

KOFAX



Kofax OmniPage Server

Folder Watcher

Folder Configuration Editor

User Guide

Contents

- Introduction 3
 - Concepts, Terms and Abbreviations 3
 - File Conversion 3
 - Folder Structure 3
 - Security 4
- Folder Configuration Editor 5
 - Function of Folder Configuration Editor 6
 - Conversion Parameters 6
- Authentication 8
 - Domain environment 8
 - Workgroup environment 8
- Settings 8
- Workflow Job Type 9
 - Workflow job type 9
 - Creation Workflow 9
 - Use Workflow 10
 - Input file mapping 10
- Document Classification job type 11
 - Perform Document Classification 11
- Tracking jobs lifecycle 11
- Troubleshooting 12
 - Folder access issue 12
 - Service connection issue 12

Introduction

This document describes the Folder Configuration Editor client of the Folder Watcher

OmniPage Server (OPS) is a solution that provides document conversion services.

OPS is an application service solution built using conventional enterprise architectural components and approaches, while still offering a high level of availability and scaling functionality.

The folder watcher system is developed for companies to provide its users with the advantage of converting their private documents through Windows-only file system usage.

Concepts, Terms and Abbreviations

The basic concept is that the folder watcher service can perceive file changes in a specified folder and send those to the conversion server after reading them. At the completion of the conversion process the new documents are downloaded to a specified subfolder of the original watched folder.

File Conversion

File conversion is the process of transforming information represented in a source file as specified by its file format into another file using a representation specified by another file format.

Note: File conversion might result in information loss if the destination file format cannot represent everything represented by the source file format.

Document Conversion

In this document, document conversion refers to the process of transforming a set of input documents into a set of output documents. The transformation is typically file conversion, OCR or both.

Folder Structure

The folder watcher system works on a predefined folder structure. The system administrator is responsible for building a watched folder structure before the watching service is started. During file system watching, the specified system directories, for example, error folder or working directories are managed by the system. These are created and deleted as necessary.

The system can read convertible files from watched folder and send those to the conversion server. After document conversion, the converted files are downloaded to a target folder directly under the watched folder.

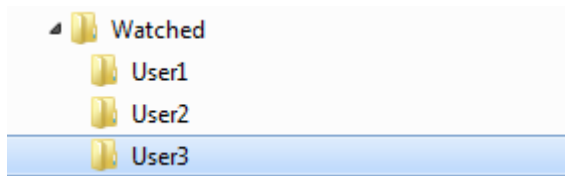
Each watched folder should have the conversion setting file which contains conversion parameters related to the container folder. In case of a missing configuration file, the parent-related settings are used for conversion.

Security

The Folder Watcher system does not contain custom security solutions; embedded Windows privacy settings are used to guarantee document integrity for users.

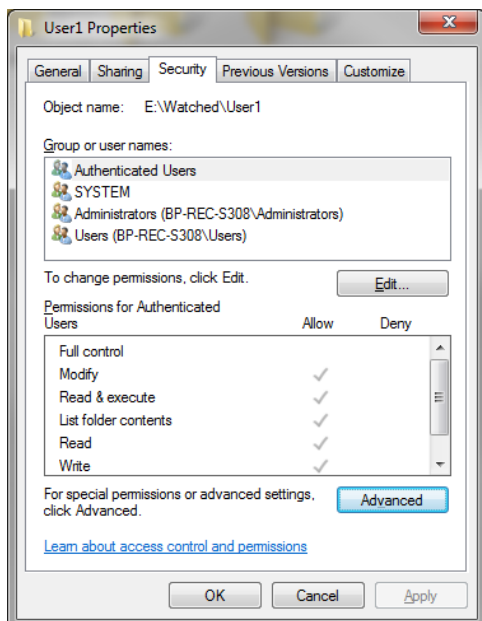
File system watching is executed on behalf of the **Folder Watcher Service** user which is defined during the installation process. The mentioned service user must have full access to the root watched folder and these access rights must be applied to subfolders and files in the subfolders as well. Keep in mind to set the proper rights when planning and creating a watched folder structure.

Example:



- In case of workgroup environment, all users, especially the Service User must have read and write permissions for the Watched folder and its subdirectories.
- In domain environment, the Service User must have read and write permissions for the root and subfolders as well, but the various users only have read and write permissions for their associated subdirectories only. The above example shows how a sophisticated security system can be created in domain environment.

Security settings can be regulated through the **Properties** dialog of the root folder.



