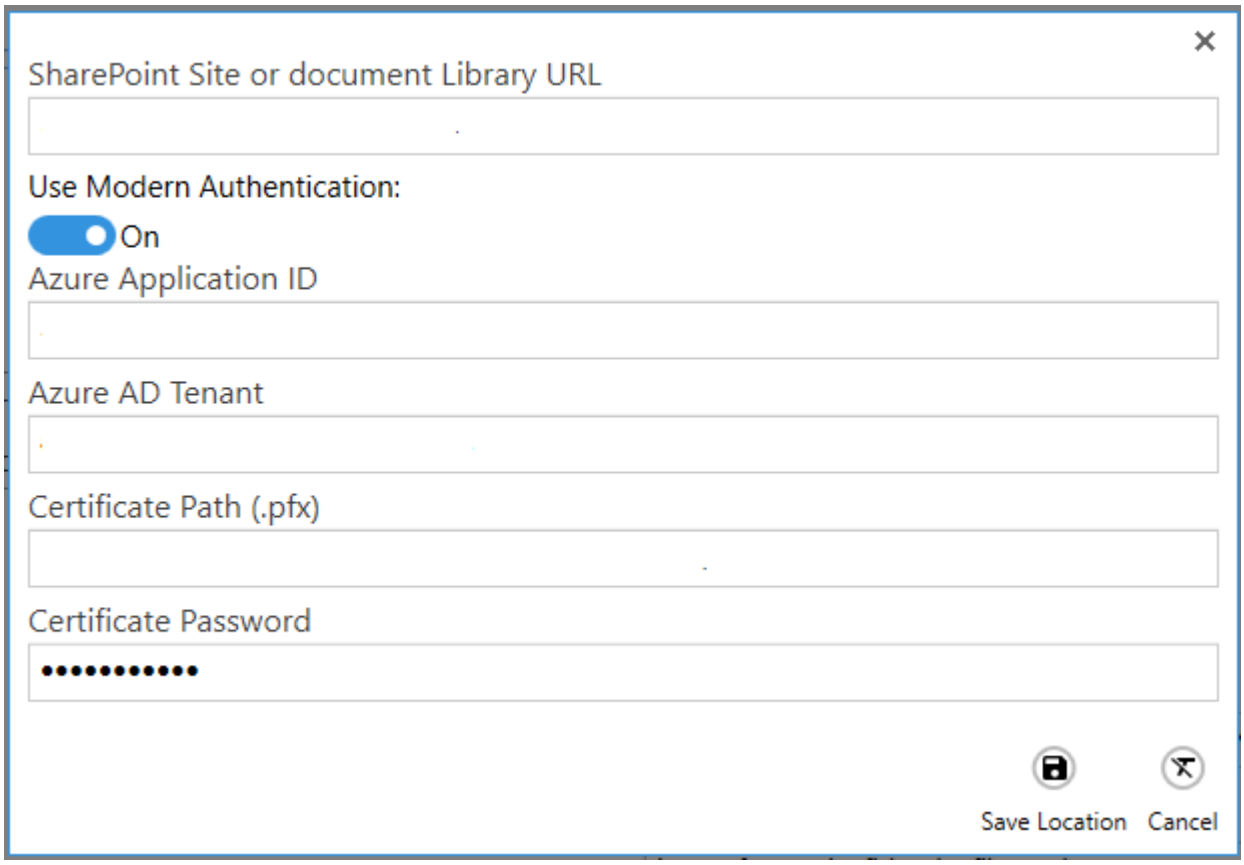
To specify what folders will and will not be processed, use the Input Filters:

- Filter Locations by Regular Expression
- Excluded Locations

Additionally if the output location is supposed to be a sub folder, ensure that the URl has EXACTLY the same format (Shared%20Documents and Shared Documents are treated as different locations and generates a nested folder structure.

2.2.2.1.3.2 Modern Authentication

For details on how to set up modern (OAuth2) authentication for SharePoint (and email) see the OAuth2 Configuration document.



Screen Field / Button	Description
SharePoint URL	Provide the URL of the SharePoint site or document library you want Kingfisher to process.
Azure Application ID	Azure Application (Client) ID
Azure AD Tenant ID	Azure Active Directory Tenant ID
Certificate Path	Path to the .pfx certificate associated with the Client ID
Certificate Password	Password for the .pfx certificate

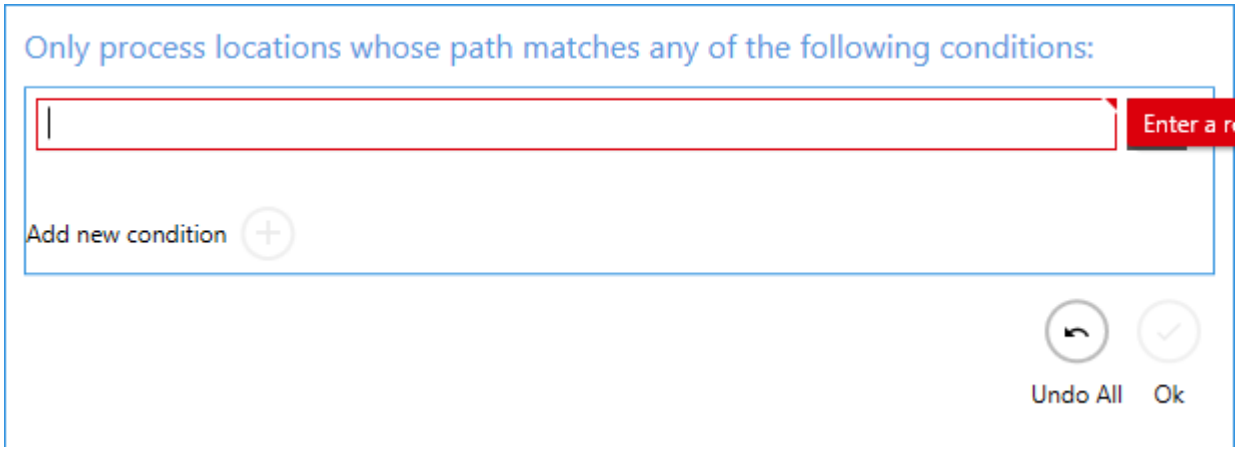
Follow the same process as the previous section if you want to save or edit the location.

2.2.2.2 Apply Input Filters

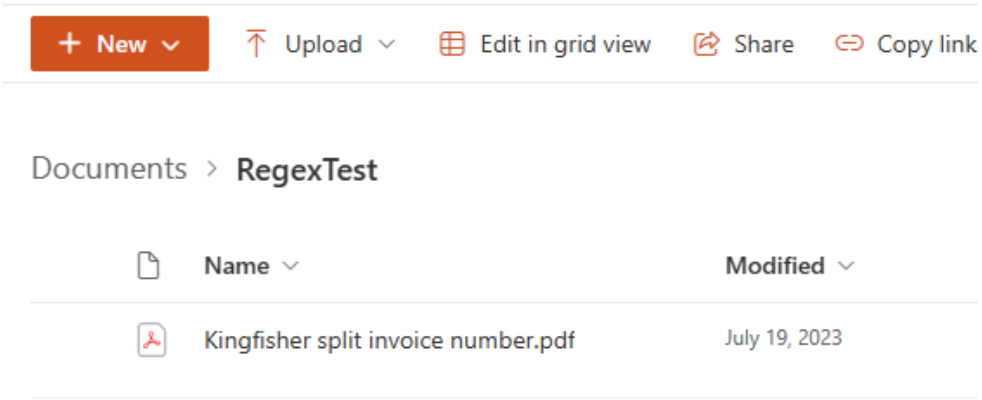
Kingfisher give you the ability to filter or exclude source files/paths based on difference criteria.

2.2.2.2.1 Filter Locations by Regular Expression

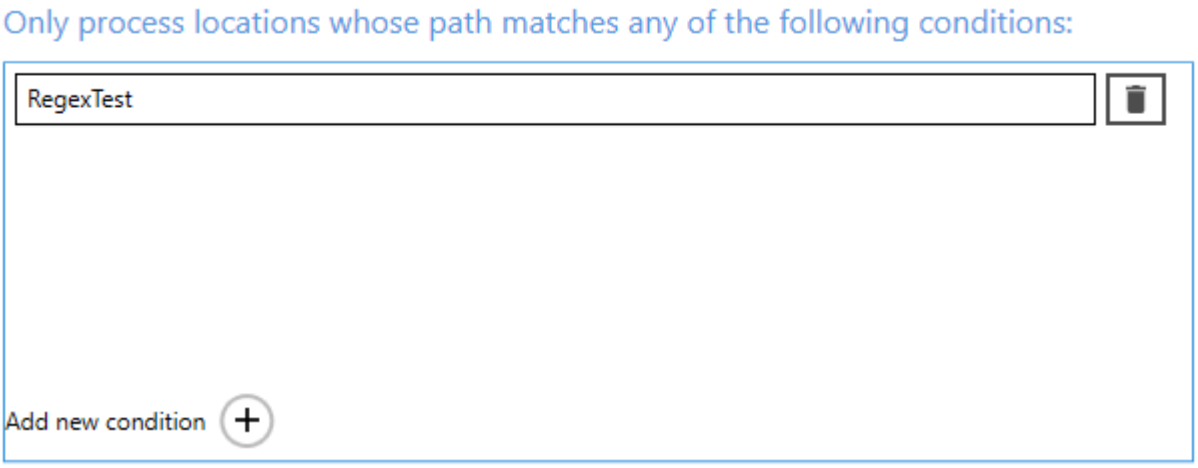
If you provide any values for this, Kingfisher will only process files that match the regular expression provided. To add values, click the "Filter Locations by Regular Expression" link, this will show the dialog below. Then click the "Add new Location" button.



Example:
This example job will process files only in the RegexTest folder:



Open the Filter Locations by Regular Expressions dialogue and add the folder name RegexTest.

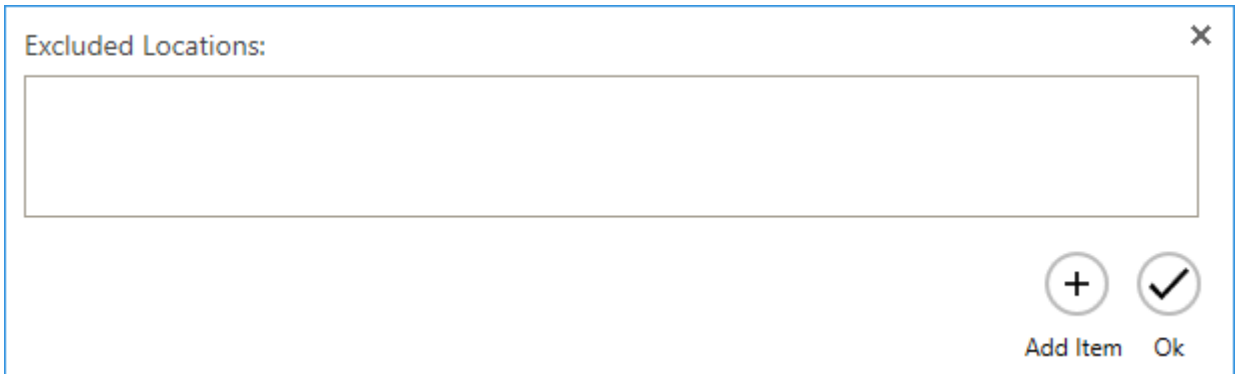


Be aware the regular expression only needs to match **something** in the path.

If the job is in Debug mode, the log will contain a record of those files that match and those that do not. This will help you with tailoring the regular expression to your environment.

2.2.2.2.2 Exclude Locations

Any path you provide here will be skipped by Kingfisher, to add values to this property, click the "Exclude Location" link. The Dialog below will be shown.



Click on the "Add Item" button to add a new value, after you are done click the Ok button.

If the job is in Debug mode, the log will contain a list of the exclusion locations and will show locations that are included and those that are excluded.

This will help you with tailoring the regular expression to your environment.

2.2.3 Select Variables

One of the most attractive features of Kingfisher is the ability to select text or barcode at a position in a PDF page, the selected text or barcode can be used to:

- Decide whether to split the document at the page where the text of barcode was found.
- Decide whether to extract the page where the text of barcode was found.
- Use these extracted text or barcode values to rename the file.

2.2.3.1 Template file

The Select Variables tab uses a template file to allow the interactive definition of the various zones. By default, it picks up either the source file selected in the Location Settings tab or one of the files located in the folder file selected in the Location Settings tab.

It is possible that during the use of Kingfisher, this file will be removed. If the job condition is opened again in the UI, a warning message will be displayed indicating that the template file is missing. It does not affect the operation of Kingfisher.

It is a good idea to define a Template File folder and copy your chosen template file into it. Use the browse file option to select this as the new template file.

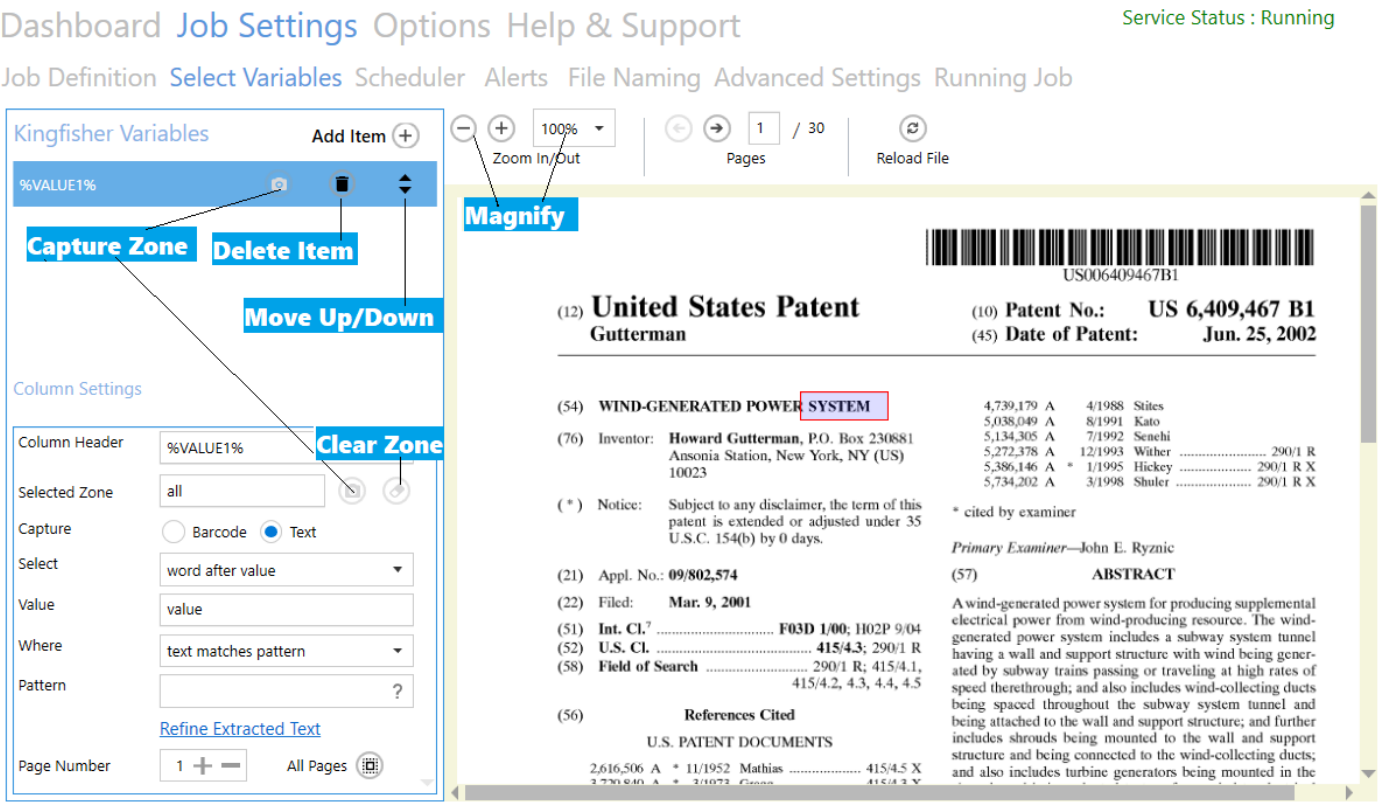
2.2.3.2 Variables

The "Select Variable" tab is used to select four variables named:

- "%VALUE1%": This is used to decide the split or extract page, it is also used to generate the output file name, note this must be selected for some jobs to run.
- "%VALUE2%": This is used to generate the output file name.
- "%VALUE3%": This is used to generate the output file name.
- "%VALUE4%": This is used to generate the output file name.

2.2.3.3 Zones

To assign a zone to a variable, you should first choose the zone on the PDF page with your mouse, click the camera icon associated to the variable you want to assign the zone to. If you do not assign a zone to the variable, Kingfisher will assume you want to use the whole page as a zone. You will notice the "Selected Zone" Text Box change from "all" to the coordinates you selected.

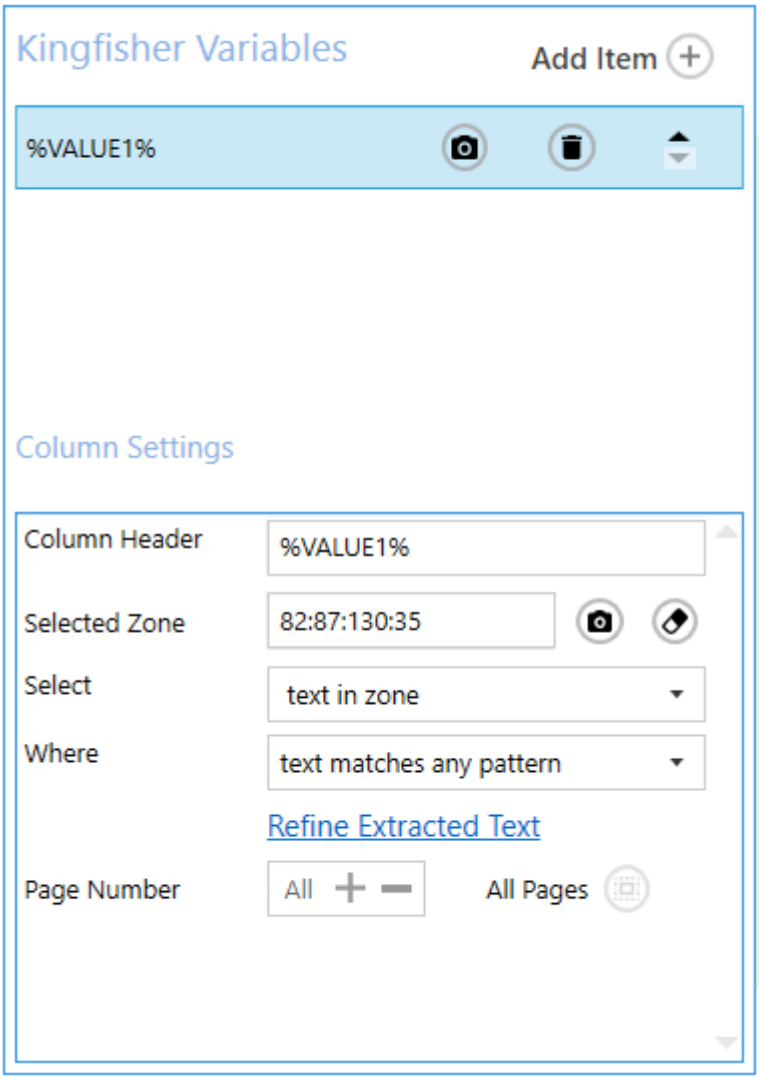


Screen Field / Button	Description

Magnifier	Use the +/- buttons to zoom in or zoom out
Next	Click to go to the next page of the PDF file
Back	Click to go to the previous page of the PDF file
Delete Item	Click this button to delete the variable or item.
Move Up/Down	You can use this to rearrange the order of the variables.
Selected Zone	Click the camera icon to assign a selected zone to a variable
Clear Zone	Click this button set the selected zone to the whole page
Text	Check this radio button to extract text values
Barcode	Check this radio button to extract barcode values
Select	Use this drop-down box to specify the text you want to extract relative to a word/value. See section 2.2.2.1 for more details.
Where	Use this Drop-down box to specify if a regular expression should be used to match the selected text. See section 2.2.2.3 for more details.
Refine Extracted Text	Click this link to refine the text extracted. See section 2.2.2.4 for more details
Page Number	You can specify specific pages to want Kingfisher to check for certain values.
All Pages	Click this button to make Kingfisher search for the values on all pages.

2.2.3.4 Text Match Option

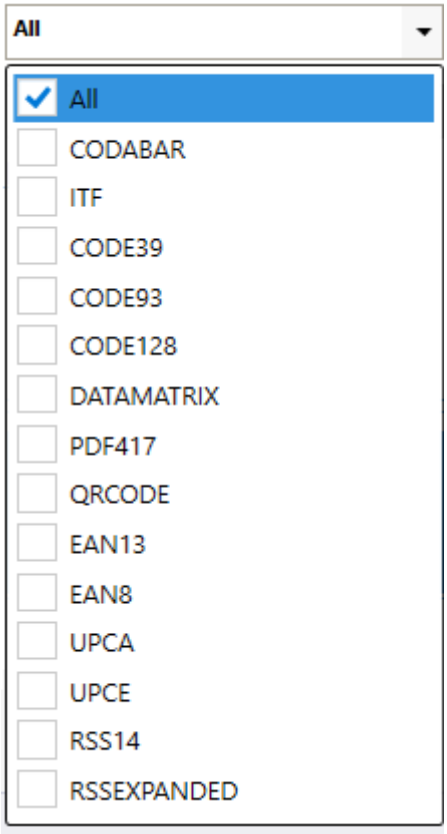
When extracting text, Kingfisher gives you the ability to select text that is relative to a word/pattern. The table below shows the options available and what they mean.



Screen Field / Button	Description
Select	• **text in zone:** This option will select all the text that was extracted. • **word after value**: If this option is selected, Kingfisher will return the word that appears immediately after the word supplied in the text box. • **Word Before value:** If this option is selected, Kingfisher will return the word that appears immediately before the word supplied in the text box. • **all text In Line after value:** If this option is selected, Kingfisher will return all the words that appear on the same line after the word supplied in the text box. • **all text In Line before value:** If this option is selected, Kingfisher will return all the words that appear on the same line before the word supplied in the text box. • **all text In Zone after value:** If this option is selected, Kingfisher will return all the words that appear in the selected zone after the word supplied in the text box. • **all text In Zone before value:** If this option is selected, Kingfisher will return all the words that appear in the selected zone before the word supplied in the text box. • **all text in paragraph before value:** If this option is selected, Kingfisher will return all the words that appear in the paragraph before the word supplied in the text box. • **all text in paragraph after value:** If this option is selected, Kingfisher will return all the words that appear in the paragraph after the word supplied in the text box.
Pattern	• **Where text matches any pattern:** This option will not match the text to any pattern • **Where text matches pattern:** Provide a Regular expression here to match the extracted text to, this will be explained more below.

2.2.3.5 Barcode

When extracting Barcode, you can select specific barcode formats or select All of the supported barcodes (default).



Screen Field / Button	Description
Barcode Type	This combo box allows the user to select the barcode type to extract, below is the list of barcode types that Kingfisher can recognize. • CODABAR 1D format. • ITF (Interleaved Two of Five) 1D format (Code 25). • Code 39 1D format. • Code 93 1D format. • Code 128 1D format. • Data Matrix 2D barcode format. • PDF417 format. • QR Code 2D barcode format. • EAN-13 1D format. • EAN-8 1D format. • UPC-A 1D format. • UPC-E 1D format. • RSS 14 • RSS EXPANDED

2.2.3.6 Text Pattern Option

Kingfisher allows the user to define the text pattern of the text to be extracted from the PDF document. This helps the user make sure only text that match a certain pattern will be extracted. For example, a user might want a phone number, but a postcode gets extracted just because it was present in the same location as the postcode on a different page. Kingfisher provide pattern matching so that a user can make sure the text that is extracted is properly validated before it is accepted.

Patterns are used to match parts of the document. Some useful examples are below. For a complete reference on patterns including advanced examples see the link below.

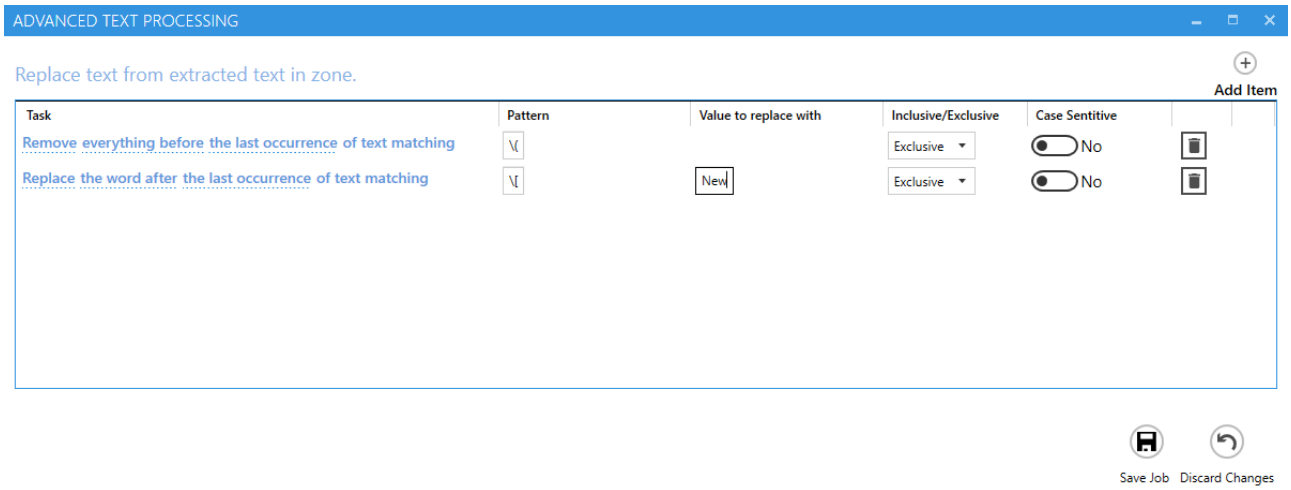
https://docs.microsoft.com/en-us/dotnet/standard/base-types/regular-expression-language-quick-reference

Pattern	Matches	Examples
invoice	The word "invoice"	
invoice no:	The phrase "invoice no:"	
.	Any 1 character	A 9 %
....	Any 5 characters	12abc $5.00
Invoice....	The word invoice followed by any 4 characters. A space is a character.	Invoice 123 Invoice No: Invoice A9-
^invoice	The word invoice at the start of a line	
invoice$	The word invoice at the end of a line	
[0-9] [0-9] [0-9]	Any 3 digits	123 900
[a-z] [a-z] [a-z]	Any 3 alphabetic characters	abc Inv

2.2.3.7 Refine Extracted Text

Kingfisher has an advanced method to refine the extracted text.

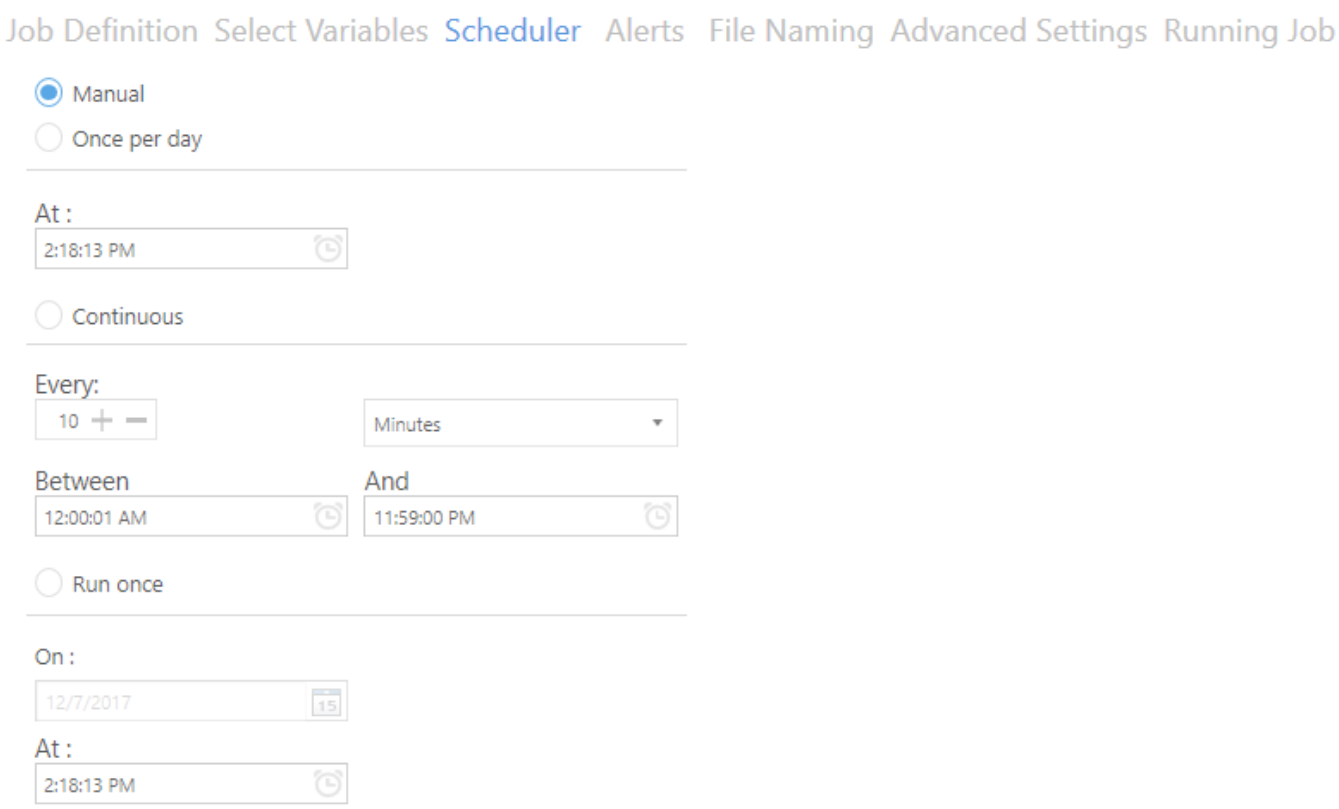
This allows the removing/replacing of certain pieces of text based on patterns or values.



Option	Explanation
Task	The first part of the task is whether to remove or replace the pattern provided, the next part deals with the position and the frequency of the removal or replacement.
Pattern	The pattern to search for, this can be a regular expression, Use the "\" literal to escape special characters. https://docs.microsoft.com/en-us/dotnet/standard/base-types/regular-expression-language-quick-reference
Values to replace with	If you choose to replace a pattern, you will have to provide a substitute value in this text box.
Inclusive/Exclusive	When you chose Inclusive, Kingfisher will remove all the text specified including the pattern being matched. When you choose Exclusive, Kingfisher will remove all the text specified without the pattern being matched.
Case Sensitive	This option tells Kingfisher whether to use a case sensitive approach or not.

2.2.4 Scheduler

This tab allows the user to schedule when jobs are executed.



Setting	Description
Manual	This means that the document library has to be run manually by clicking on the “Run” button on the dashboard.
Once per day	This allows the document library to be scheduled to run at a specified time each day.
Continuous	This allows the document library to be scheduled to run periodically between a start time and end time each day. The periods may be minutes, hours, days or months. For example, a document library may be specified to run every 1 hour between 9:00 and 17:00.
Run Once	This allows the document library to be scheduled to run only once at a specified time.

Please note that Kingfisher’s default behavior is that a Scheduled job that generates an error will not run again until either the error is cleared (right click on the job entry on the Dashboard and select clear error), or the job is successfully run manually.

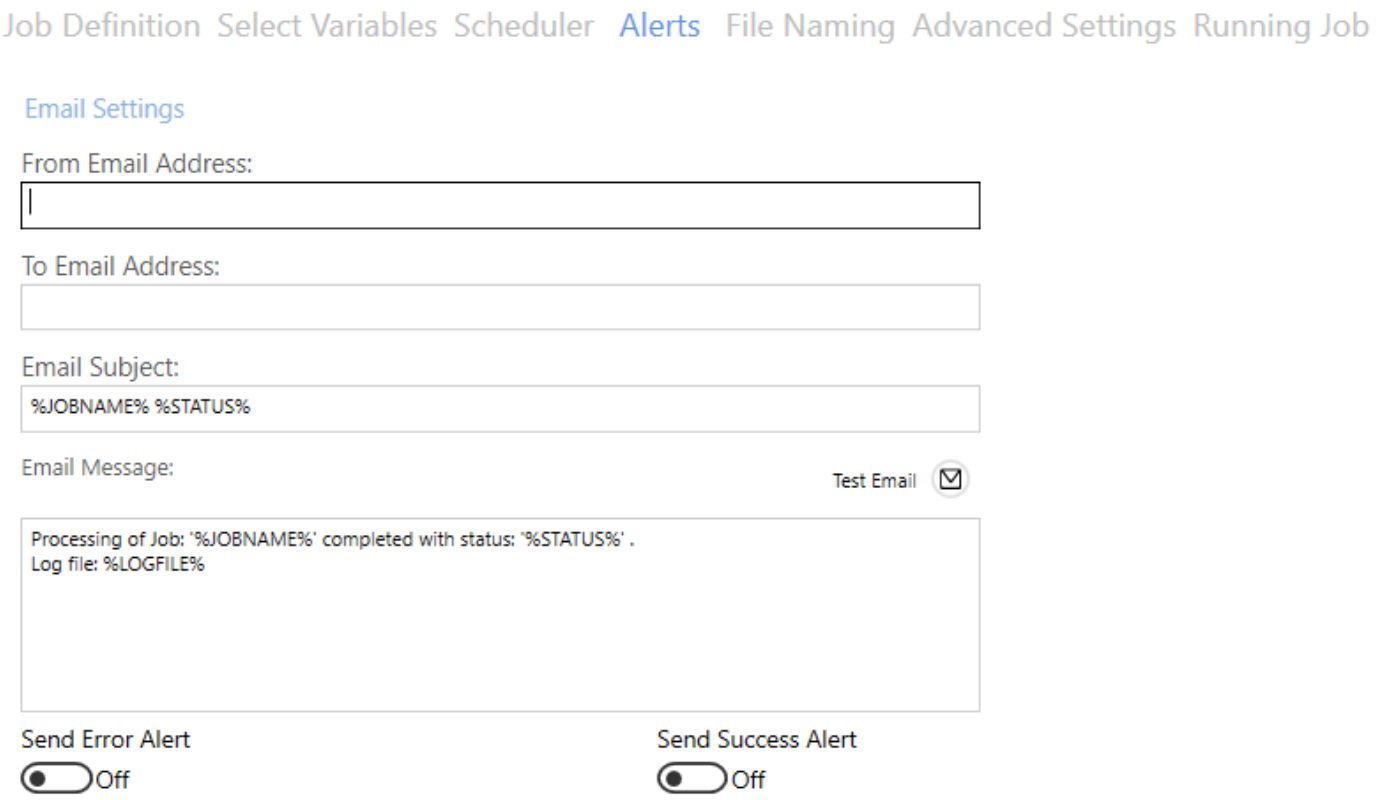
This behavior can be changed by changing two entries in the Kingfisher.Config file.

For jobs scheduled to run once per day, change the value of the setting ScheduleOncePerDayContinueAfterError to true.

For jobs scheduled to run continuously, change the value of the setting ScheduleContinuousContinueAfterError to true.

2.2.5 Alerts

Use this tab to set email alerts for jobs. Email settings (SMTP or OAuth2) are set in the Options/Email tabs (see section 2.3.2).



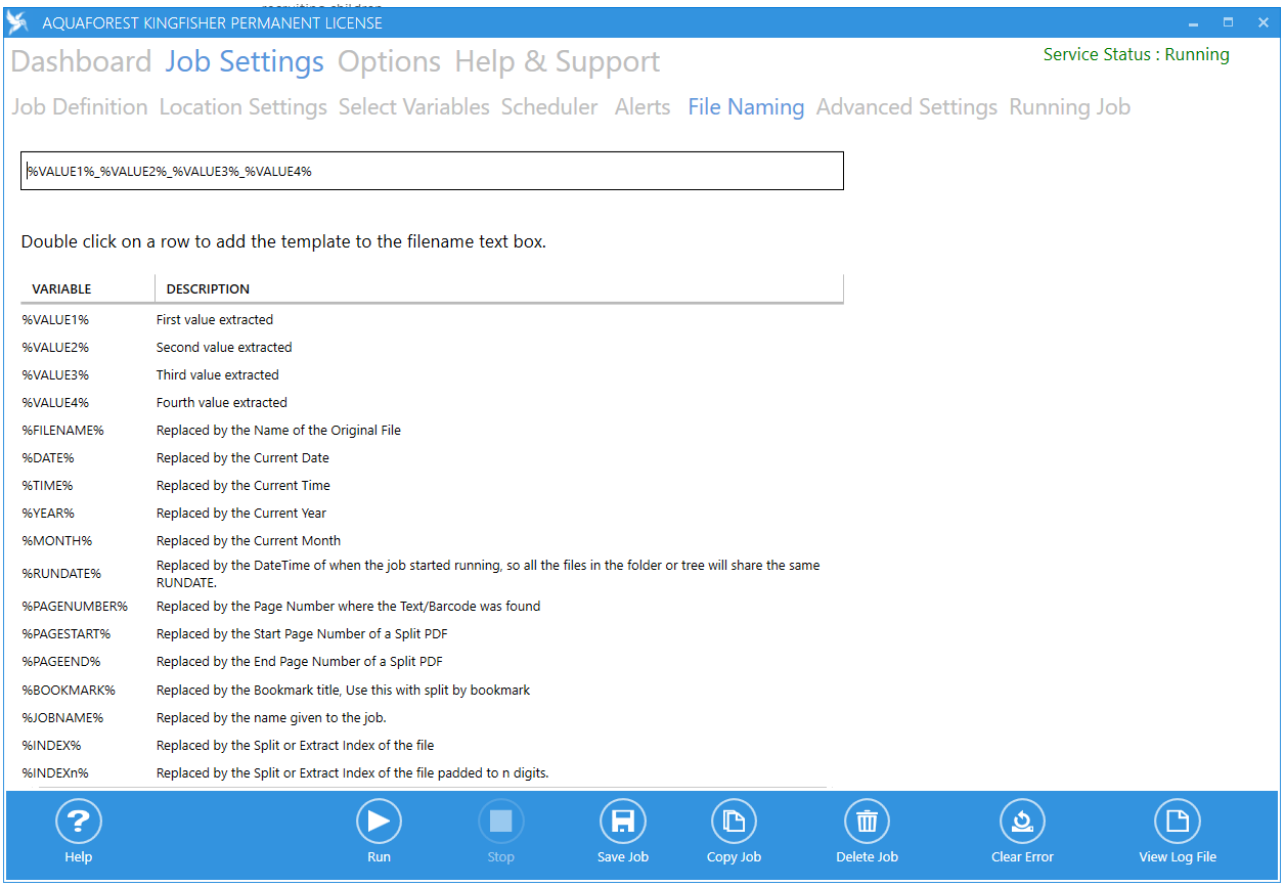
Setting	Description
From Email Address	The email address to send the email from. This needs to be a valid email address for the email service employed (unless Anonymous is chosen as an SMTP option)
To Email Address	The email address to receive the email.
Email Subject	The email subject. You can use the following templates: %JOBNAME% - will be replaced by the name of the library %STATUS% - will be replaced by “success” or “error” depending on whether the job ran successfully or not
Email Message	The email message to send. You can use the following templates: %JOBNAME% - will be replaced by the name of the library %STATUS% - will be replaced by “success” or “error” depending on whether the job ran successfully or not %LOGFILE% - will be replaced by the path of the log file.
Send Error Alert	Switch this on to send an alert when a job ends in an error.
Send Success Alert	Switch this on to send alerts when the job successfully completes.

2.2.6 File Naming

Kingfisher offers various variables to customize the final output name. These variables include:

- Text/Barcode values extracted from the PDF document (including by OCR)
- Index value for split or extract operations (with zero padded options).
- Page numbers (original, first original page of block, last original page of block)
- Bookmarks (for split by bookmark)
- Time and Date values (including the date and time that a job started)

Note that the index value is very useful where the values extracted on different pages during the split/extract are the same.



2.2.6.1 Name Template

You can enter a string in the name template text box, any occurrence of variables in the table below will be replaced by the appropriate value at runtime.

%VALUE1%	The primary barcode/text value.
%VALUE2% %VALUE3% %VALUE4%	The barcode/text value(s) found.
%INDEX%	The index of the output file in a split or extract page operation.
%INDEXn%	The index of the output file in a split or extract page operation padded with zeroes to length n.
%PAGENUMBER%	The page number of the extracted page in the original PDF file.
%PAGESTART%	Start page number.
%PAGEEND%	End page number.
%FILENAME%	The original file name without the .pdf extension.
%BOOKMARK%	This represents the bookmark name and can be used with the "Split by Top Level Bookmarks" split option only.
%TIME%	The current Time.
%DATE%	The current Date.
%RUNDATE%	The Time at which the job was executed, the difference between this and the %TIME% variable is that this value will be same for all files processed in a single job run.
%YEAR%	The current Year in the form yyyy
%MONTH%	Current Mo0nth in form MM.
%JOBNAME%	The Name of the Job.

For non-Tree file sources, there is an option to include folders in the file naming by using a double backslash (escaping a single backslash).

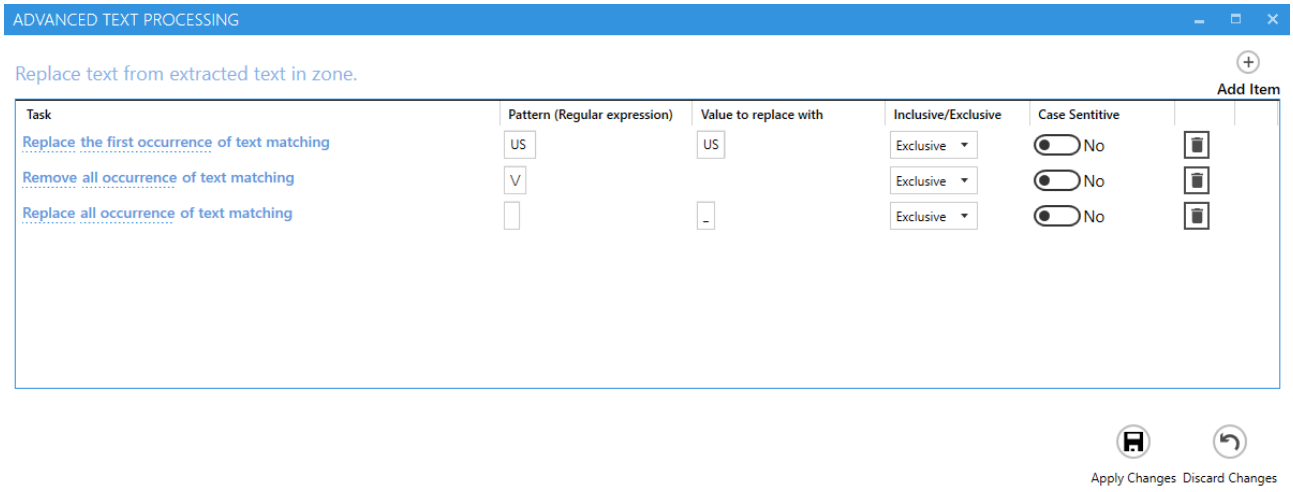
%DATE%\\%VALUE1%

Be aware it is possible to use this file name template with Tree file sources, but the folder structure will include the source folder structure before the file naming one.

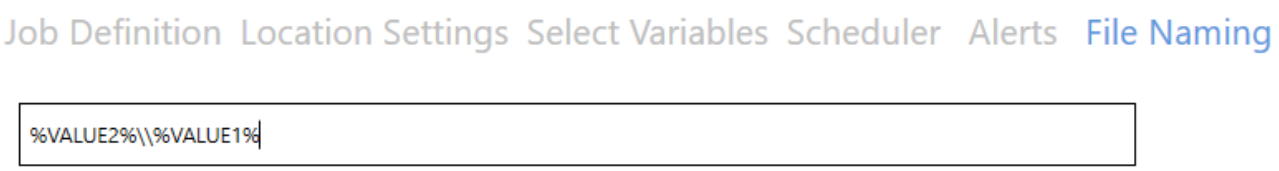
2.2.6.1.1 File naming example

Two areas are selected:

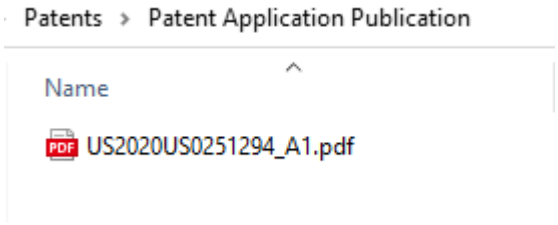
%VALUE1% has text refinement added:



The File Naming tab has the value %VALUE2\\%VALUE1. The double backslash ("\\") is an escaped backslash that indicates the file should be written to a folder with the value of %VALUE2%

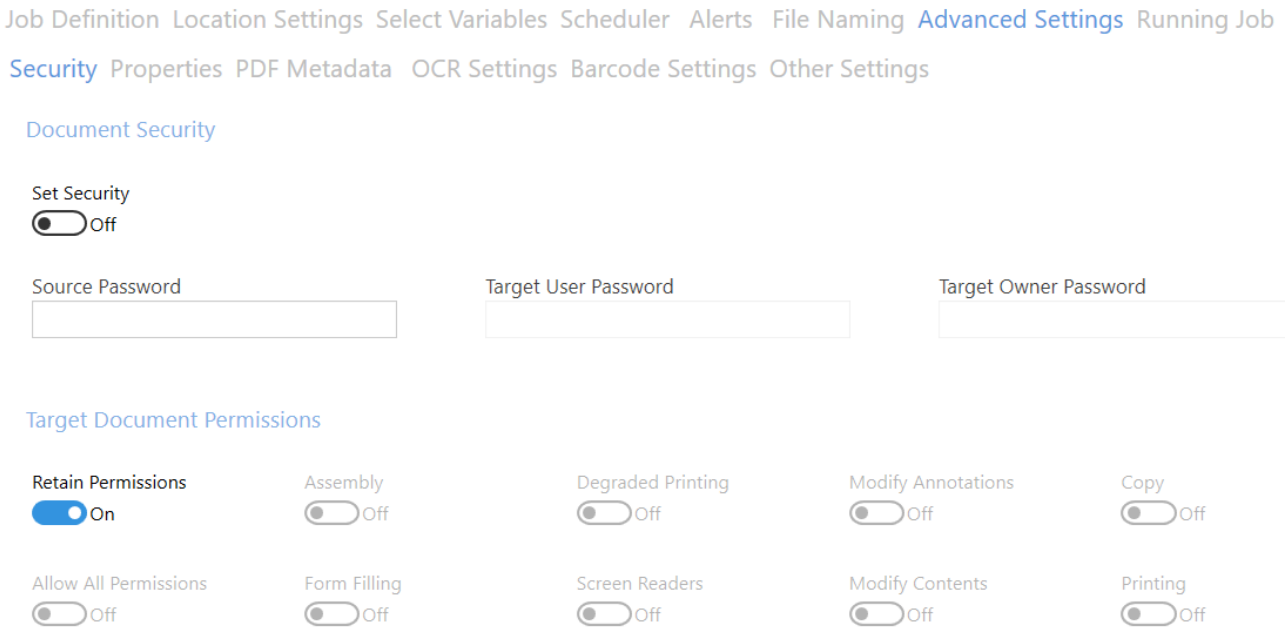


The output file from the example above is created in a sub-folder called "Patent Application Publication".



2.2.7 Advanced Settings

2.2.7.1 Security



2.2.7.1.1 Source Document Security

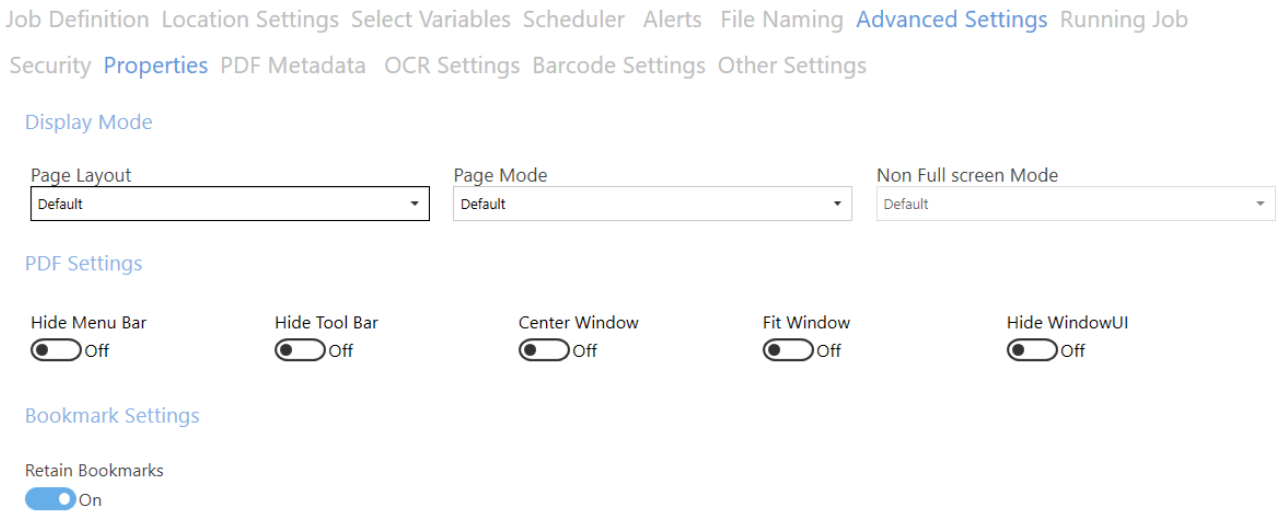
Screen Field / Button	Description
Source Password	If the source document requires a password it should be entered here.

2.2.7.1.2 Target Document Security

Option	Description
User Password	A password that will be required to open the document.
Owner Password	A password that will be required to change the document permissions.
Permissions	**Permission** / **Description** Allow ALL Permissions — All the permissions below. Printing — Allow high-quality printing. Modify Contents — Allow assembly (see below) and other document medications. Copy — Allow text and graphic copying and extraction. Modify Annotations — Allow modification of annotations. Form Filling — Allow filling of form fields. Screen Readers — Allow extraction of text and graphics in support of accessibility. Assembly — Allow rotation, insertion or deletion of pages. Degraded Printing — Allow low-quality printing.

2.2.7.2 Properties

The default display properties of a PDF can be set here.



2.2.7.2.1 Display Modes

Option	Description
Page Layout	The setting for initial document page display. Choose from: • Single Page • Continuous • Continuous Facing (odd pages left) • Continuous Facing (odd pages right)
Page Mode	The setting for initial viewer mode. Choose from: • Neither Bookmarks nor Thumbnails Open • Bookmarks Open • Thumbnails Open • Bookmarks & Thumbnails Open • Full Screen
Non-Full Screen Mode	Only applicable where Page Mode=Full Screen. The setting for document page display when exiting Full Screen mode. One of: • Neither Bookmarks nor Thumbnails Open • Bookmarks Open • Thumbnails Open

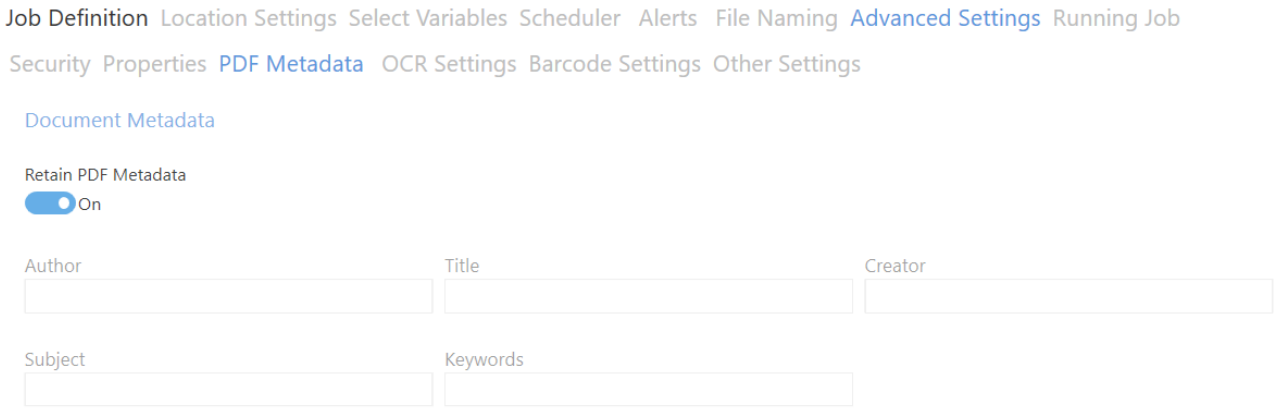
2.2.7.2.2 Settings

Option	Description
Hide Tool Bar	The viewer's tool bar will be hidden.
Hide Menu Bar	The viewer's menu bar will be hidden.
Hide Window UI	The viewer's UI elements (scrollbars etc.) will be hidden.
Fit Window	The viewer will resize the document's window to fit the size of the first displayed page.
Centre Window	The document window will be positioned in the center of the screen.

2.2.7.2.3 Bookmark Settings

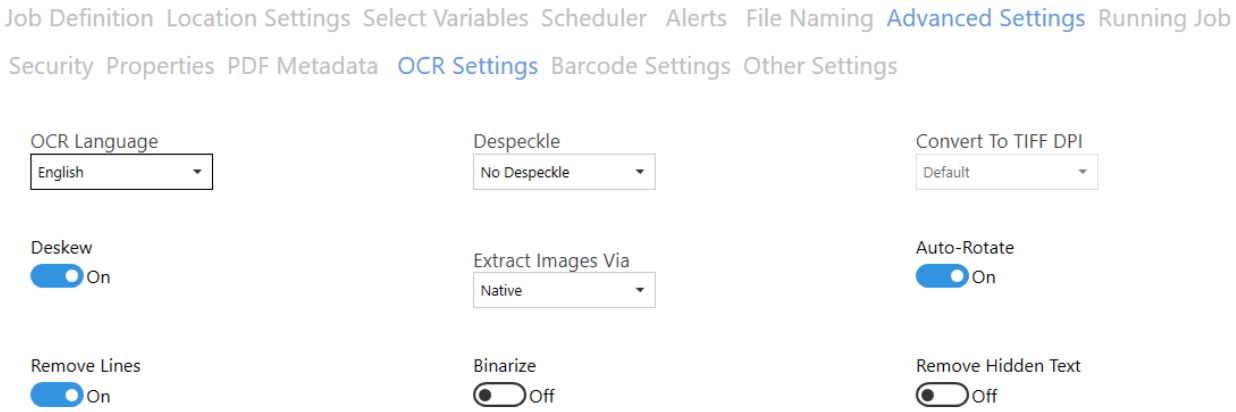
Option	Description
Retain Bookmarks	Select this option to carry over the bookmarks in the original PDF to the split PDF. Default is set to true.

2.2.7.3 PDF Metadata



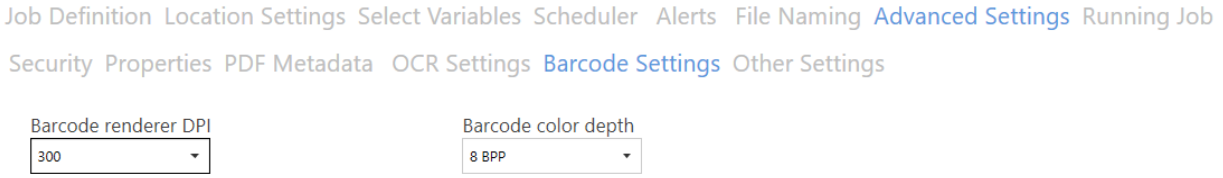
Option	Description
Retain PDF Metadata	Set to yes if you want the output PDF file to retain the metadata of the input PDF, when you set this to true, Kingfisher will disable all the metadata options below.
Author	Provide a new value for the **Author**
Title	Provide a new value for the **Title**
Creator	Provide a new value for the **Creator**
Subject	Provide a new value for the **Subject**
Keywords	Provide a new value for the **Keywords**

2.2.7.4 OCR Settings



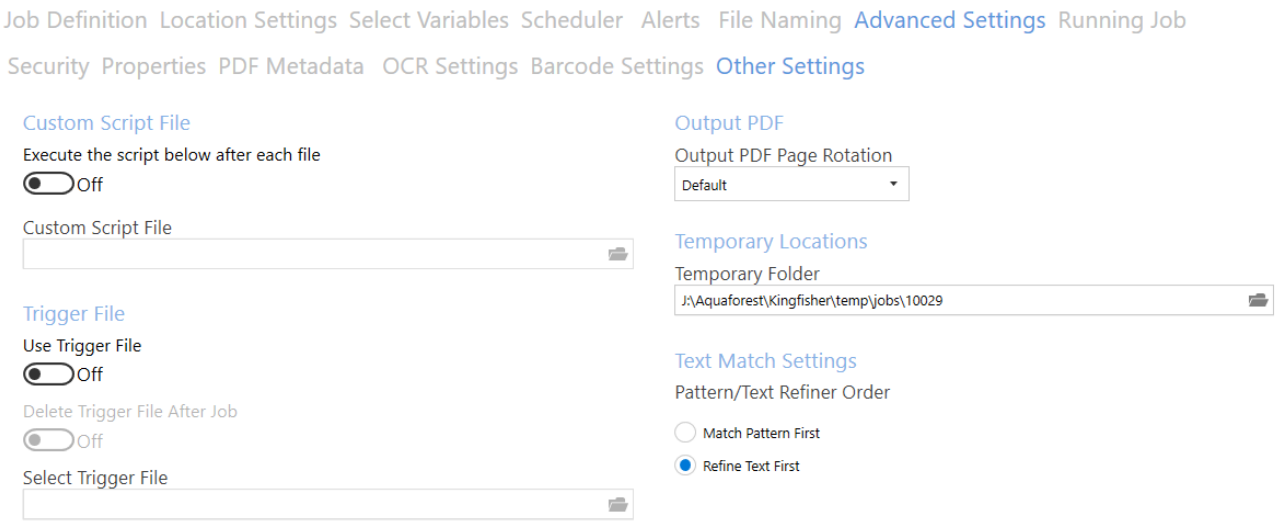
Option	Description
OCR Language	Select expected language for OCR
Despeckle	Despeckle the image – The method removes all disconnected elements within the image that have height or width in pixels less than the specified figure. The maximum value is 9 and the default value is 0.
Convert To TIFF DPI	Select DPI for intermediate TIFF file for OCR when extracting images via the non-native mode.
Deskew	Deskew (straighten) the image.
Extract Images Via	Extract images for OCR as native or by converting the PDF to an intermediate TIFF file. The latter is useful if the file contains complex content (Forms, text and images etc.). Use in conjunction with the Convert To TIFF DPI to change the DPI of the image that is OCR'd. This can improve OCR accuracy depending on the file.
Auto Rotate	Auto rotates images to an orientation where more text can be recognized.
Remove Lines	Remove lines from images.
Binarize	This option is useful when processing colored images. It converts the page into an intermediate black and white image which can improve accuracy.
Remove Hidden Text	Remove existing hidden text. Some scanners produce PDFs with partially OCR'd text which can lead to incorrect results.

2.2.7.5 Barcode Settings



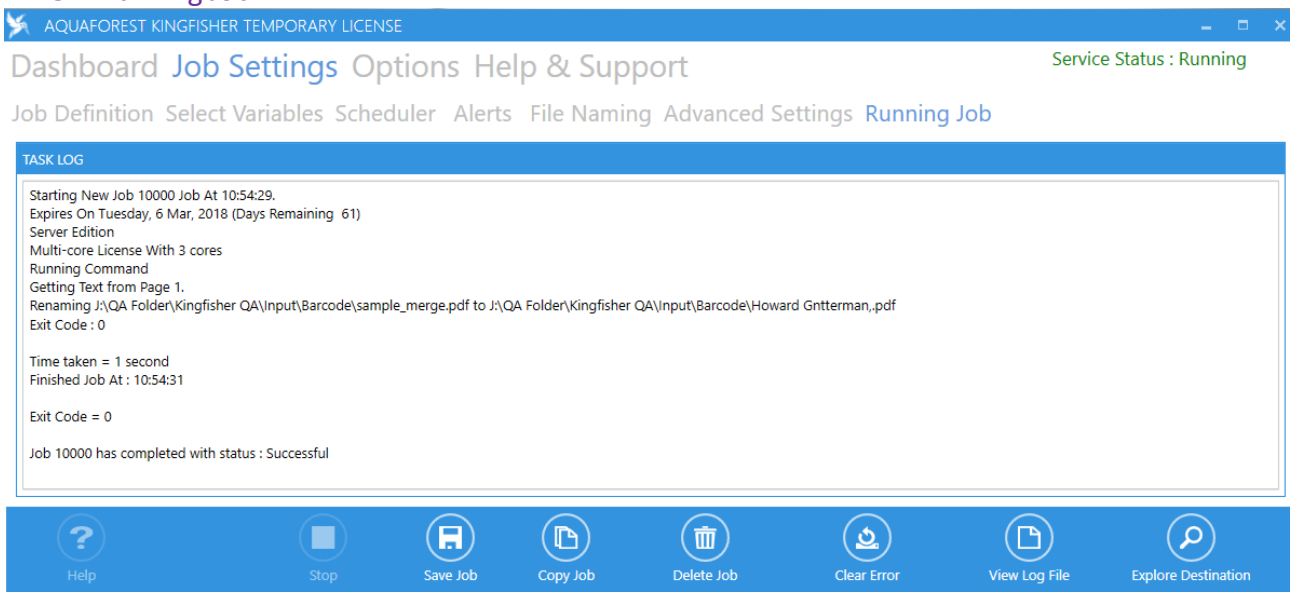
Option	Description
Barcode renderer DPI	Set the DPI used by the barcode engine to read the barcode. With certain images results can be improved by increasing the DPI above the nominal image DPI.
Barcode color depth	Set the color depth (bits per pixel) used by the barcode engine to read the barcode. Though barcodes are traditionally in black and white, color images may include barcodes where the lines or background are in shades or are crossed by colored lines. Changing the color depth can improve results.

2.2.7.6 Other Settings



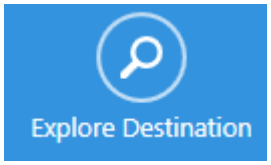
Option	Description
Custom Script File	
Execute the script below after each file	Switch this on if you want to execute a Windows Script File (.wsf) after each output file is produced. We send the following arguments: "newfilename" and "oldfilename".
Custom Script File	The location of the script file.
Trigger File	
Use Trigger File	When this option is switched on, a Kingfisher job will not start if a trigger file specified is not available.
Delete Trigger File After Job	When the Use Trigger File option is on, Kingfisher will delete the trigger file after the job completes.
Select Trigger File	Use this control to provide the trigger file name and location.
Output PDF	
Output PDF Page Rotation	Choose the rotation of the output PDF pages.
Retain Bookmarks	Set to true to retain the bookmarks in the output PDF file.
Temporary Folder	You can change the temp folder using this directory browser control.
Text Match Settings	
Match Pattern First	Kingfisher will match the extracted text with the provided pattern before it applies the text refiner
Refine Text First	Kingfisher will apply the text refiner to the extracted text before matching it to the pattern.

2.2.8 Running Job



The **Task Log** displays the progress of the current or most recent run of the selected job.

The toolbar at the bottom has an additional **Explore Destination** button



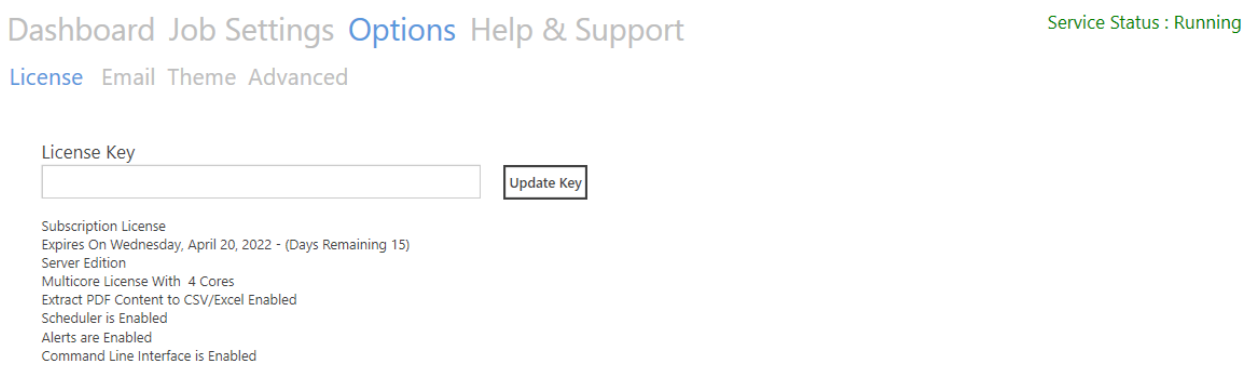
This opens the destination folder if it exists.

2.3 Options

The options tab contains general settings in Kingfisher like license keys, Smtp Settings, Themes etc.

2.3.1 License Keys

Enter the license key in the License key text box and click the “UPDATE KEY” button



Note if the entered license key is invalid it is not saved – the previous license key is retained.

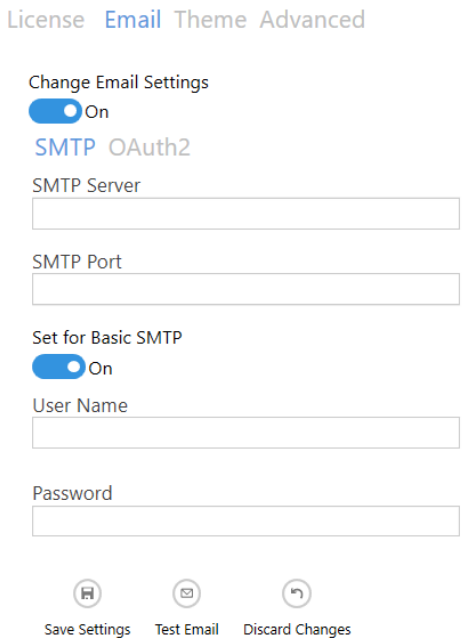
2.3.2 Email

Kingfisher supports SMTP and OAuth2 email authentication.

2.3.2.1 Username and password

Username and password authentication is still supported by Kingfisher, though this is being deprecated by email providers.

Click on the Change Settings toggle to change the settings.

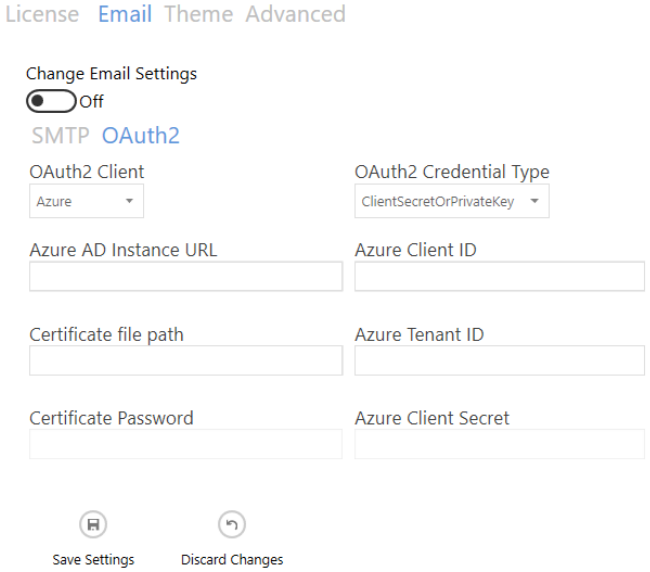


Option	Description
SMTP Server	Enter the address of your SMTP server.
SMTP Port	Enter the SMTP port number. Port 587 is the standard port number
Set for Basic SMTP	Set On for Basic SMTP (username and password)
Username	Enter your SMTP Username.
Password	Enter your SMTP password.
Save Settings	Click to update the email settings and set email type to SMTP.
Test Email	Click to test the SMTP settings by sending an email from the Username to the Username using the SMTP settings on this tab.
Discard Changes	Click this button to revert to the last saved email settings.

2.3.2.2 Modern (OAuth2) Authentication

The Modern (or OAuth2) Authentication provides a more secure method to access email.

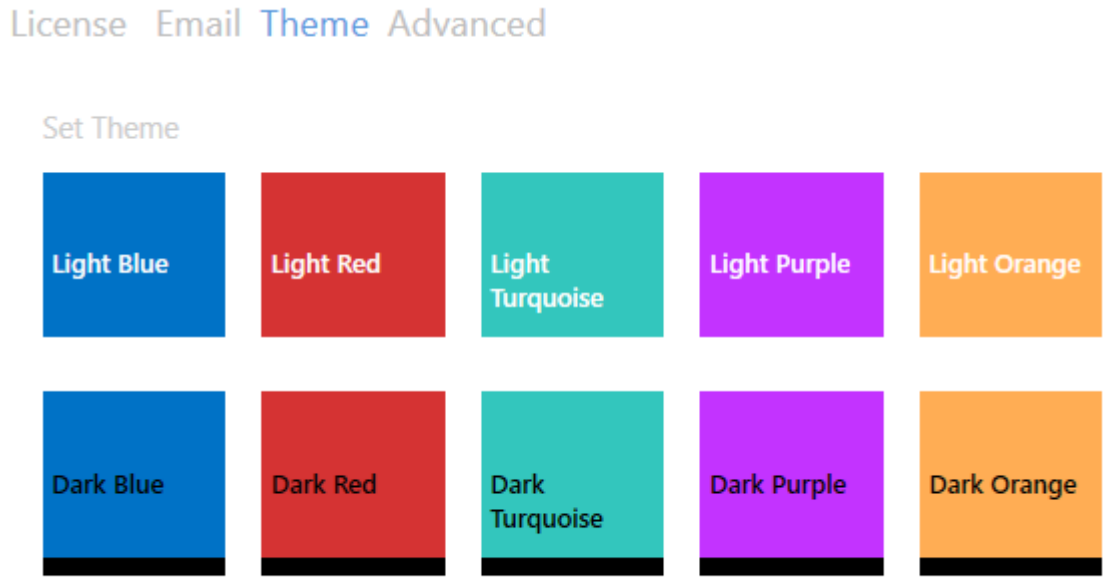
See https://go.microsoft.com/fwlink/?linkid=2083908 for details on setting up the required Microsoft registration.



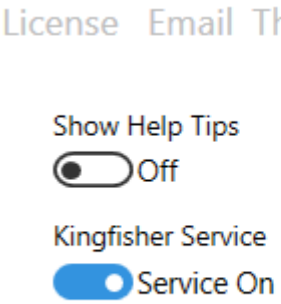
Screen Field / Button	Description
OAuth2 Client	Currently only Azure.
OAuth2 Credential Type	Two options: • ClientSecretOrPrivateKey • Certificate
Azure AD Instance URL	Azure Active Directory authentication URL.
Azure Client ID	Azure Application (Client) ID
Azure Tenant ID	Azure Active Directory Tenant ID
Certificate File Path (Certificate credential type)	Path to the .pfx certificate associated with the Client ID
Certificate Password (Certificate credential type)	Password for the .pfx certificate
Azure Client Secret {ClientSecretOrPrivateKey credential type)	Client secret (Private Key) for the Azure Client
Save Settings	Click to update the email settings and set email type to OAuth2.
Discard Changes	Click this button to revert to the last saved email settings.

2.3.3 Theme

Use this tab to switch between the different themes available in Kingfisher.



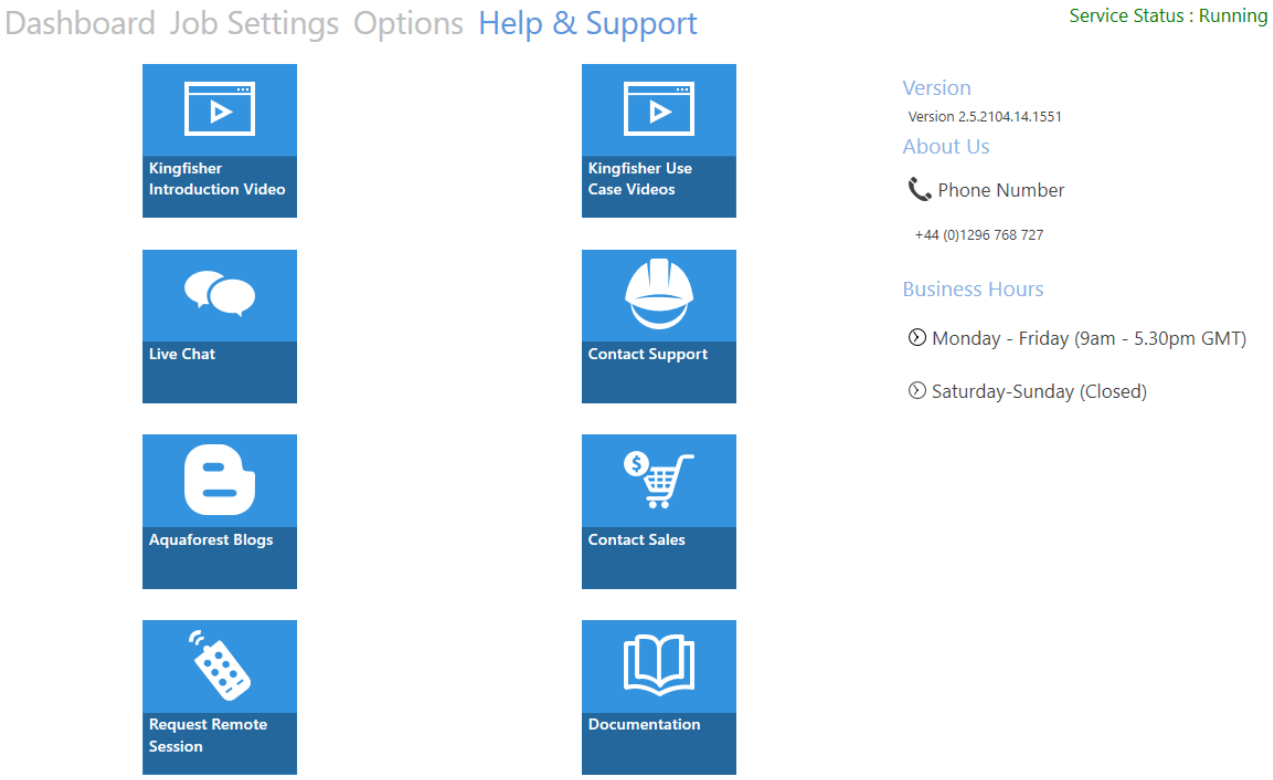
2.3.4 Advanced Settings



Option	Description
Show Help Tips	Switch this on if you want Kingfisher to show you tips as you use it.
Kingfisher Service	Use this switch to On or Off the Kingfisher service.

2.4 Help

The help Tab give you access to all the different ways you can get support on how to use Kingfisher.



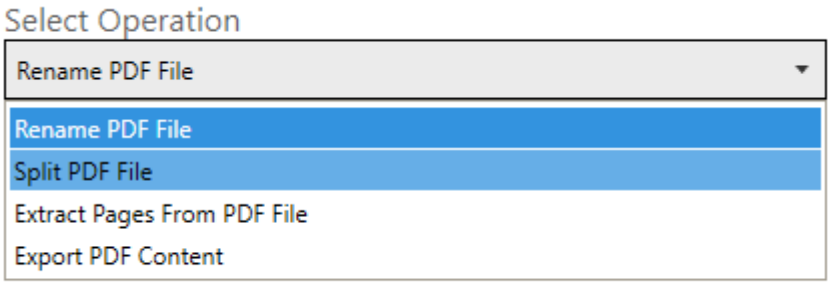
Option	Description
Kingfisher Introductory Video	Opens a folder containing an introductory video.
Kingfisher Use Case Videos	Opens a folder containing the use case videos.
Live Chat	Opens the Kingfisher product page including a live chat option in your default web browser.
Contact Support	Launches the default Email Client with the email address of our support team.
Aquaforest Blogs	This takes the user to the Aquaforest Blogs.
Contact Sales	Launches the default Email Client with the email address of our sales team.
Request Remote Session	Opens the Aquaforest Request page on the web site. Enter your details and click on Send Message.
Documentation	Opens the folder containing the Kingfisher documentation.

3 Kingfisher Operations

Kingfisher has 4 main operations which are:

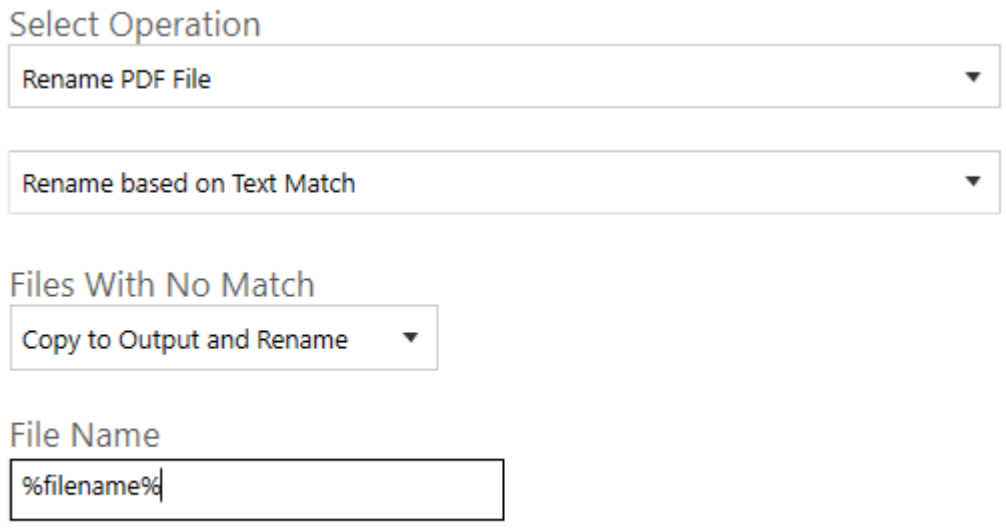
- Rename PDF File
- Split PDF File
- Extract Pages from PDF File
- Extract Content from PDF File

To select an operation, select one of the Main Operations in the “Select Operation” combo-box.



Each of the operation listed above are further divided into sub-categories, the section below explains each of these operations in more details.

3.1 Rename PDF File



<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td>Files With No Match</td><td>This gives the user three options of what to do with files that do not match the text or barcode criteria specified by the user.

• Do Not Copy to Output
• Copy to Output
• Copy to Output and Rename</td></tr>
<tr><td>File Name</td><td>If the “Copy to Output and Rename” option is selected you will need to provide a filename mask like the one in the “File Naming” tab, the table below gives you a breakdown of the variables available.

<table>
<tr><td>%FILENAME%</td><td>The original file name without the .pdf extension.</td></tr>
<tr><td>%BOOKMARK%</td><td>This represents the bookmark name and can be used with the “Split by Top Level Bookmarks” split option only.</td></tr>
<tr><td>%TIME%</td><td>The current Time.</td></tr>
<tr><td>%DATE%</td><td>The current Date.</td></tr>
<tr><td>%RUNDATE%</td><td>The Time at which the job was executed, the difference between this and the %TIME% variable is that this value will be same for all files processed in a single job run.</td></tr>
<tr><td>%JOBNAME%</td><td>The Name of the Job.</td></tr>
</table>
</td></tr>
</table>

3.1.1 Rename Based on Barcode

This operation will take a PDF file that has barcodes in its pages as an input and renames the PDF file based on the value of the barcode. You can also specify Areas on the PDF page; Kingfisher will check these Areas and use the barcode present there. If no Areas are selected, Kingfisher will use the first barcode it encounters in the PDF document.

3.1.2 Rename Based on Text Match

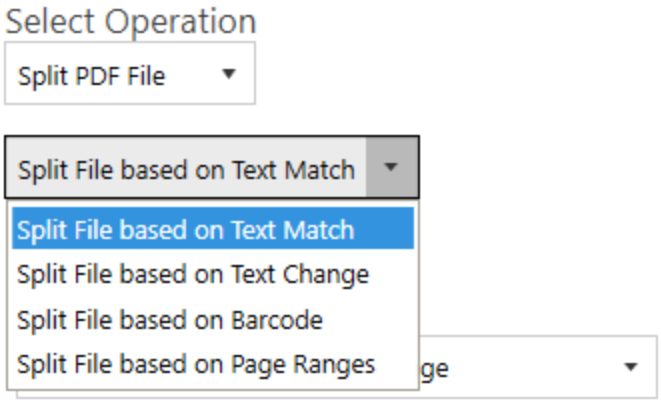
This step is like the step above except that it uses text in the PDF document to rename the file.

Unlike the barcode step, an Area must be selected for this step to work. You can also specify Regular Expressions and Words that appear before/after the word you want the file to be named to. These extra features are to help you validate the text you select.

3.1.3 Rename Based on Barcode and Text Match

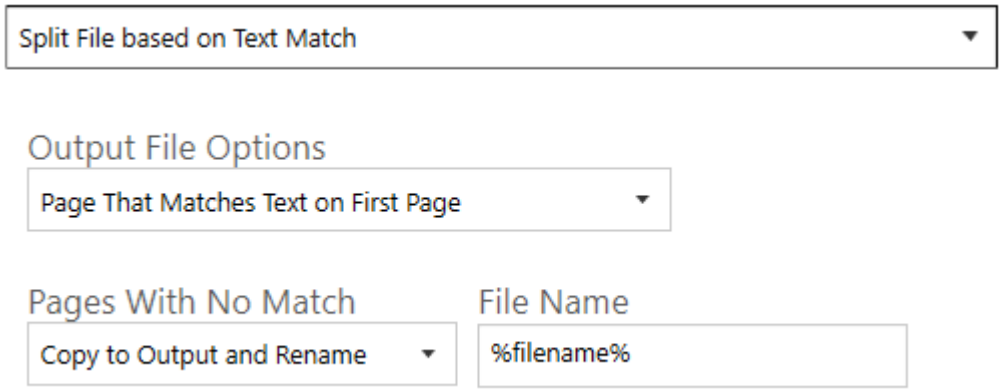
This step is a combination of the two steps above, it allows a user to specify barcodes and text on a PDF page that can be used to create the filename.

3.2 Split PDF File



3.2.1 Split File based on Text Match

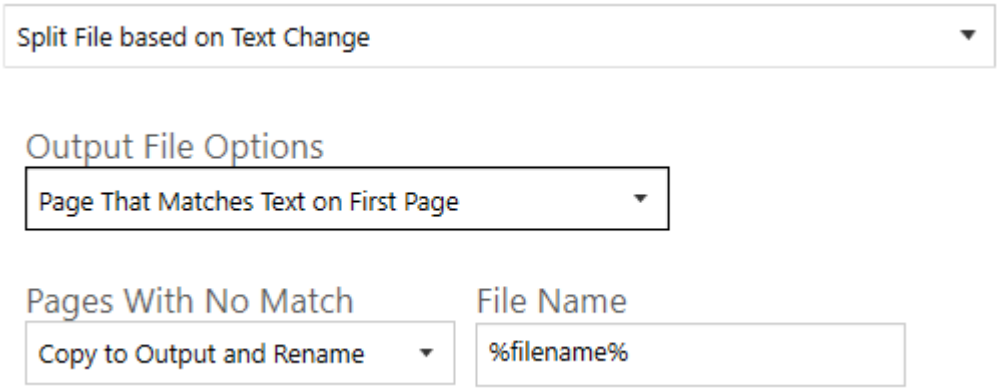
The split option is like the extract option except that instead of extracting a single page, it will split the file into pages using the pages that matches the text value as a divider.



<table>
<tr><th>Screen Field / Button</th><th>Description</th></tr>
<tr><td>Pages That Matches on First Page</td><td>This option puts the page that matches the text value as the first page in the output file.</td></tr>
<tr><td>Pages That Matches on Last Page</td><td>This option puts the page that matches the text value as the last page in the output file.</td></tr>
<tr><td>Remove Page That Matches Text</td><td>This option removes the pages that match the text expressions in the output files.</td></tr>
<tr><td>Pages with No Match</td><td>This gives the user three options of what to do with pages that do not match the text or barcode criteria specified by the user.
Do Not Copy to Output
Copy to Output
Copy to Output and Rename</td></tr>
<tr><td>File Name</td><td>If the “Copy to Output and Rename” option is selected you will need to provide a filename mask like the one in the “File Naming” tab, the table below gives you a breakdown of the variables available.<table>
<tr><td>%FILENAME%</td><td>The original file name without the .pdf extension.</td></tr>
<tr><td>%BOOKMARK%</td><td>This represents the bookmark name and can be used with the “Split by Top Level Bookmarks” split option only.</td></tr>
<tr><td>%TIME%</td><td>The current Time.</td></tr>
<tr><td>%DATE%</td><td>The current Date.</td></tr>
<tr><td>%RUNDATE%</td><td>The Time at which the job was executed, the difference between this and the %TIME% variable is that this value will be same for all files processed in a single job run.</td></tr>
<tr><td>%JOBNAME%</td><td>The Name of the Job.</td></tr>
</table></td></tr>
</table>

3.2.2 Split File based on Text Change

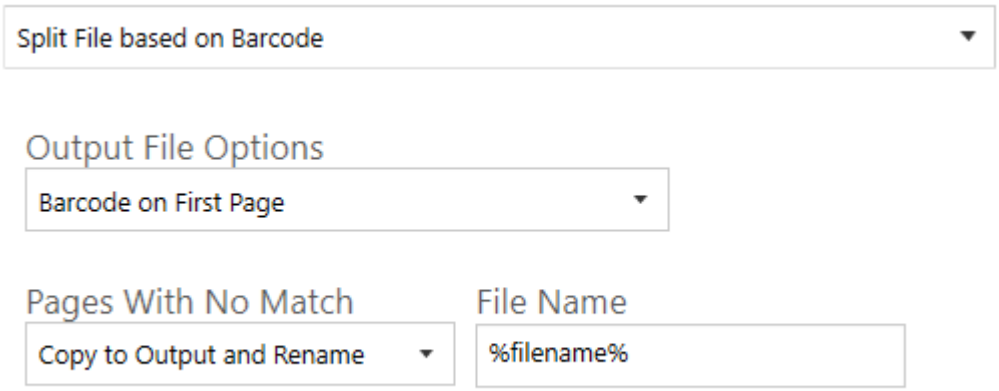
The split option is similar to the "Split File based on Text Match" option except that instead of splitting everywhere there is a text match, it waits until the text that was matched has changed from the previous text match.



Screen Field / Button	Description
Pages That Matches on First Page	This option puts the page that matches the text value as the first page in the output file.
Pages That Matches on Last Page	This option puts the page that matches the text value as the last page in the output file.
Remove Page That Matches Text	This option removes the pages that match the text expressions in the output files.

3.2.3 Split File based on Barcode

The split option is similar to the extract option except that instead of extracting a single page, it will split the file into pages using the pages that matches the barcode value as a divider.



Screen Field / Button	Description
Barcode on First Page	This option puts the page that matches the barcode value as the first page in the output file. Subsequent pages are added to that file. A new barcode match will create a new file and the subsequent pages are added to this new file.
Barcode on Last Page	This option puts the page that matches the barcode value as the last page in an output file. Subsequent pages will be added to a document based on the next matching barcode value
Remove Barcode Page	This option splits on and removes the first page that contains a barcode (which can be specified by type or location). Subsequent pages are output to the output files. Each new barcode creates a new output file. No post processing of barcode value is possible with this option.

3.2.4 Split File based on Page Ranges

This split operation uses characteristics like page numbers and bookmarks to split PDFs documents.

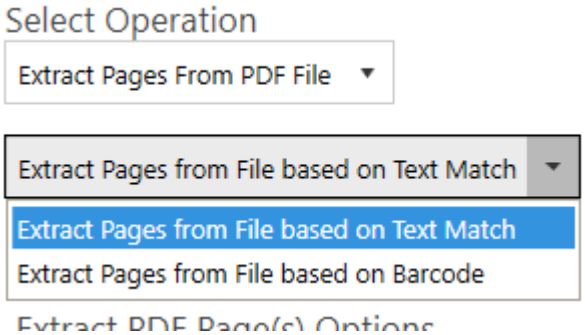
Has four split options.

- Split into Single Pages: Split the PDFs into single pages
- Split by Top Level Bookmarks
- Split by Page Ranges
- Split by Repeating Page Ranges

<table>
<tr><th>Screen Field / Button</th><th>Description</th></tr>
<tr><td>Split Type</td><td>One of:
Split into single pages
Split by top-level bookmarks
Split by page ranges (See below)
Split by repeating page ranges (See below)</td></tr>
<tr><td>Page Ranges</td><td>Set of page ranges separated by commas that defines which pages from the original should be extracted. The following types of page ranges are allowed:

<table><tr><td>1</td><td>Specifies a single page.</td></tr><tr><td>1-3</td><td>Specifies a range of pages.</td></tr><tr><td>10-20even</td><td>Specifies even pages within the range.</td></tr><tr><td>10-20odd</td><td>Specifies odd pages within the range.</td></tr><tr><td>20-lastpage</td><td>Note that lastpage can be used to represent the last page in the document.</td></tr></table></td></tr>
<tr><td>Repeating Range</td><td>Apply the page range at an interval of the repeating number. For example, if 2-4 is specified for page ranges, and 4 is specified as the repeating range, then the range is re-applied every 4 pages.
Hence the file is split such that the first output file contains pages 2-4 from the original document, the second contains pages 6-8 and so on.</td></tr>
</table>

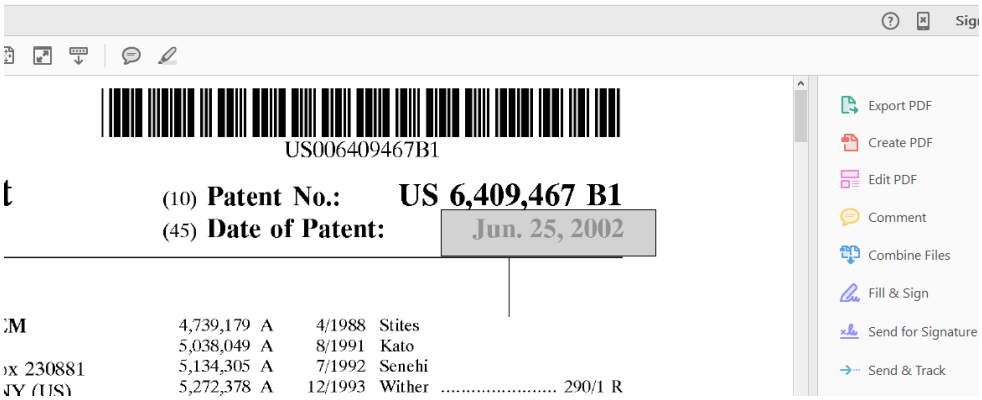
3.3 Extract Pages from PDF File



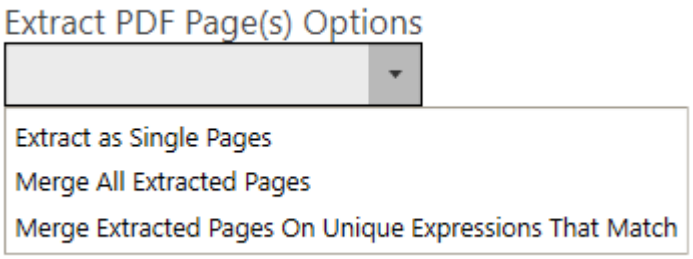
3.3.1 Extract Pages from File based on Text Match

This step allows you to extract single pages from a PDF file based on the presence of a text value in a zone. For example, in the image below you might want to extract any page that has a date value (Jun. 25, 2002) at that location, all you need to do is to is to set up an "Extract Pages from File based on Text Match" job, select the zone and provide some pattern matching to make sure a date is extracted.

Kingfisher also gives you the ability to rename the output file based on the text extracted from the PDF page.



The extract step has the following settings

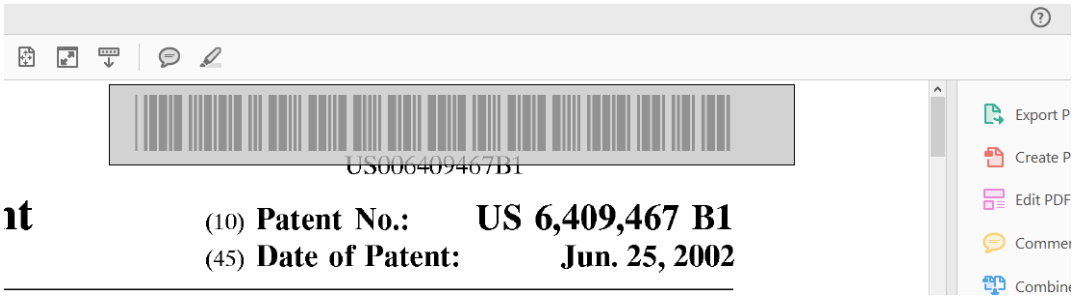


Screen Field / Button	Description
Extract as Single Pages	This option copies all the extracted pages as single page PDF files to the output folder.
Merge All Extracted Pages	Merges all the pages that were extracted into one document.
Merge Extracted Pages on Unique Expressions That Merge	Merges All the Extracted pages that have equal "%VALUE1%" values

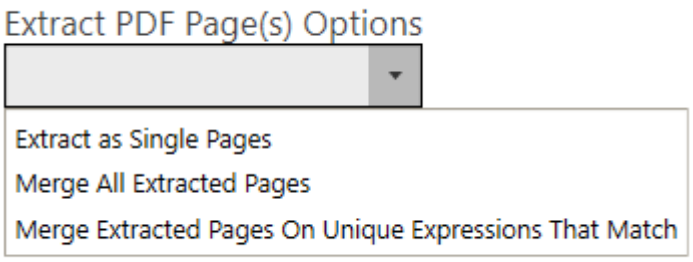
3.3.2 Extract Pages from File Name based on Barcode

This step allows you to extract single pages from a PDF file based on the presence of a barcode value in a zone. For example, in the image below you might want to extract any page that has a barcode at that location, all you need to do is to set up an “Extract Pages from File based on Barcode” job and select the zone you are interested in. If no zone is selected the first barcode found in a page will be used.

Kingfisher also give you the ability to rename the output file base on the value extracted from the PDF page.

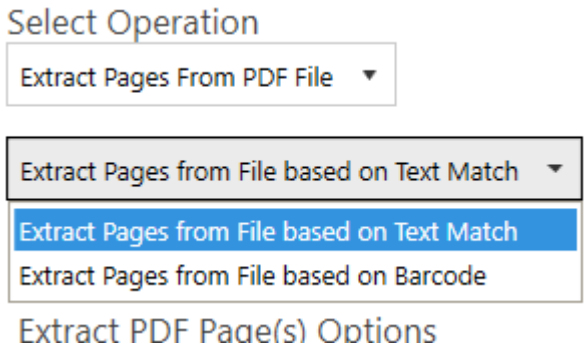


The extract step has the following settings



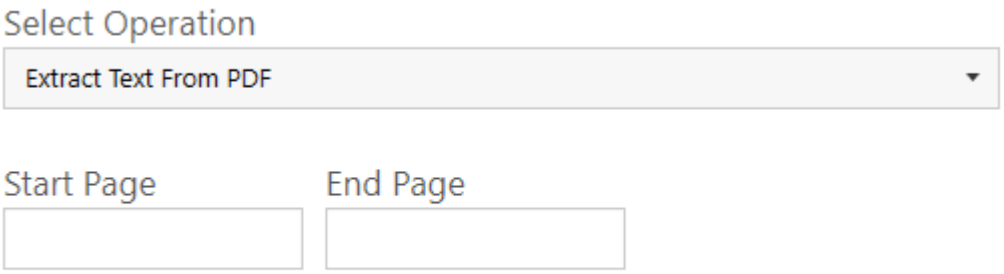
Screen Field / Button	Description
Extract as Single Pages	This option copies all the extracted pages as single page PDF files to the output folder.
Merge All Extracted Pages	Merges all the pages that were extracted into one document.
Merge Extracted Pages on Unique Expressions That Merge	Merges All the Extracted pages that have equal “%VALUE1%” values.

3.4 Extract Content from PDF File



3.4.1 Extract Text from PDF File

This step simply extracts all the text in a PDF file, Kingfisher is intelligent enough to detect image PDF pages and OCR before extracting any text from it. The only type of files we can't extract meaningful text from by default are the ones with font encoding, we advise users to switch OCR for these file types.



Screen Field / Button	Description
Start Page	Page number of the page you want Kingfisher to start extracting text from.
End Page	Page number of the page you want Kingfisher to stop extracting text from.

3.4.2 PDF to CSV/XLSX

This step is used to extract tabular data from PDF files, see section 5 for more details.

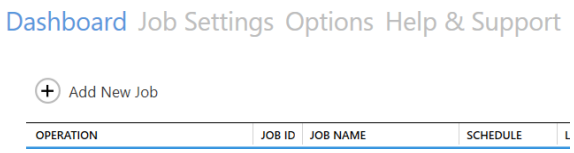
3.4.3 Advanced Export to CSV/XLSX

This step extracts text that appears before/after certain expressions, see section 5 for more details.

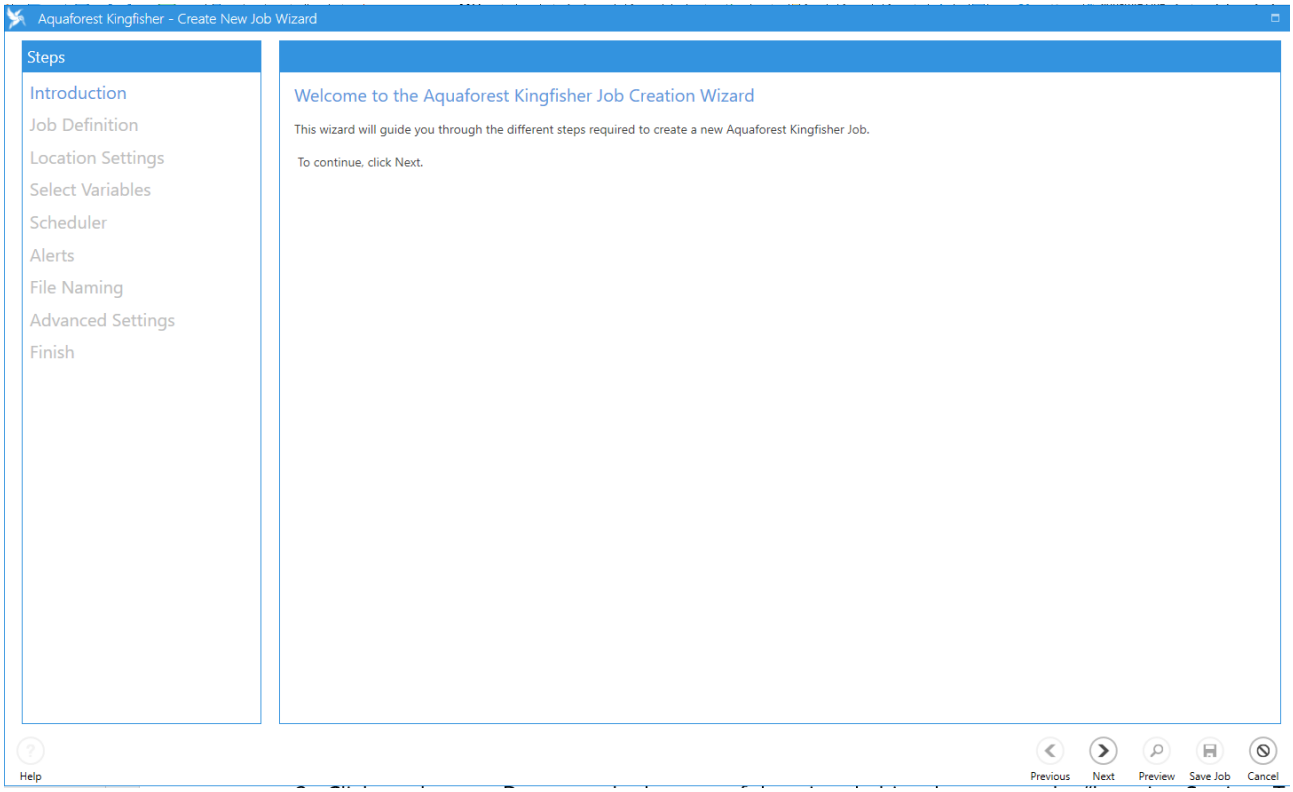
4 Creating and running Jobs with the GUI

4.1 Creating a new Job

To create a new Kingfisher job, go to the Dashboard.

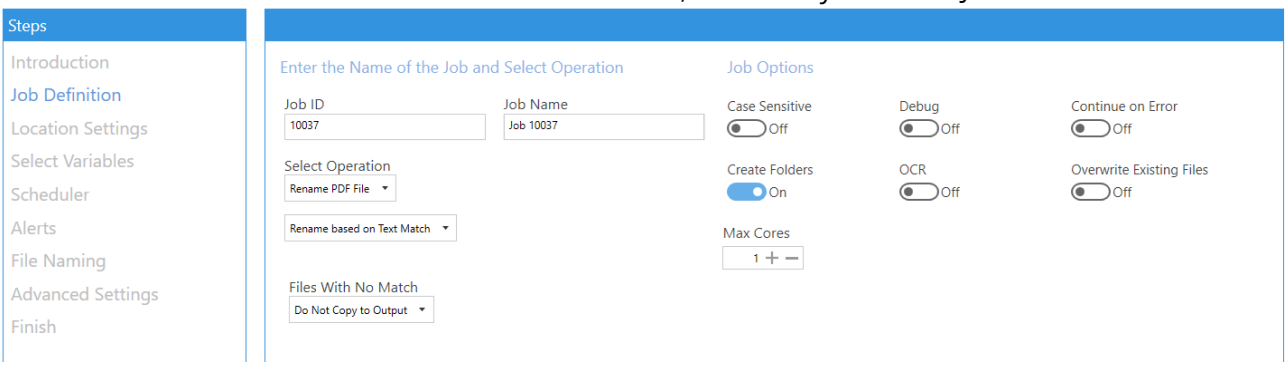


Click on the “Add New Job” button, this will launch the new job wizard shown below.



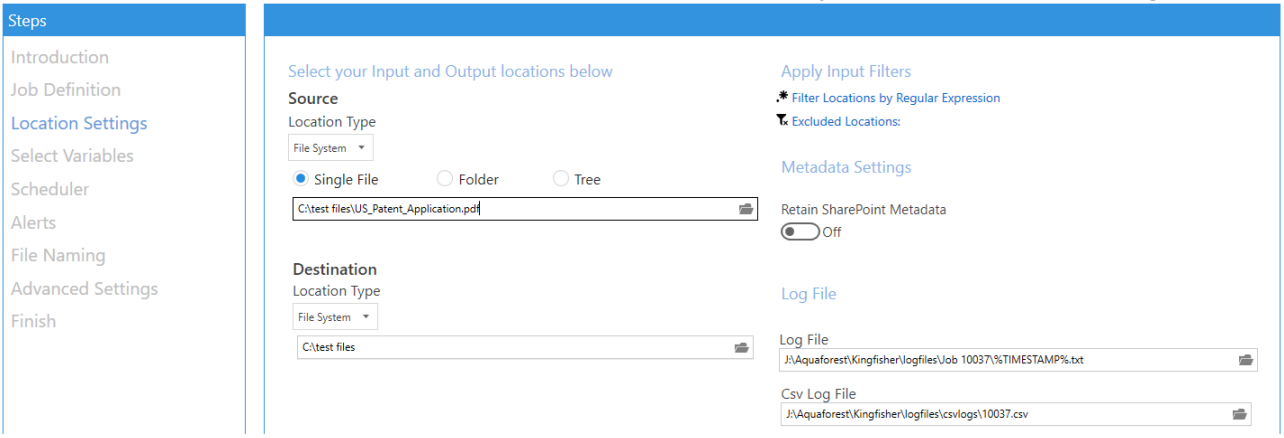
Follow the steps below to create a new job.

1. Click on the next Button at the bottom of the wizard, this takes you to the “Job Definition Tab”.



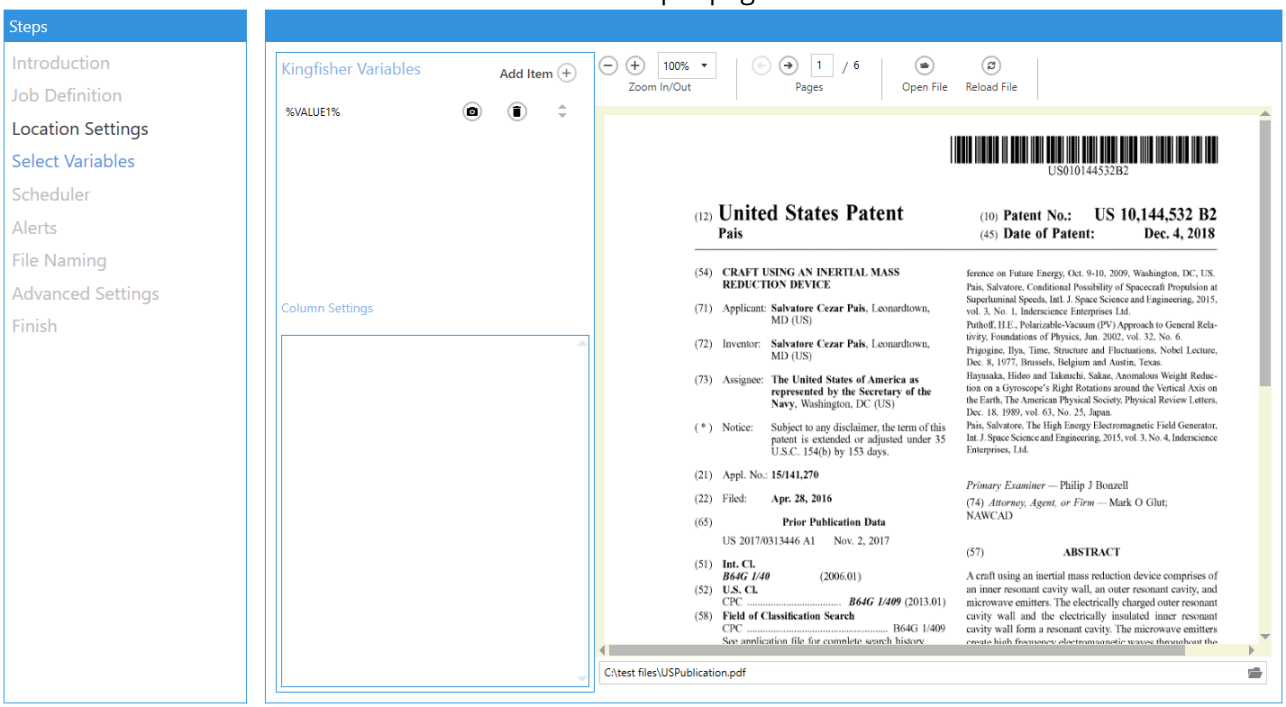
Check section 2.2.1 for explanation of the fields in this window, choose the suitable options for the job.

2. Click on the next Button at the bottom of the wizard, this takes you to the “Location Settings Tab”.



Check section 2.2.2 for explanation of the fields in this window, choose the suitable options for the job and press next.

3. One of the main advantages of Kingfisher is the ability to process pdf files based on the file content. A Zone Definer allows the user to select areas on the pdf page to extract text or barcode values.

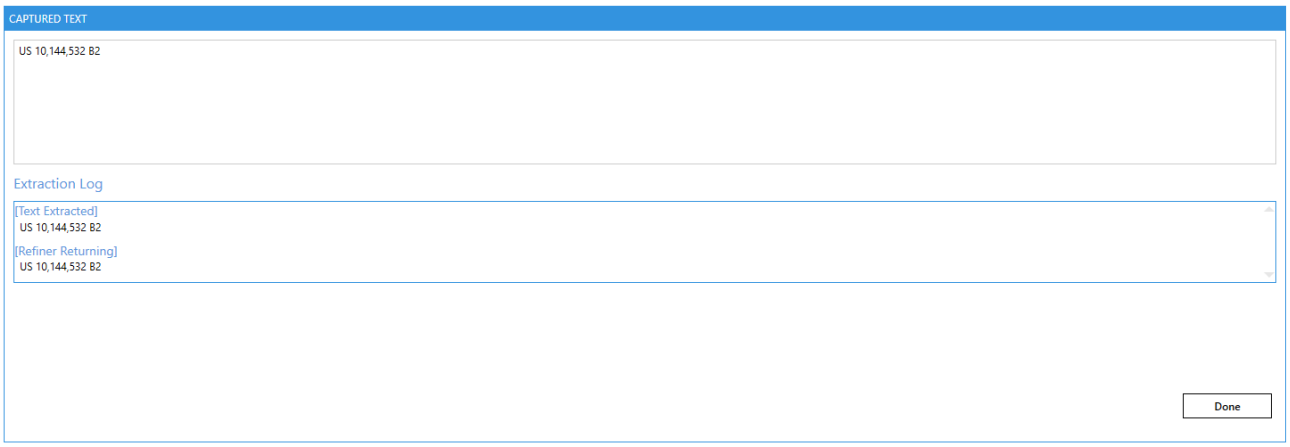


Initially there is one variable (%VALUE1%). If there is no variable click on the Add Item button.
Draw on the image of the PDF on the right-hand side by starting at one corner of the area required and holding down the left mouse button and moving it to cover the area.

If the area is not correct, repeat the procedure.

Once the area is selected, click on the camera icon.

This will display the captured text.

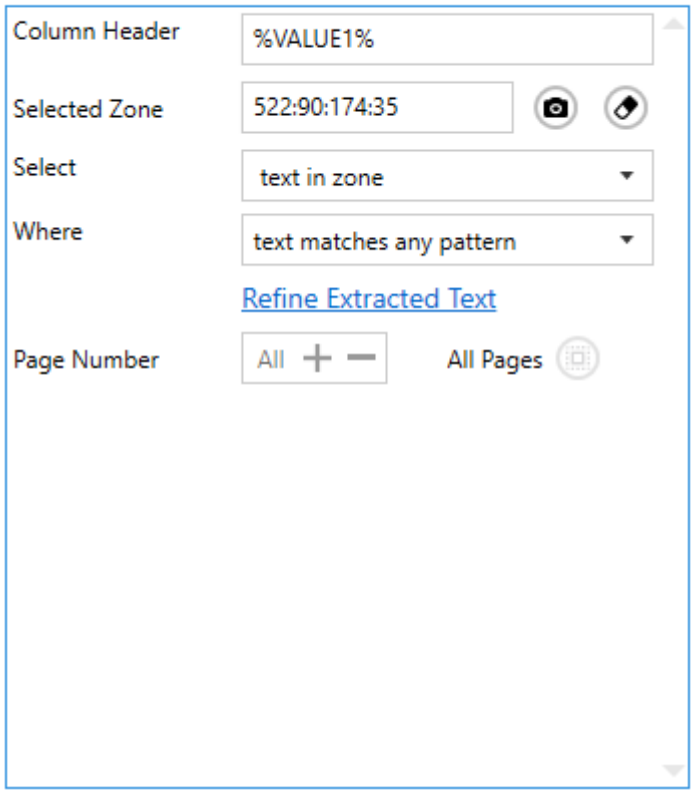


Initially the Extraction Log will just show the same text in the Text Extracted and Refiner Returning sections. The Refiner can be used to remove extraneous text, spacing and punctuation.

Click on the Done button.

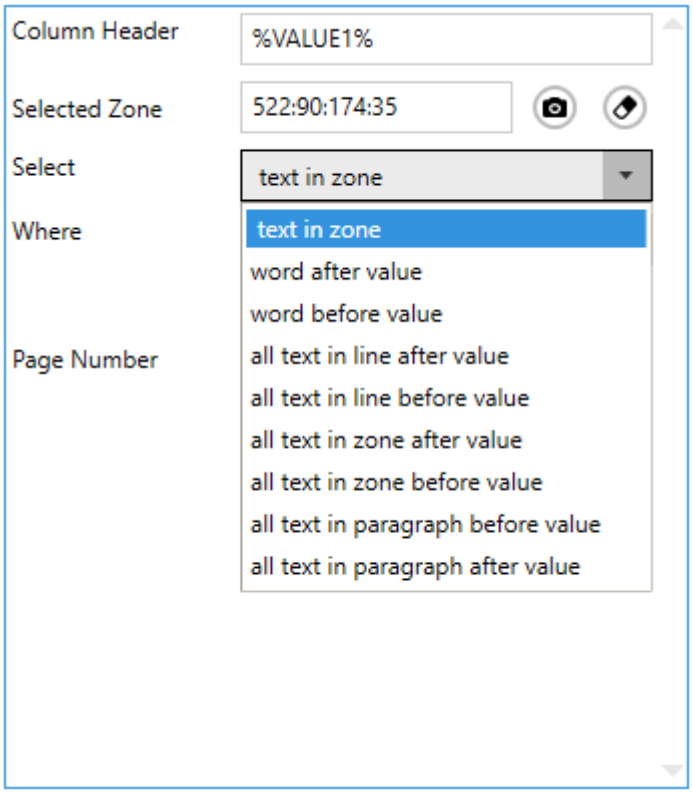
Details of the area will be shown in the Column Settings area.

Column Settings



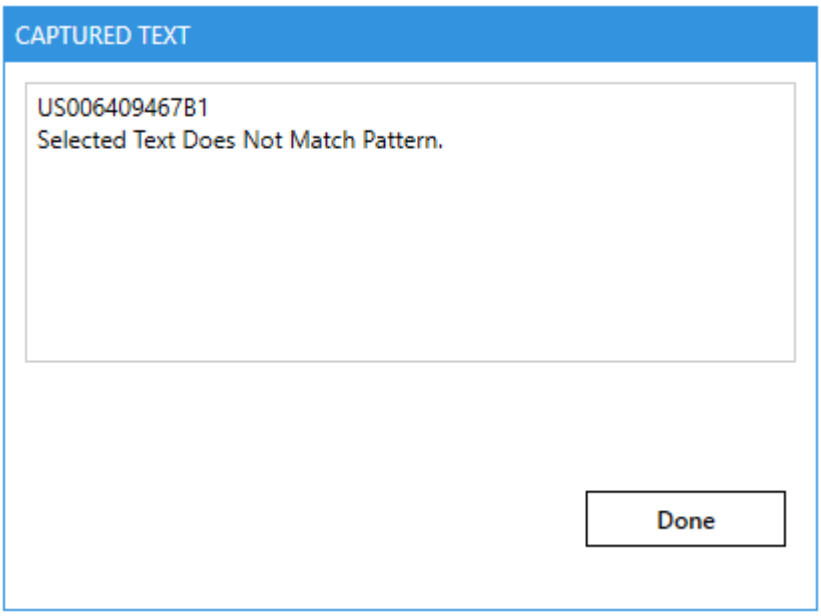
The Selected Zone is the area displayed on the PDF viewer.

The Select option determines how the text is handled

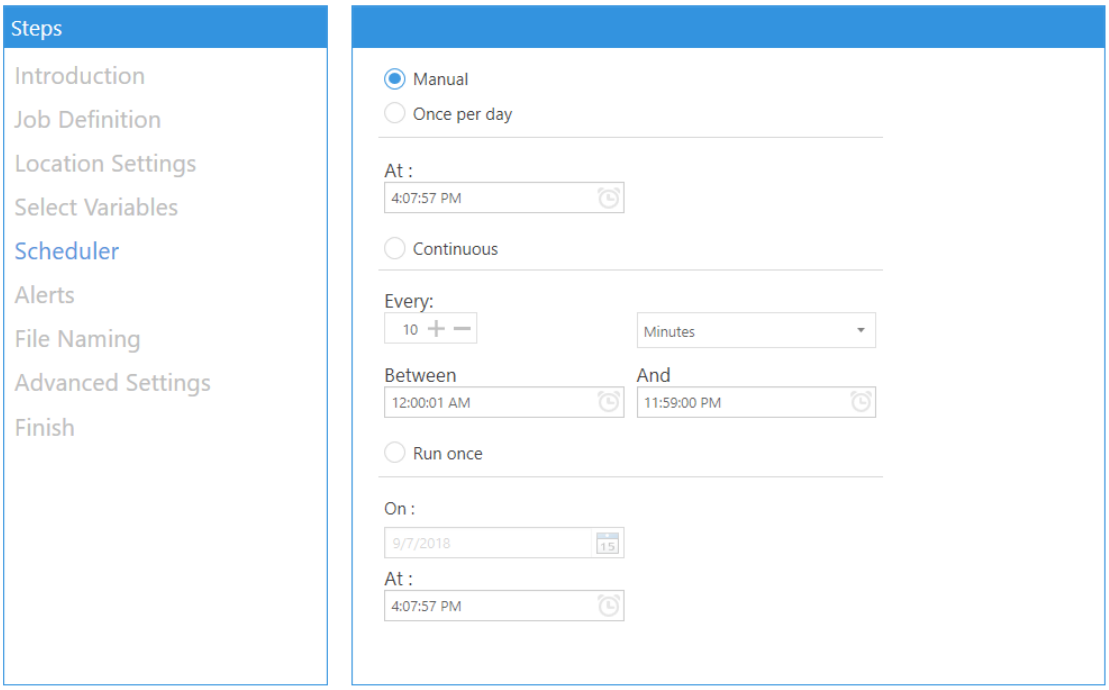


For this example select text in zone.

- Click on the “Camera Icon” button, you will see the text below:

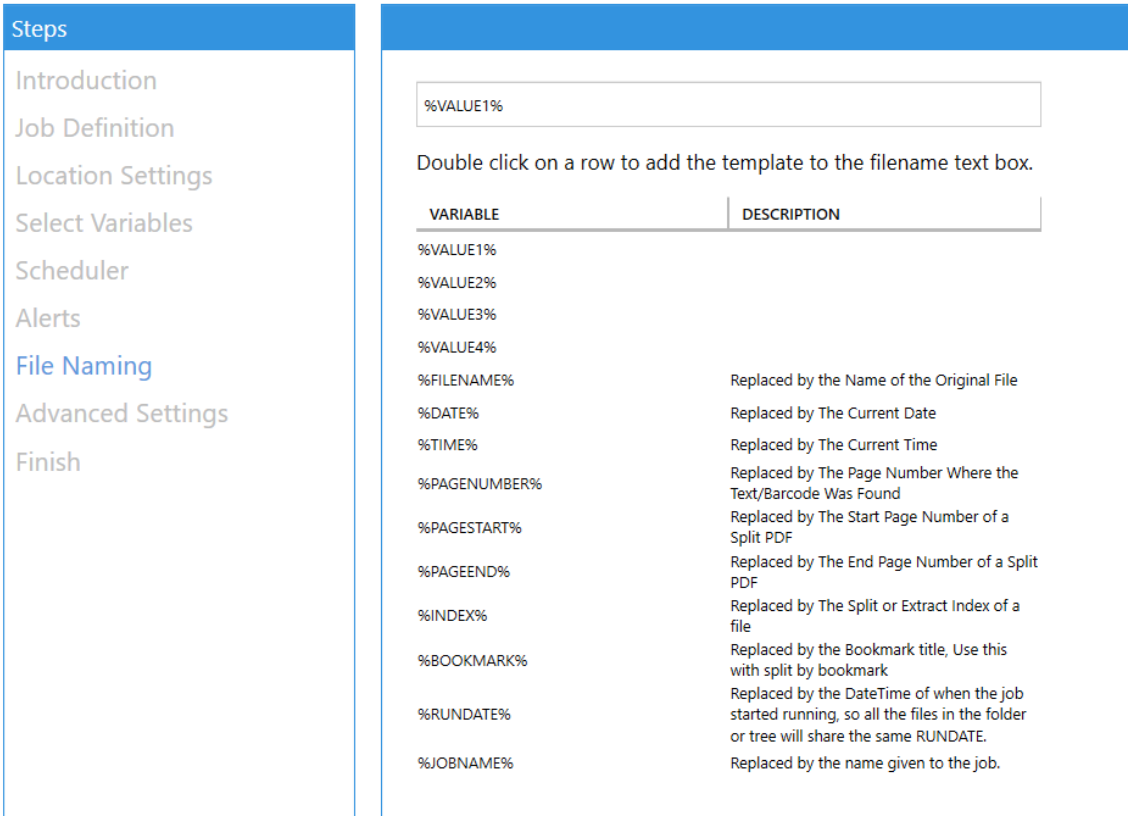


Set up the scheduler, we will use manual for this test.

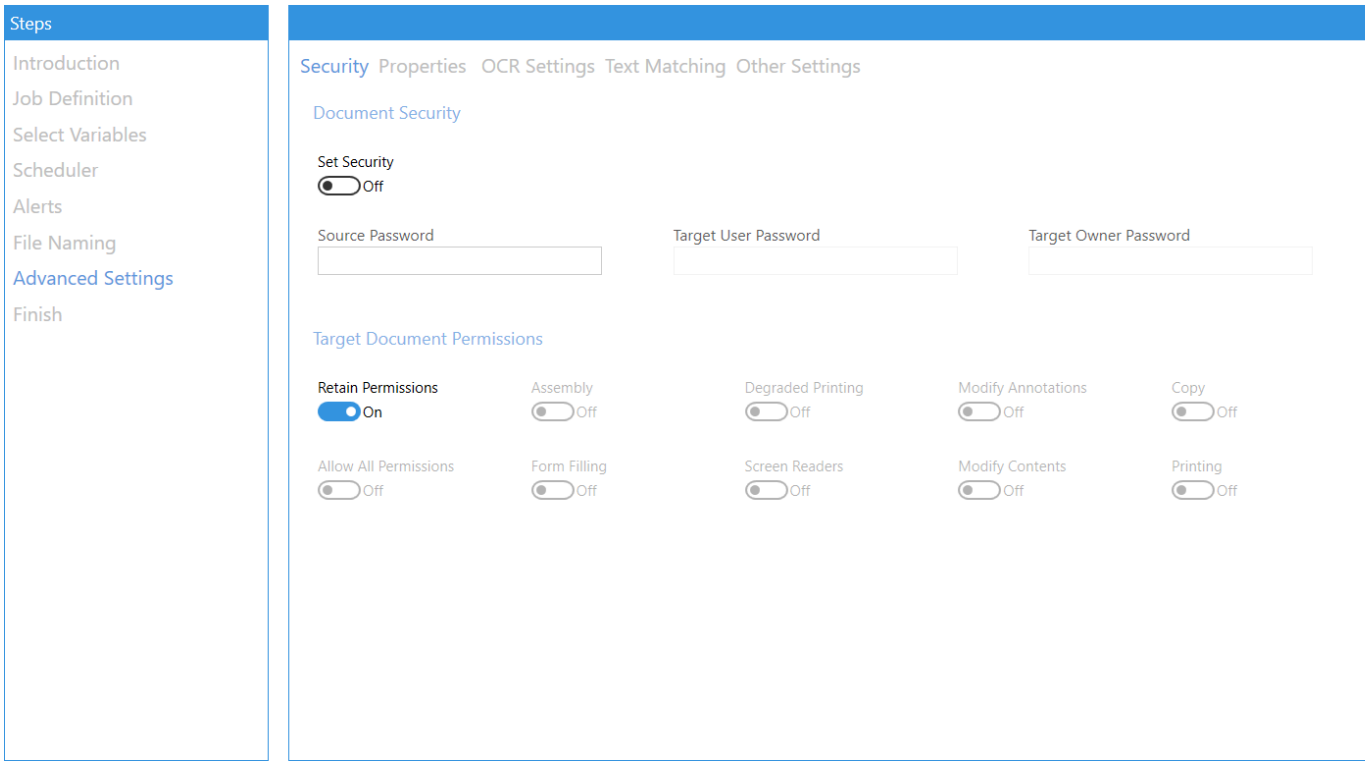


4. Click Next to skip the Alerts for this example.
5. Click Next to skip the Schedule for this example.

6. The “File Naming” tab allows the definition of a template for the output filename

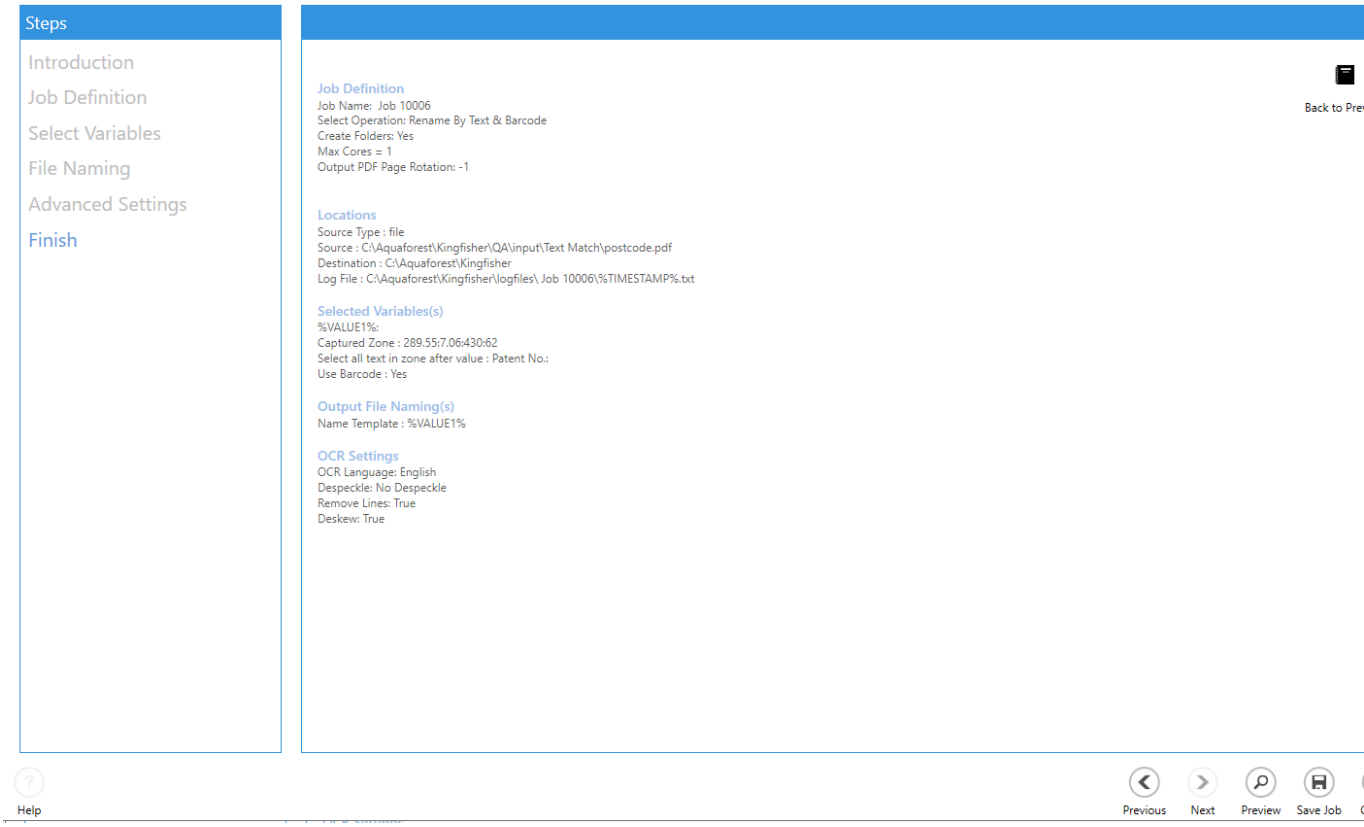


The “Advanced Settings” tab allows the user to provide properties of the output PDF file, OCR settings and post-processing script execution. For more information check section 2.2.7.

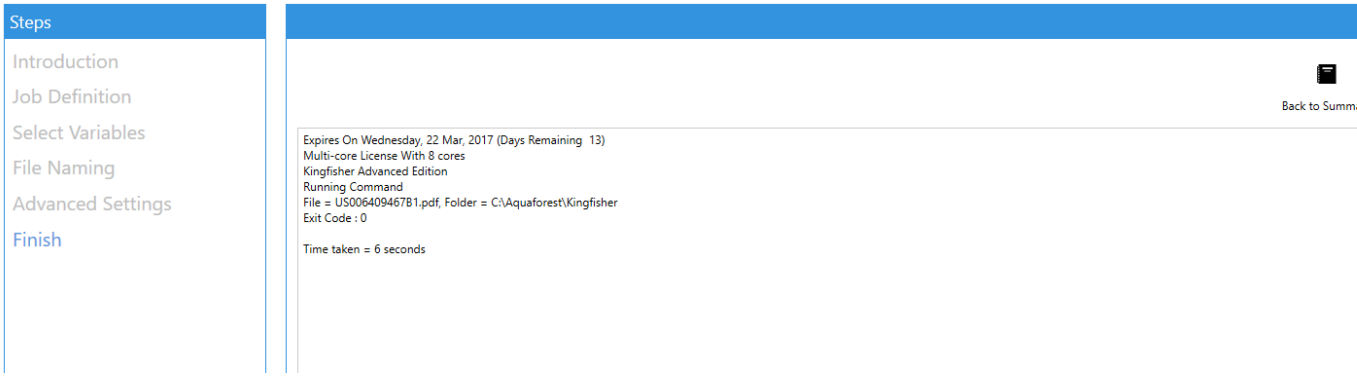


For this example, leave the default settings, click next.

7. The “Finish” tab will show you a summary of the job you set up



8. To confirm that you are happy with the job, click the preview button to see what the outputs generated will look like, note this does not execute the “Extract PDF Contents” step.



9. Click on the “create” button to create the new job.

4.1.1 Error Codes

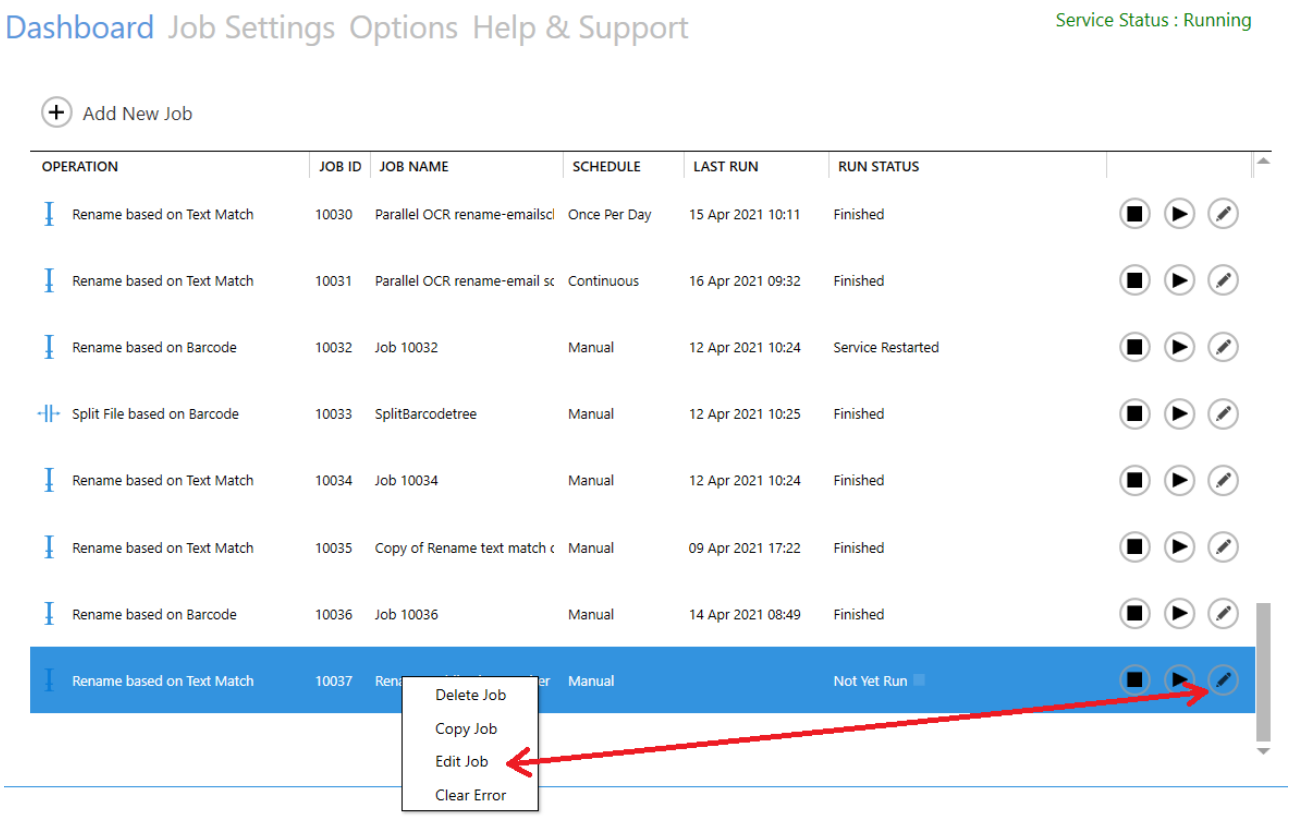
Error Code	Description
0	Job executed successfully
1	Error executing job. See log for details (also see section 13).
2	License related error. See log for details (also see section 13).
3	Server error: Contact support@aquaforest.com
4	Trigger file not present

4.2 Editing a job

To edit an existing Kingfisher job, you should go to the “Dashboard” tab.

From here there are four methods:

- Select the job and click on the “Job Settings” tab.
- Double click on the job you want to edit
- Click the edit button beside the Job name or by
- Right click the job and select “Edit Job” from the context menu.



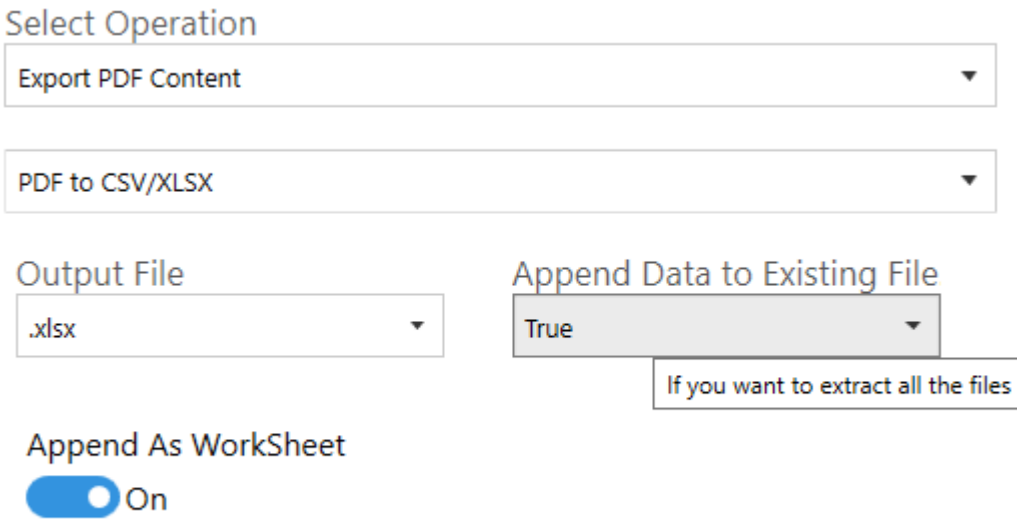
The “Job Settings” tab is explained in section 2.2.

After making changes to the job, click on the save button to update the Job condition file.

5 Extract Tabular Data From PDF

Aquaforest Kingfisher has the capability of recognizing tables in PDF files with very minimal user intervention. The tabular data it extracts is usually written to a csv or xlsx file.

5.1 Export to CSV/XLSX



Screen Field / Button	Description
Output file	csv: Produces a simple csv file. xlsx: Produces and Excel file.
Append Page Data to Existing File	If set to true, instead of overwriting the output file, Kingfisher will append the contents to the end of the current file.
Append as WorkSheet	When on, this will make Kingfisher add an extra worksheet for each PDF file, instead of appending the data at the end on the first Worksheet.

Aquaforest Kingfisher has the capability of recognizing tables in PDF files with very minimal user intervention. The tabular data it extracts is usually written to a csv or xlsx file.

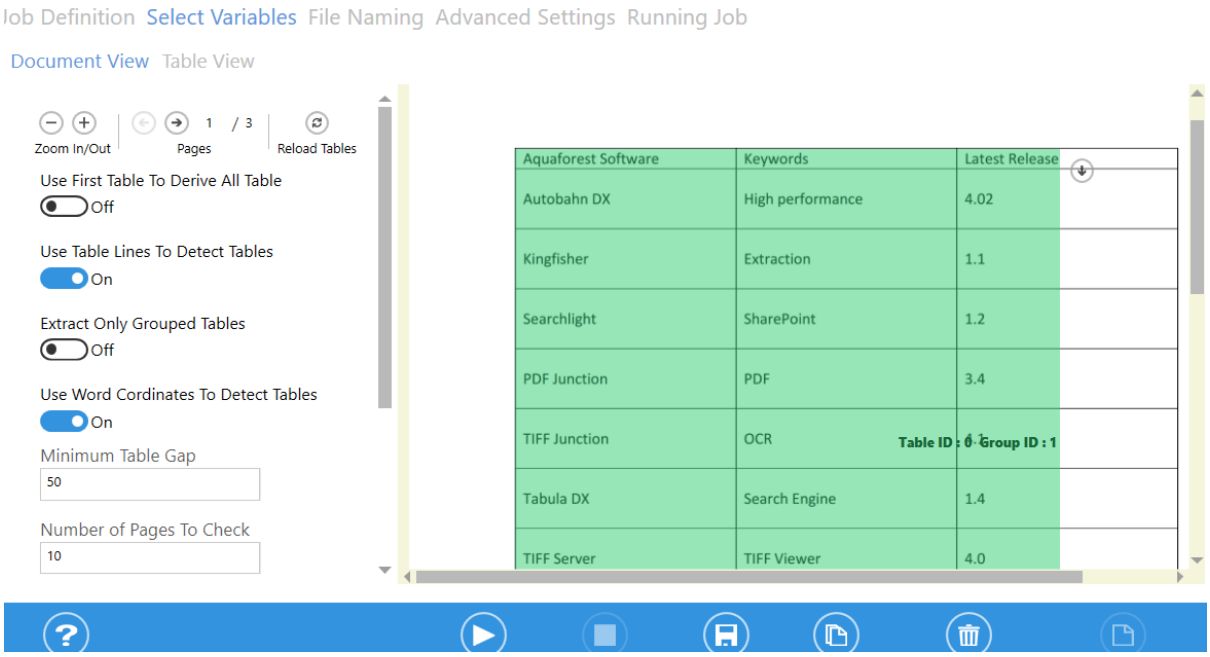
To extract all the tables recognized in the PDF file to the csv file you can just load the input file and run the job. If you want to tailor the extraction further and group the tables to your own liking, you can read the section below.

The Kingfisher Table control is divided into 2 tabs called the “Document View” and the “Table View”

5.1.1 The Document View

This tab shows a visual representation of the PDF file and the table recognized on the file, the table are usually highlighted. Below is a screen shot showing the Document View tab.

The document view usually shows the table recognized in the first 10 pages, this value can be changed as shown below. It also provides some properties that can be changed to alter the way the tables are recognized. These properties are described below.

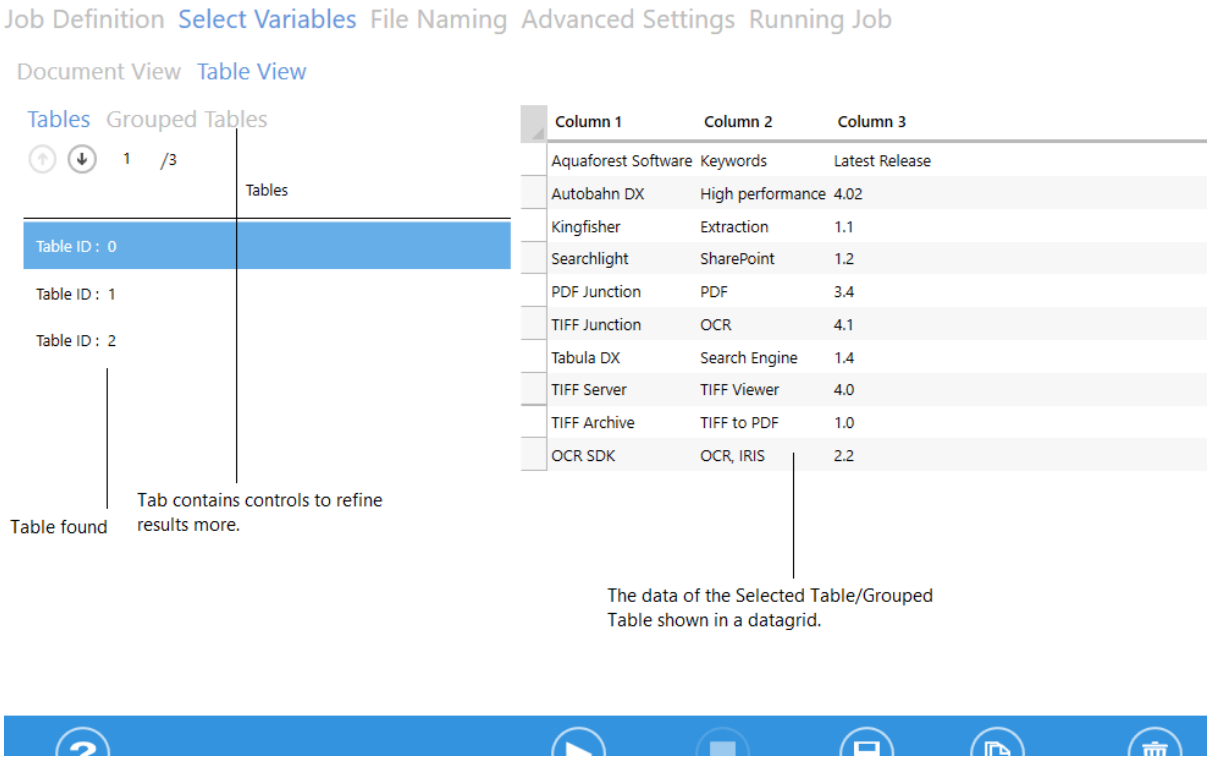


Value	Details
Zoom in/Out	You can use this to zoom in or zoom out of the file loaded in the table viewer control.
Pages	Use this control to move through the pages of the loaded file.
Reload Tables	Clicking this button will make Kingfisher reload the tables in the PDF file, this allows you to change some settings and reload the tables in the file after the file has been processed.
Crop Table	Switching this on will allow the user to specify an area where Kingfisher should check for tables, this method produces more accurate result because it eliminates any possible noise.
Do Not Write Header to Output	Switching this on will not write the Group Header to the output file, this is important if you want to write similar table data from multiple files to one output file, it will preserver continuity.
Use First Table to Derive All Tables	This setting instructs Kingfisher to use the first table in the file as a template to recognise all the other tables in the file.
Use Table Lines to Detect Tables	If set to true, Kingfisher will use the PDF graphic line to find the tables in the file.
Use Word Coordinates to Detect Tables	This setting instructs Kingfisher to use the coordinates of words on a page to recognise tables.
Extract Only Grouped Tables	Kingfisher gives the user the ability to group tables depending on different rules, turning this setting on will write only the tables grouped to the csv file.
Number of Pages to Check	The process of table recognition is CPU intensive; thus, Kingfisher limits the number of tables displayed in the GUI to the first 10 pages. You can increase or reduce it using the textbox.
Minimum Table Gap (pt)	If the space between two preceding lines is greater than the value given here, that space will be interpreted as a table break.
Table Space (pt)	If the space between two words is greater than this value, Kingfisher will recognise that space as a column. Note if the value is 0 or less, Kingfisher will ignore this and work out the space automatically.
Minimum Number of Rows	Kingfisher will only recognize tables with a higher number of rows than this value.
Minimum Number of Columns	Kingfisher will only recognize tables with a higher number of columns than this value.

5.1.2 Table View

The Table View shows the data that was successfully extracted from the PDF file, it is divided into two parts explained below:

- The controls on the left are divided into two tabs named “Tables” and “Grouped Tables”, the tabs. They contain a list of tables found and some other controls for grouping tables.
- The right part contains a Data grid which displays the data in the Table/Grouped Table Selected.



The next section will dive deeper into the “Table View”.

5.1.3 Tables

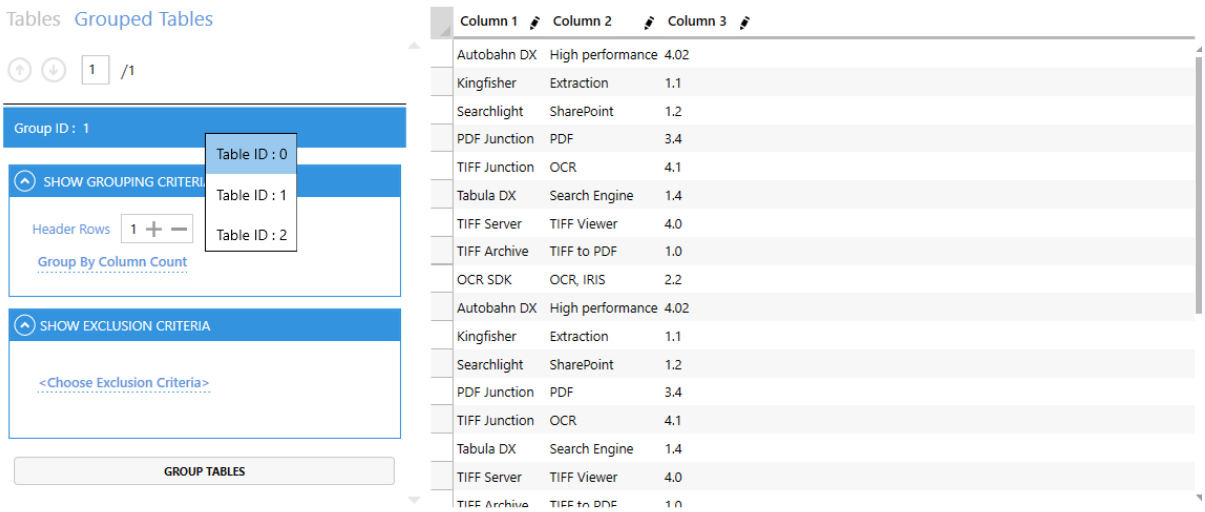
This is a very straight forward tab; it contains only two controls.

Value	Details
UP/Down Control	Use these buttons to select the next table in the list. Note the data grid on the right shows the contents of the selected table.
The table list	This contains a list of all the tables recognized in the document. Clicking on any one of them will load the selected table in the data grid on the right.

5.1.4 Grouped Tables

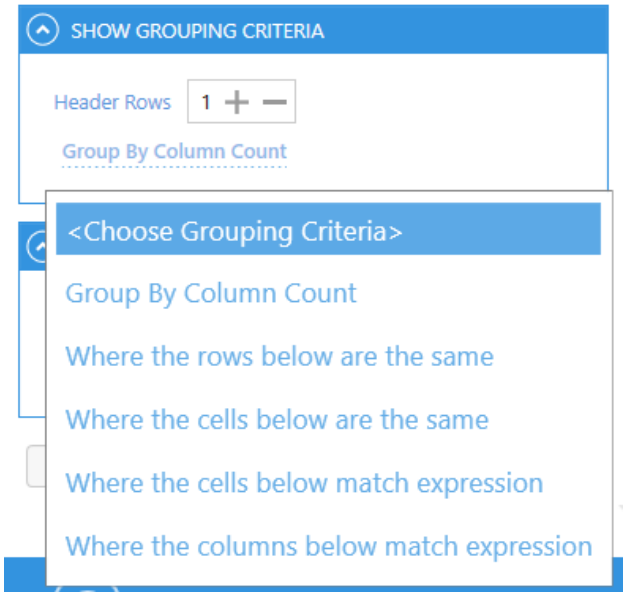
By default, Kingfisher can extract all structured data in a PDF file and write it to a ".csv" or ".xlsx" file. We also understand that users might not want to extract all the tables found in a document or the user might want to group some similar tables into one group with a common header.

Because of the issues raised in the paragraph above, we have decided to give the user the ability to group/exclude tables and edit the headers. The following sections will explain this in more depth.



5.1.4.1 Grouping Criteria

This section explains how users can group various tables based on the grouping criteria offered by Kingfisher.



- Header Rows: If you are writing a group of say 5 tables to a ".csv" file, the table header will be repeated for each table thus the final CSV file won't be very clean. As result, Kingfisher will skip the first "n" rows provided in the header rows control.

5.1.4.1.1 Group by Column Count

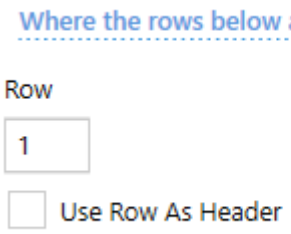
This is a straight forward option; it groups all the tables with the same number of columns. This option will also allow to change the provide the header rows (See 1.2.3).

5.1.4.1.2 Where the rows below are the same

This criterion will check all the tables for the row specified and group the tables based on identical rows. The example below will group all the tables with the same first row as one group.

Unfortunately, you can't save headers for this type of criteria, this is because the table types recognized here are not very predictable and will vary across different files. We decided to disable that feature to avoid putting wrong headers on tables.

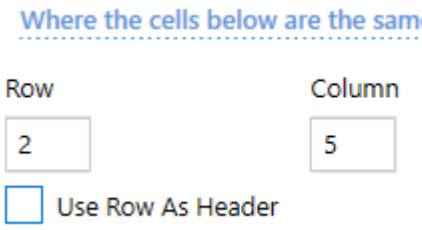
Checking the Use Row as Header Check Box will use the Row Matched as the header row.



5.1.4.1.3 Where the cells below are the same

This is the same as the previous option, the only difference is that it compares a single cell instead of the whole row. Unfortunately, you can't save headers for this type of criteria, this is because the table types recognized here are not very predictable and will vary across different files. We decided to disable that feature to avoid putting wrong headers on tables.

Checking the Use Row as Header Check Box will use the Row Matched as the header row.

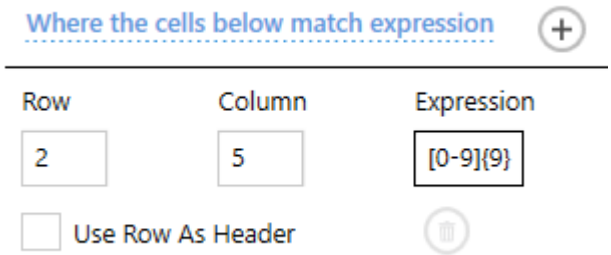


5.1.4.1.4 Where cells below match expression

This option groups all the tables where the cells provided below matches the Regular Expression given. Note, you can use the '+' button to add more items

This option will also allow to change the provide the header rows (See 1.2.3).

Checking the Use Row as Header Check Box will use the Row Matched as the header row.

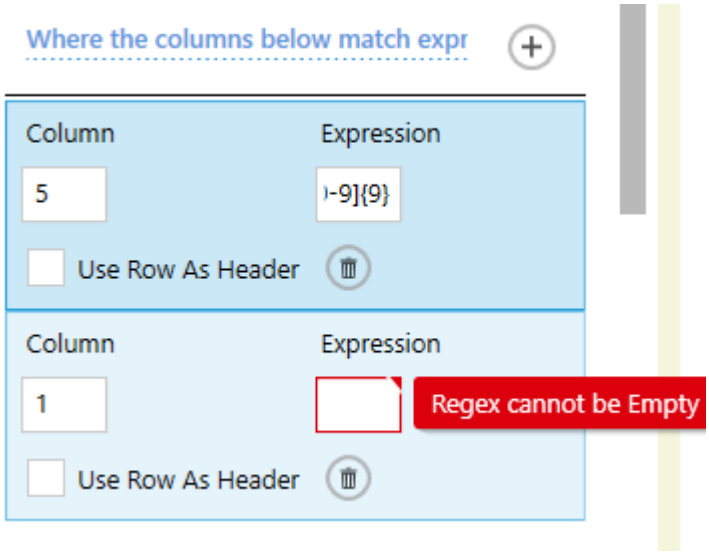


5.1.4.1.5 Where Columns below match expression

This option groups all the tables where the columns provided below matches the Regular Expression given. Note, you can use the '+' button to add more items

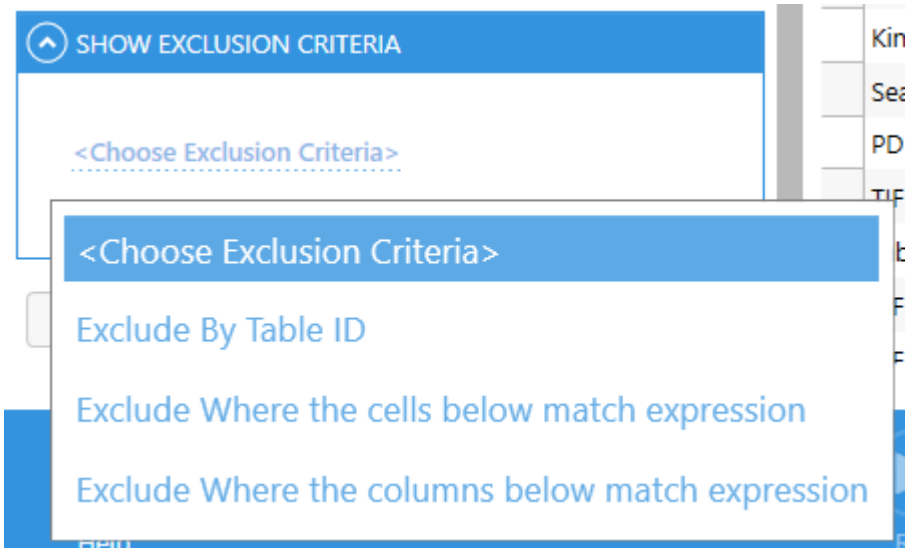
This option will also allow to change the provide the header rows (See 1.2.3).

Checking the Use Row as Header Check Box will use the Row Matched as the header row.



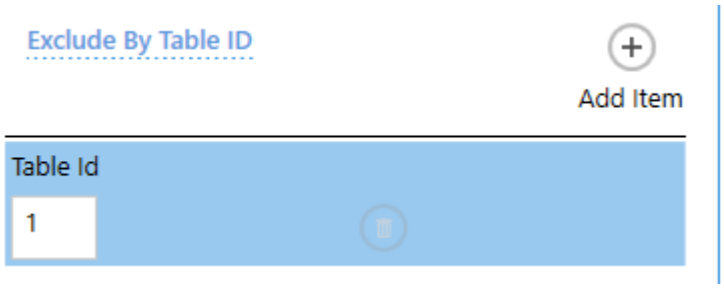
5.1.4.2 Exclusion criteria

Kingfisher also allows the Users to exclude table with characteristics they are not interested in, to use this feature you will have click on the "Show Exclusion Criteria" Expander to view the various option



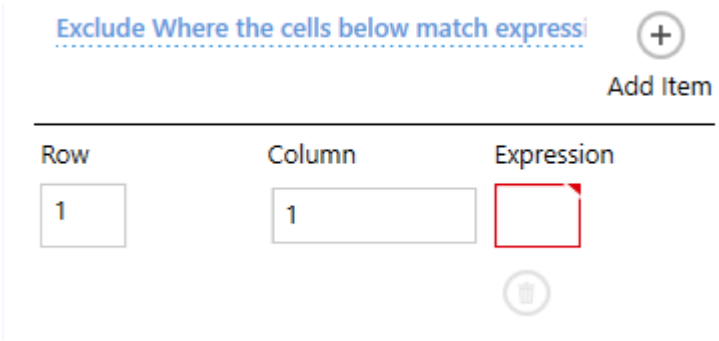
5.1.4.2.1 Exclude by Table ID

This option will exclude all the tables that their ID's are provided below.



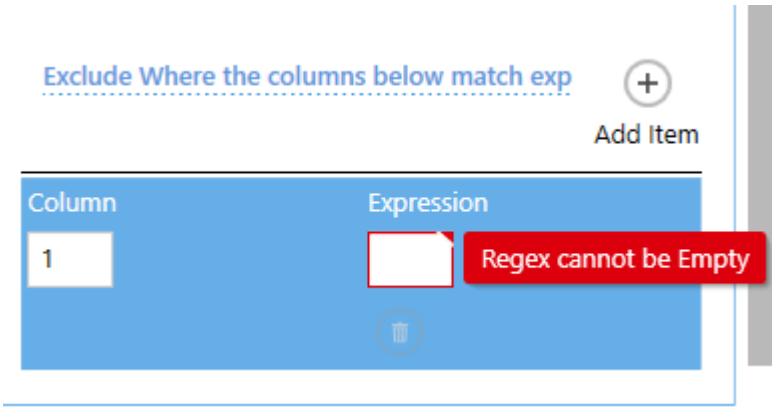
5.1.4.2.2 Exclude cells below match expression

This option excludes all the tables where the cell provided below matches the Regular Expression given. Note, you can use the '+' button to add more items



5.1.4.2.3 Exclude Where Columns below match expression

This option excludes all the tables where the columns provided below matches the Regular Expression given. Note, you can use the '+' button to add more items.



5.1.5 Data grid

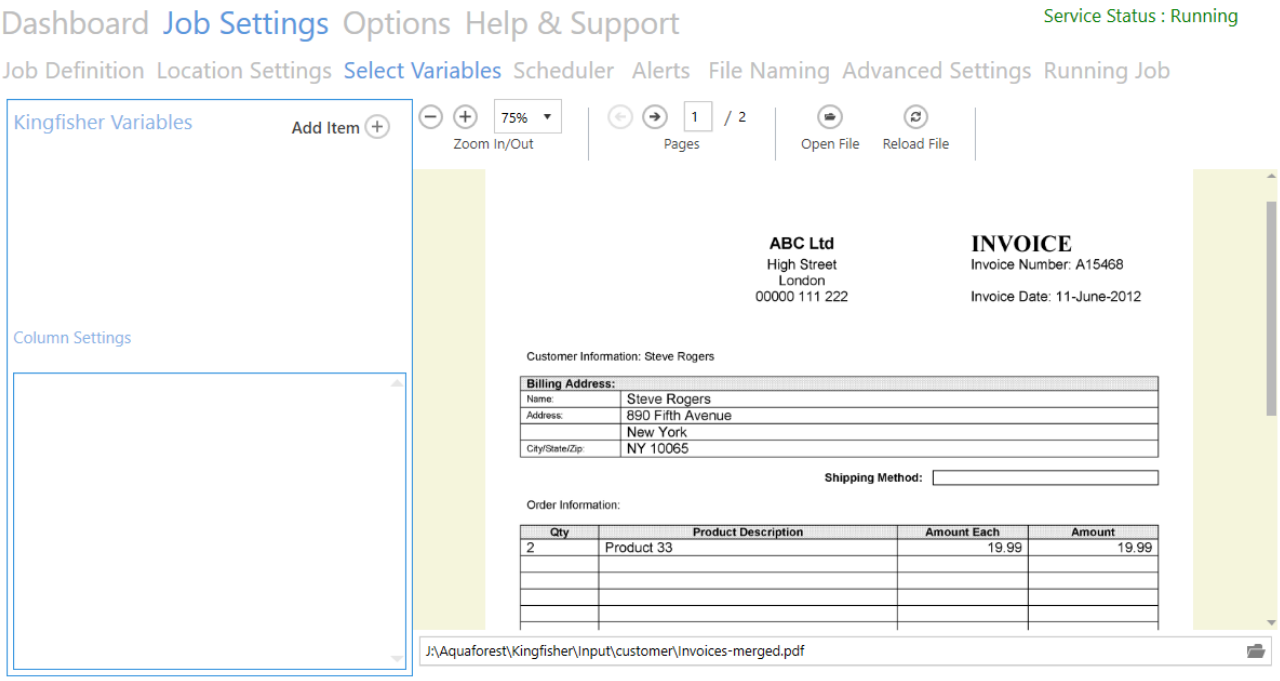
The main use of the data grid is to show the users the data that has been extracted from the PDF file, you can also use it to Edit headers (If the pen symbol appears next to the header) as shown Below. After editing, just click the save job button.

AQUAFOREST SOFTWARE	KEYWORDS	Column 3
Autobahn DX	High performance	4.02
Kingfisher	Extraction	1.1
Searchlight	SharePoint	1.2

5.2 Advanced Export to csv/xlsx

Advanced Export allows the export of areas on the PDF page to CSV or XLSX files.

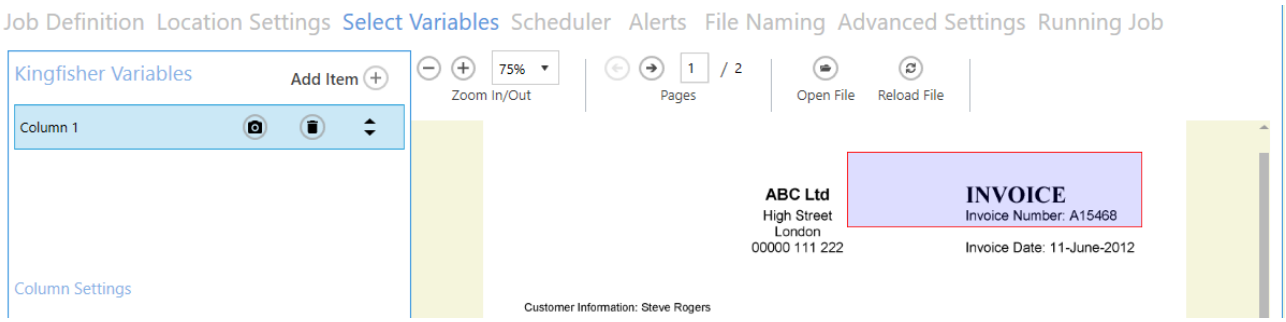
For an Advanced Export job, the Select Variables tab looks like this.



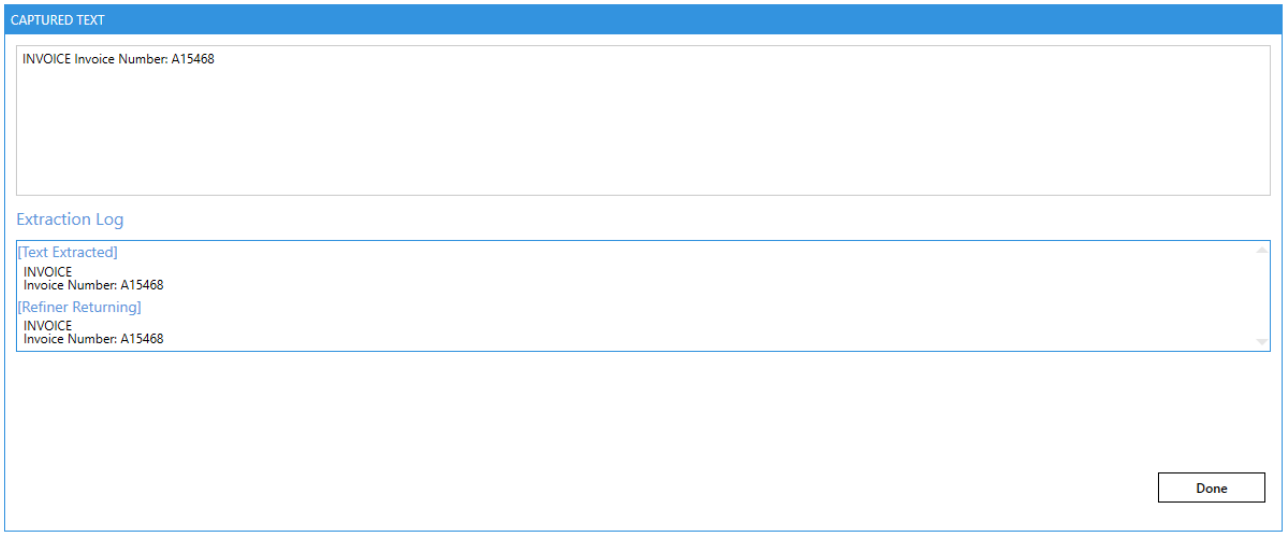
The file displayed is either the one selected in the Location Settings tab or can be selected by clicking on the Open File button.

If there is no Kingfisher Variable, click on the Add Item button.

Select the area on the displayed file containing the information required.

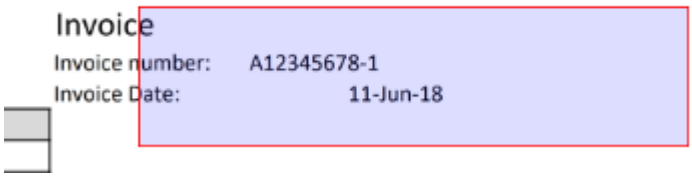


Click on the camera icon to capture the text.



The extracted text contains more than the actual invoice number, this is because the area is designed to cover slightly different format invoices:

This is the same area on the invoice that makes up the second page of the example document.



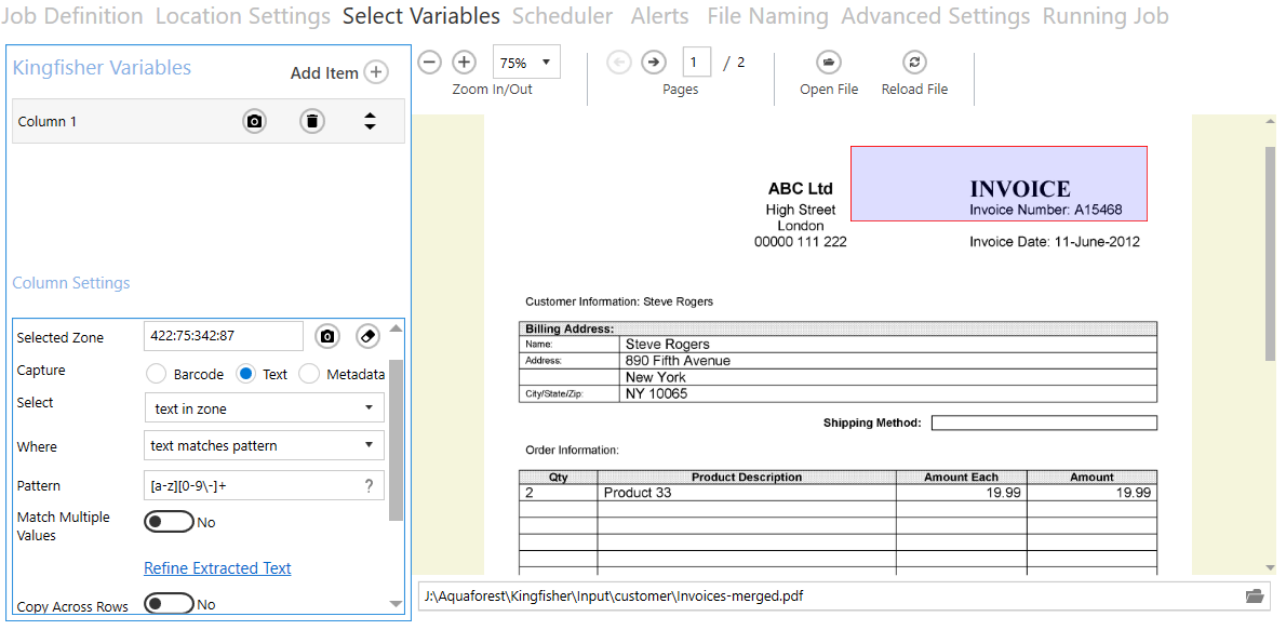
Click Done.

The text needs to be refined.

For this example file, the invoice number is of the (regular expression) format [a-z][0-9\-]+.

The first part selects a string starting with the alphabetic characters A-Z. The second part is one or more numeric characters or a hyphen.

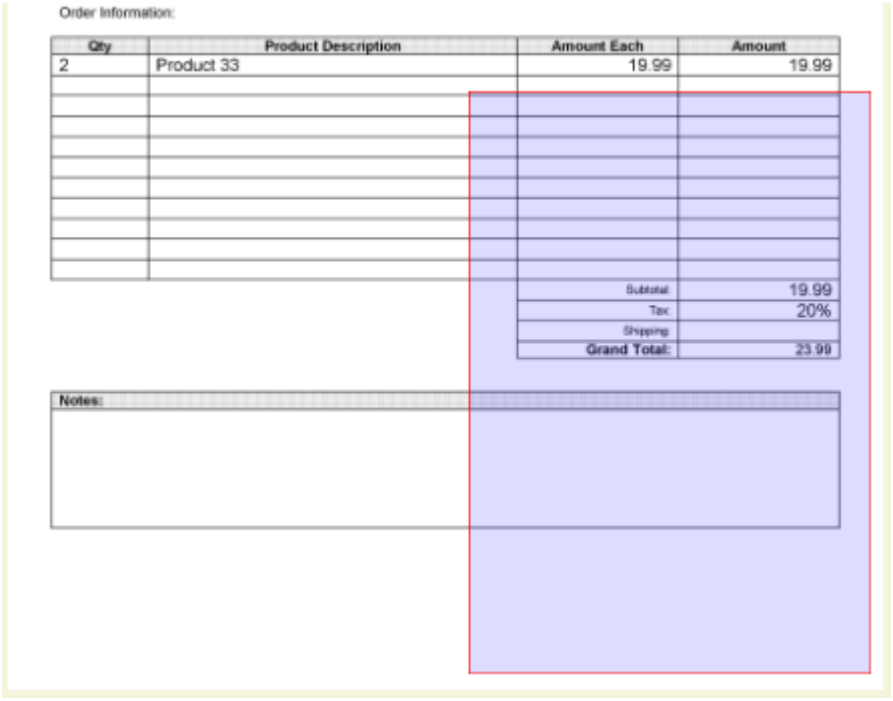
This is added to the column settings by choosing Select “text in zone”, Where “text matches pattern” and entering the pattern in the box. There are tips on the regular expressions available by clicking on the “?” button.



Check the extracted text using the camera icon in the column settings.

Add another item.

Select an area that covers the Grand Total on both pages.

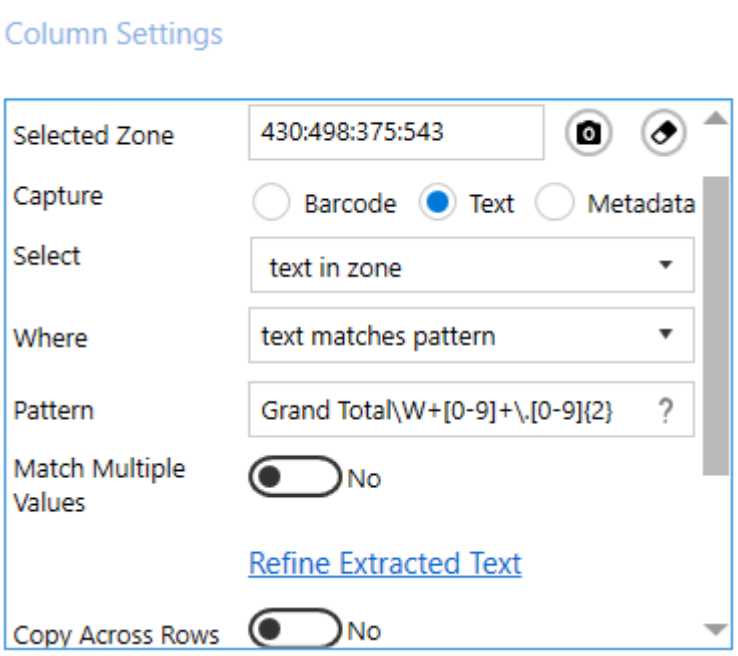


Click on the camera icon on Column 3 to select the area.

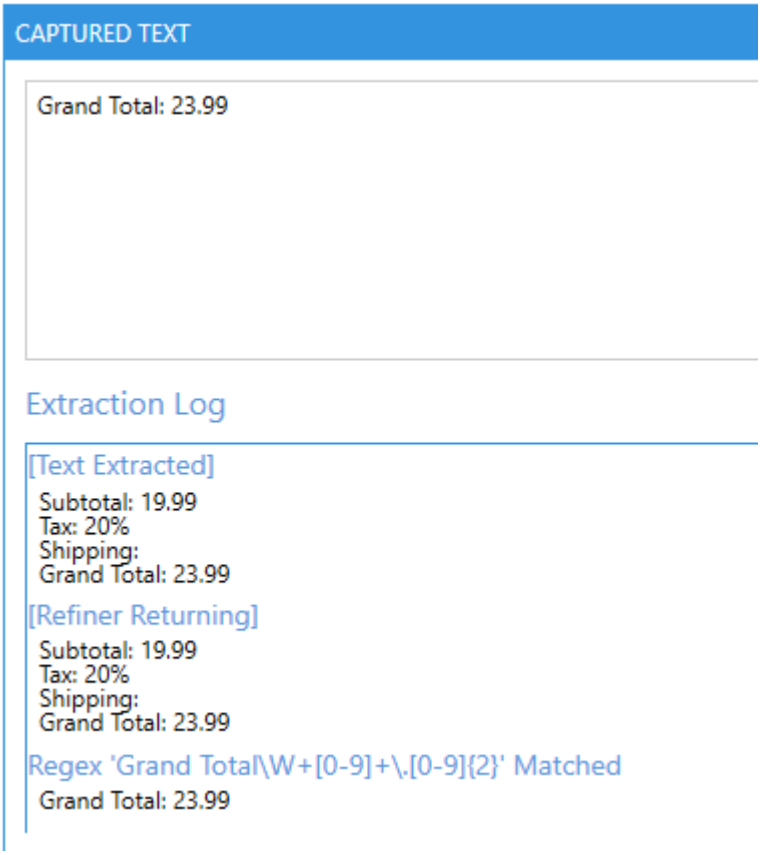
Check that the text contains the value you are after.

Click on Done.

The pattern for this selection is more complicated. The literal text “Grand Total” identifies the beginning of the selection. Next there is one or more whitespace (space or punctuation) characters followed by one or more digits, a decimal point then two digits.

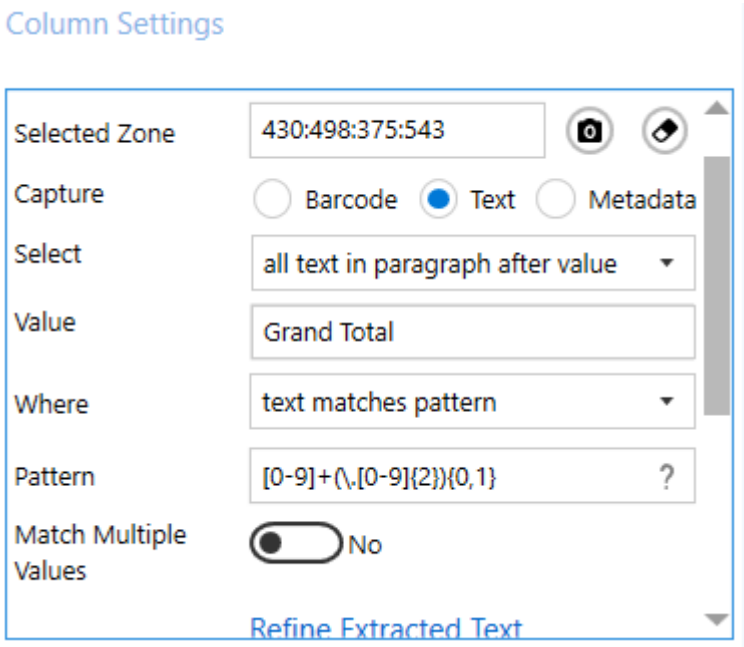


Click on the Column Settings camera icon to see the extracted text.

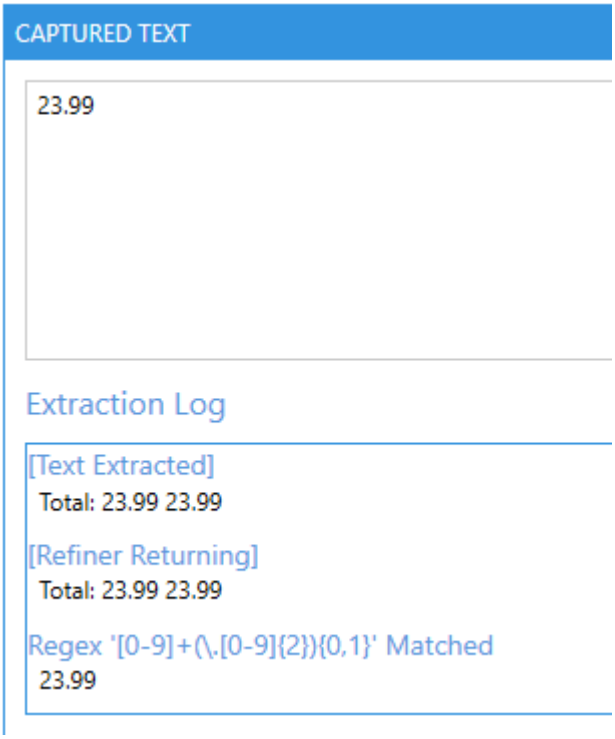


The next task is to refine this text further by removing everything before the “Grand Total” part.

Select “all text in paragraph after value” “Grand Total” where “text matches pattern” one or more digits optionally followed by a decimal point and two digits.

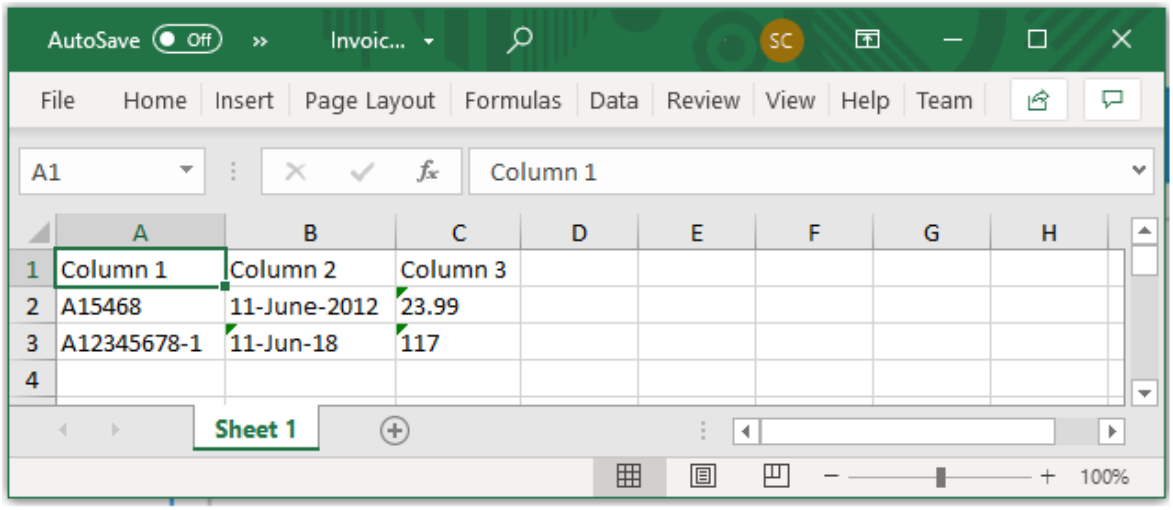


Click on the Column Settings camera icon to see the extracted text.



Save the job and run it.

The output file will contain:



6 Custom Scripts

In Kingfisher, the user can execute custom scripts after each file is successfully processed, this gives the user a chance to do some custom post processing on the file input or output files. This gives the user the power to perform any type of operations that are not available in Kingfisher.

We use **Windows Script Host** (specifically the command line utility CScript) to execute these scripts, below is an example of the command line generated to execute a custom script.

If the Debug option is set on Job Settings you will be able to see the arguments.

```
cscript  "C:\Aquaforest\Kingfisher\Custom\testarguments.wsf"
"/Newfilename:00000006.pdf" "/Oldfilepath:C:\Aquaforest\Kingfisher\In\invoice_6.pdf"
"/TempOutputFile:C:\Aquaforest\Kingfisher\temp\jobs\10012\bb3e649b-6cda-4de4-bd45-
c69d12f23227\target\00000006.pdf"
"/TempTarget:C:\Aquaforest\Kingfisher\temp\jobs\10012\bb3e649b-6cda-4de4-bd45-
c69d12f23227\target" "/FinalTarget:C:\Aquaforest\Kingfisher\Out\Split\Text Change"
"/JobSource:C:\Aquaforest\Kingfisher\In" "/LocationType:file" "/value1:00000006"
```

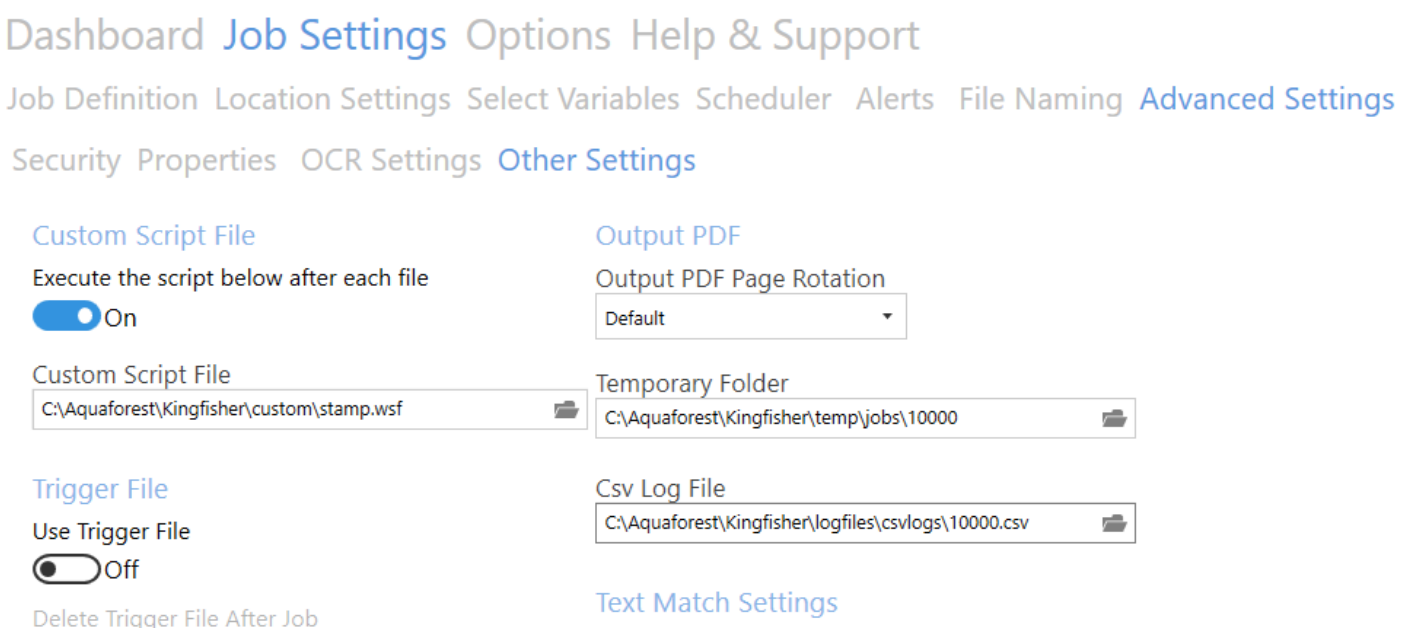
6.1 Custom Script Arguments

Value	Description
Newfilename	The file name generated by Kingfisher. This is the name, so does not contain the whole path.
Oldffilepath	The input file path, in a SharePoint step this will the temporary download location.
`TempOutputFile`	This is the temporary output file. It is the same as concatenating the **TempTarget** and **Newfilename** variables above.
TempTarget	This is the temporary folder where Kingfisher stores all the target files until the job is finished, this is where you will get the output file. **i.e. outputfilepath= TempTarget+ Newfilename**
FinalTarget	This is the destination path.
JobSource	This is the source path that the user provided when setting up the job, it can be a file name, a folder name or a tree name.
LocationType	Source location type: File, Folder or Tree
value1	This represents the “%VALUE1%” selected variable
value2	This represents the “%VALUE2%” selected variable (if found)
value3	This represents the “%VALUE3%” selected variable (if found)
value4	This represents the “%VALUE4%” selected variable (if found)

6.2 Enabling Custom Scripts

To enable custom scripts in a job you should

- Go to **Job Settings → Advance Settings → Other settings.**
- Switch the "Execute the script below after each file" switch on.
- Enter or paste a valid ".wsf" file path in the "Custom Script File" text box or use the file browser to select the file.



The settings above will execute a script named "C:\Aquaforest\Kingfisher\custom\stamp.wsf" after each file.

6.3 Custom Scripts Samples

There are four sample scripts in a zip file in the <installation folder>\docs folder called custom.zip.

Extract these to a suitable location (such as <installation folder>\Custom).

6.3.1 testarguments.wsf

This just lists the named arguments supplied by Kingfisher.

Using the example call above results in the following output:

```
Microsoft (R) Windows Script Host Version 5.812
Copyright (C) Microsoft Corporation. All rights reserved.

Named arguments
Newfilename has value 00000006.pdf
Oldfilepath has value C:\Aquaforest\Kingfisher\In\invoice_6.pdf
TempOutputFile has value C:\Aquaforest\Kingfisher\temp\jobs\10012\bb3e649b-6cda-4de4-bd45-c69d12f23227\target\00000006.p
df
TempTarget has value C:\Aquaforest\Kingfisher\temp\jobs\10012\bb3e649b-6cda-4de4-bd45-c69d12f23227\target
FinalTarget has value C:\Aquaforest\Kingfisher\Out\Split\Text Change
JobSource has value C:\Aquaforest\Kingfisher\In
LocationType has value file
value1 has value 00000006
```

Included (but commented out) is code to list all arguments supplied.

This sample is useful to check the call to the script from Kingfisher and assist in the development of additional custom scripts.

6.3.2 archiveinputs.wsf

This script copies (or moves) the input file to an archive location.

Please ensure you fully understand and have tested this process if you choose to use the move option.

This has been tested with the file system job only.

Open the file with a text editor of your choice and apply the following changes to use the script.

- Line 14: var ArchiveTarget = "%ARCHIVELOACTION%";

Change this to your archive location.

- Line 15: var move = "false";

Set to true to move the file instead of copying it.

This script returns certain exit codes when executed:

Exit Code	Description
0	Success
1	Destination folder not found
2	Error in process
3	Source file not found

This is a revised script. The previous script is also supplied.

6.3.3 stamp.wsf

This script applies a stamp to the first page of the output PDF file, this stamp can contain the values extracted from the file. Open the file with a text editor of your choice and apply the following changes to use the script.

Line 18: var stampExe = "C:\\Aquaforest\\Kingfisher\\custom\\stamp\\stamp.exe". Provide the path to the stamp.exe file, this can be found in the “stamp” subfolder of the custom.zip file.

Line 19: var stampText = "Processed by %value1%". This is the text that will be stamped on the PDF file, the script will replace the “%valuex%” in the string with the values extracted from the document by Kingfisher. E.g. if the text extract from the file is “Admin Department”, the stamp will be “Processed by Admin Department”

6.3.4 converttoxml.wsf

This script converts a CSV output file from Kingfisher to an XML and copies the file to the target directory. This has been tested with the file system job only. Open the file with a text editor of your choice and apply the following changes to use the script.

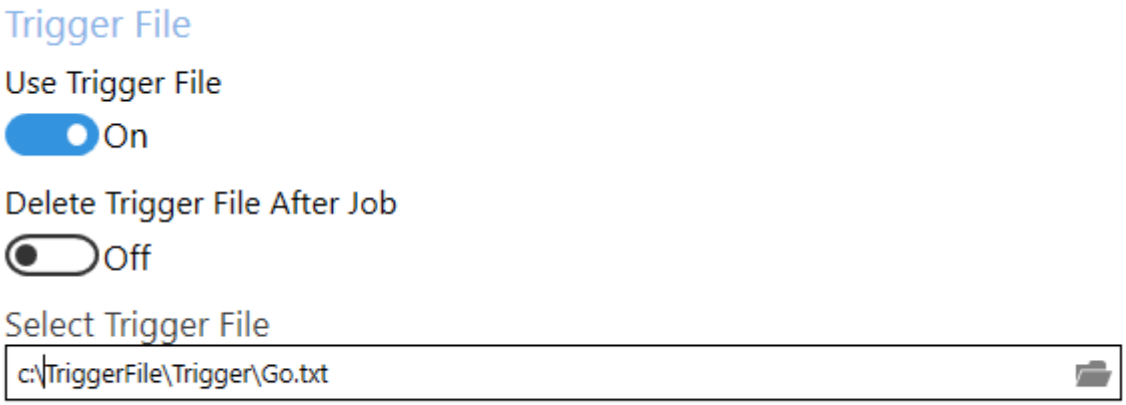
- Line 14: var XmlRoot = "books". This represents the root xml node.
- Line 15: var XmlElement = "book". This is the name of the node that represents each record in the CSV.
- Line 16: var HasColumnNames="true". This will use the first row of the CSV file as the names of the xml nodes that represent the CSV Records. If you use false, the xml nodes will be named column1,column2 etc.
- Line 17 var csv2XmlExe = "C:\\Aquaforest\\Kingfisher\\custom\\csvtoxml\\CSV2XML.exe". Provide the path to the CSV2XML.exe file, this can be found in the “csvtoxml” subfolder of the custom.zip file.

7 Trigger Files

There are Use Cases where files need to be processed together, or only after a specific file has been supplied.

The Trigger File option is found at Job Settings/Advanced Settings/Other Settings.

Set the Use Trigger File to on and paste/type or browse to your trigger file for this job.



Value	Details
Use Trigger File	Set to On to hold the job until the trigger file is found
Delete Trigger File After Job	Once the job has completed, delete the trigger file
Select Trigger File	Full path to the trigger file.

8 Kingfisher Configuration File (Kingfisher.Config)

Kingfisher stores common setting in the Kingfisher.config file located in the [installation folder]\bin folder. Most of these entries are controlled by the UI, as such they should not be edited in the file.

Field	Description
<add key="culture" value="en-US" />	Culture for default number and time/date format
<add key="Version" value="2.5.2104.22.1023" />	Kingfisher version
<add key="ShowWelcome" value="False" />	Display welcome screen when starting the UI
<add key="showHelp" value="false" />	Show tool tips
<add key="SearchableThreshold" value="50" />	Threshold defining when a document is searchable
<add key="Morph" value="c2.2" />	Contact support@aquaforest.com for details of use
<add key="OCRDPI" value="300" />	DPI of TIFF in Convert to Tiff OCR
<add key="UIThreads" value="4" />	Contact support@aquaforest.com for details of use
<add key="LineNoise" value="0" />	For text extraction contact support@aquaforest.com for details of use
<add key="UseHocrForText" value="true" />	For text extraction contact support@aquaforest.com for details of use
<add key="LineSpace" value="2" />	For text extraction contact support@aquaforest.com for details of use
<add key="BarcodeDPI" value="200" />	Default Dots Per Inch for barcode rasterization (overridable at job level)
<add key="PdfToImageBpp" value="1" />	Default Bits per pixel (default is 8) for barcode (overridable at job level)
<add key="BarcodeMultipleRead" value="true" />	Set the option to multiple read bar codes to improve accuracy
<add key="GetBarcodeDirection" value="true" />	Set to true to get bar code direction
<add key="GetBarcodeBoundingRectangle" value="true" />	Set to true to get the rectangle around the bar code
<add key="GetBarcodeScore" value="true" />	Set to true to get the confidence score for the bar code
<add key="GetBarcodeType" value="true" />	Set to true to get the bar code type
<add key="SpaceWidth" value="10" />	For text extraction contact support@aquaforest.com for details of use
<add key="DateFormat" value="d-MMM-yyyy" />	Date format
<add key="TimeFormat" value="HH-mm-ss" />	Time format
<add key="LogUI" value="false" />	Log UI actions (contact support@aquaforest.com for more information)
<add key="LogPath" value="..\logfiles\GeneralLogs" />	Relative location of the general log folder
<add key="EmailServer" value="" />	Email server (SMTP)
<add key="EmailPortNumber" value="" />	Email port number (SMTP)
<add key="EmailUsername" value="" />	Email server (SMTP)
<add key="EmailPassword" value="" />	Encrypted password for the email user (SMTP)
<add key="ProcessThrottleTimout" value="180" />	Process Throttle time out (total)
<add key="ProcessThrottleInterval" value="10" />	Process Throttle check interval
<add key="NetworkPathTimout" value="1" />	Network timeout period when checking for file or folder
<add key="AzureADInstance" value="" />	Azure Active Directory authentication end point (OAuth2)
<add key="AzureClientSecret" value="" />	Encrypted Azure Client Secret (OAuth2 - client secret)
<add key="AzureTenant" value="" />	Azure Tennant ID (OAuth2)
<add key="AzureClientId" value="" />	Azure Client ID (OAuth2)
<add key="CertificatePassword" value="" />	Encrypted certificate password (Oauth2 - certificate)
<add key="OAuth2CredentialType" value="" />	OAuth2 Credential type (Certificate/Client Secret)
<add key="OAuth2Client" value="" />	OAuth2 Client (Azure)
<add key="GooglePrivateKey" value="" />	For future use
<add key="GoogleServiceAccountEmail" value="" />	For future use
<add key="CertificateFilePath" value="" />	OAuth2 certificate file path
<add key="AuthenticationMode" value="" />	Email Authentication mode (SMTP/OAuth2)
<add key="ScheduleOncePerDayContinueAfterError" value="false" />	Set value to true to continue after scheduled run error
<add key="ScheduleContinuousContinueAfterError" value="false" />	Set value to true to continue after scheduled run error

9 XML Job Condition Files

Any job created in Kingfisher is saved in a job condition file, these are xml files usually created in the conditions folder. These files can be used to execute jobs from the Command line using the –JobFile parameter.

Developers may wish to create, edit or generate XML job files using the information below.

jobspecialoptions	**Set of options for the specified operation. Described in detail in section 6.2.1 and 6.2.2**
outputnamemask	**The Name Template explained in section 2.2.6**
usenametemplate	**If the value is true, the Name Template will be used instead of the script.**
Metadata	**Metadata settings for result PDF files. By default, metadata fields are left blank. Described in detail in section 2.2.7.3**
Security	**Security settings for result PDF files. By default, files are not secured. Described in detail in section 2.2.7.1**
docoptions	**Document option settings for result PDF files. Described in detail in section 2.2.7.2.1**
Sourcefiles	**Source PDF file, folder or tree. May be multiple files for merge operations.**
advancedflags	**Used to pass extra parameters, do not use unless directed by Aquaforest Technical team.**
logfile	**Each run of a job is output will be logged to a file based on this template. See section 2.2.2 for details. The filename by default uses the %TIMESTAMP% variable. On completion of a run, the file referenced by kingfisherlivelog is copied to a new file based on this logfile template**
kingfisherlivelog	**When a job is running, output is written to this file. At the end of a job, the logfile template is used as a destination for a copy.**
textextraction	**This contains the definitions of text to be extracted .**
forceocr	**When set to true, any file processed will be OCRd (see the OCR options for other related settings) irrespective of whether it would normally be considered searchable.**

10 Using Kingfisher from the command line

Kingfisher can be accessed via the command line to run rules that have already been created in the GUI. You can also run Kingfisher directly from the command line.

The executable used for running Kingfisher from the command line is “Kingfisher.Commandline.exe”

10.1 Using Job Definition File

This is the preferred and safest method to run Kingfisher from the command line, this method ensures you set up the job correctly and you are sure of what it does.

If you create a Job in Kingfisher, an xml Job Definition file is automatically created in the “conditions” folder in the installation Directory.

Note: When using the command line option, the xml file can actually be stored in a different location if required. e.g. c:\mydocuments\1001.xml. Thus, a working copy of definitions can be stored in a different (possibly less secure) location to the original.

Usage:

Kingfisher.Commandline.exe -jobfile=“..\\conditions\\1001.xml”

11 Autobahn DX Integration

Autobahn DX is a Workflow orientated Document Processing product also available from Aquaforest.

It can process and convert a variety of different types of document including TIFF images, Microsoft Office documents, email messages and HTML pages, convert those documents to searchable PDFs and perform a variety of other customizable activities. For more information on Autobahn DX, check the product page by clicking the link below.

https://www.aquaforest.com/en/autobahn.asp

A Kingfisher Job can be used as a step in Autobahn DX, for more details, see the blog post below.

https://blog.aquaforest.com/adding-a-kingfisher-job-as-an-autobahn-dx-step/

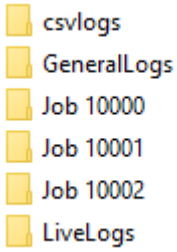
12 Kingfisher Files and Directories

After installation, the following subdirectories are created in the Kingfisher install directory:

Directory	Contents
bin	Executables and Scripts.
conditions	This folder contains the job condition files (XML format).
docs	Contains Reference Guide and release notes
license	Contains the license key file.
logfiles	Log files.
regex	This folder contains files that can be used to store predefined regular expressions.
samples	This folder contains sample files used in the sample jobs.
temp	This folder contains temporary folders and files used during processing.
theme	This folder contains the current theme information file.
status	Contains status file, for progress reports.
video	Contains introductory and use case videos.

12.1 Log Files

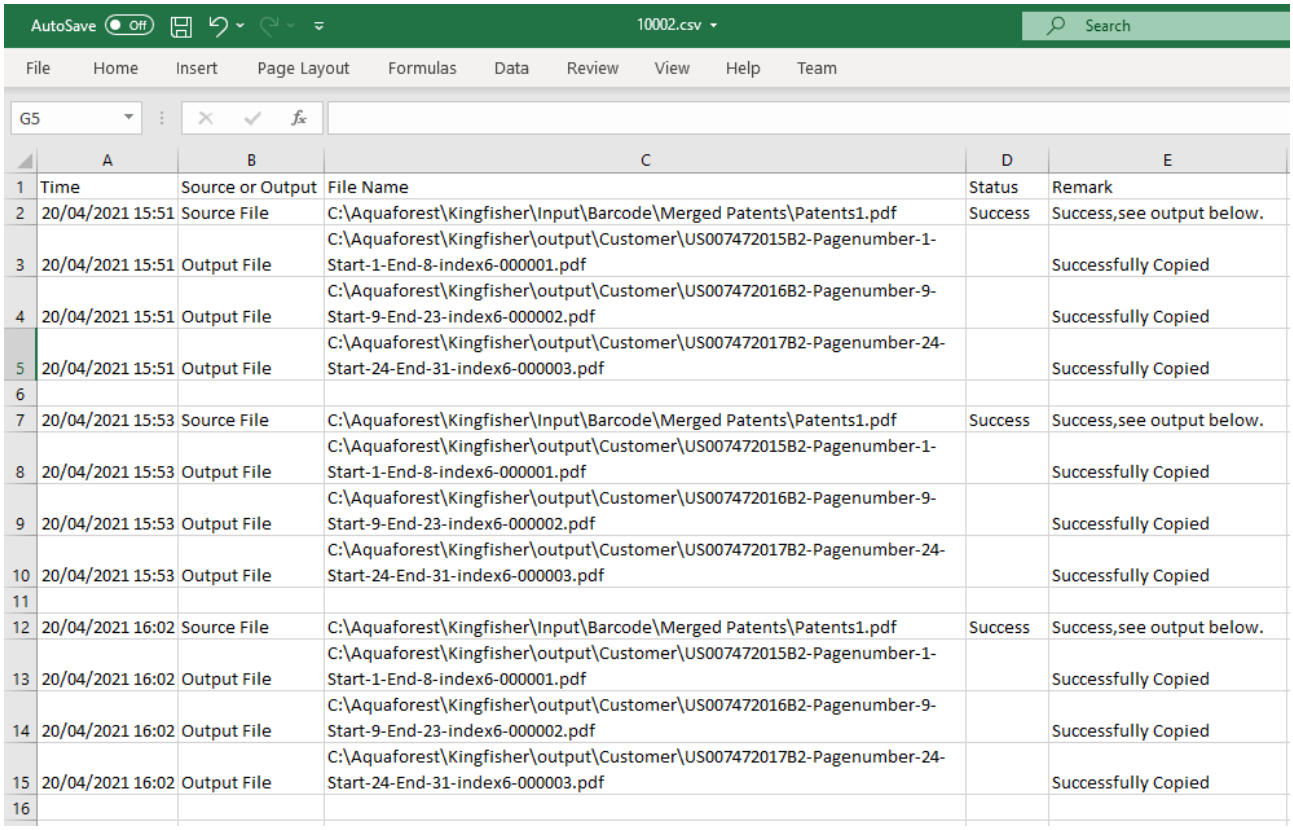
The log file folder is the location for the various log files generated by Kingfisher.



In this example, there is the folder for the CSV Logs, a folder for general logs, three folders for previous runs of the three jobs (10000, 10001 and 10002) and the folder containing the last logs of the jobs.

12.1.1 CSV logs

The CSV log file contains a list of the source files and the resultant output files. By default each job has a csv log file of the form <jobid>.csv. This can be changed manually – see section 2.2.2.



AutoSave Off | 10002.csv | Search

File Home Insert Page Layout Formulas Data Review View Help Team

G5

	A	B	C	D	E
1	Time	Source or Output	File Name	Status	Remark
2	20/04/2021 15:51	Source File	C:\Aquaforest\Kingfisher\Input\Barcode\Merged Patents\Patents1.pdf	Success	Success,see output below.
3	20/04/2021 15:51	Output File	C:\Aquaforest\Kingfisher\output\Customer\US007472015B2-Pagenumber-1-Start-1-End-8-index6-000001.pdf		Successfully Copied
4	20/04/2021 15:51	Output File	C:\Aquaforest\Kingfisher\output\Customer\US007472016B2-Pagenumber-9-Start-9-End-23-index6-000002.pdf		Successfully Copied
5	20/04/2021 15:51	Output File	C:\Aquaforest\Kingfisher\output\Customer\US007472017B2-Pagenumber-24-Start-24-End-31-index6-000003.pdf		Successfully Copied
6					
7	20/04/2021 15:53	Source File	C:\Aquaforest\Kingfisher\Input\Barcode\Merged Patents\Patents1.pdf	Success	Success,see output below.
8	20/04/2021 15:53	Output File	C:\Aquaforest\Kingfisher\output\Customer\US007472015B2-Pagenumber-1-Start-1-End-8-index6-000001.pdf		Successfully Copied
9	20/04/2021 15:53	Output File	C:\Aquaforest\Kingfisher\output\Customer\US007472016B2-Pagenumber-9-Start-9-End-23-index6-000002.pdf		Successfully Copied
10	20/04/2021 15:53	Output File	C:\Aquaforest\Kingfisher\output\Customer\US007472017B2-Pagenumber-24-Start-24-End-31-index6-000003.pdf		Successfully Copied
11					
12	20/04/2021 16:02	Source File	C:\Aquaforest\Kingfisher\Input\Barcode\Merged Patents\Patents1.pdf	Success	Success,see output below.
13	20/04/2021 16:02	Output File	C:\Aquaforest\Kingfisher\output\Customer\US007472015B2-Pagenumber-1-Start-1-End-8-index6-000001.pdf		Successfully Copied
14	20/04/2021 16:02	Output File	C:\Aquaforest\Kingfisher\output\Customer\US007472016B2-Pagenumber-9-Start-9-End-23-index6-000002.pdf		Successfully Copied
15	20/04/2021 16:02	Output File	C:\Aquaforest\Kingfisher\output\Customer\US007472017B2-Pagenumber-24-Start-24-End-31-index6-000003.pdf		Successfully Copied
16					

13 Support

Kingfisher has been tested on a diverse selection of PDF files as well as a range of hardware and operating systems.

https://www.aquaforest.com/en/kingfisher.asp#techspec

In the event of a problem, the first action should be to try to process the affected job (or files) with the Debug flag set to Yes (see 2.2.1). This will produce additional information that may assist you in solving the issue.

If you need additional help, please email support@aquaforest.com . It will help us if you include the log file (with the Debug flag set), condition file and sample files that are causing the problem.

14 Acknowledgements

This product makes use of many Open-Source components which are included in binary form. The appropriate acknowledgements and copyright notices are given below.

Name	Homepage
BitMiracle.LibTiff.NET	Homepage | GitHub
BouncyCastle.Crypto	Homepage
Camelot	Homepage
Cuneiform	n/a (Copyright (c) 1993-2008, Cognitive Technologies)
FreeImage.NET	Homepage
IKVM.NET	Homepage | Sourceforge
Leptonica	Homepage
Libjpeg	Homepage
Libpng	Homepage
Libtiff	Homepage
MahApps MahApps.Metro MahApps.Metro.IconPacks	Homepage GitHub GitHub
Newtonsoft.Json	Homepage | GitHub
PDFBox	Homepage
PDFMiner	GitHub
veraPDF	Homepage | GitHub
Zlib	Homepage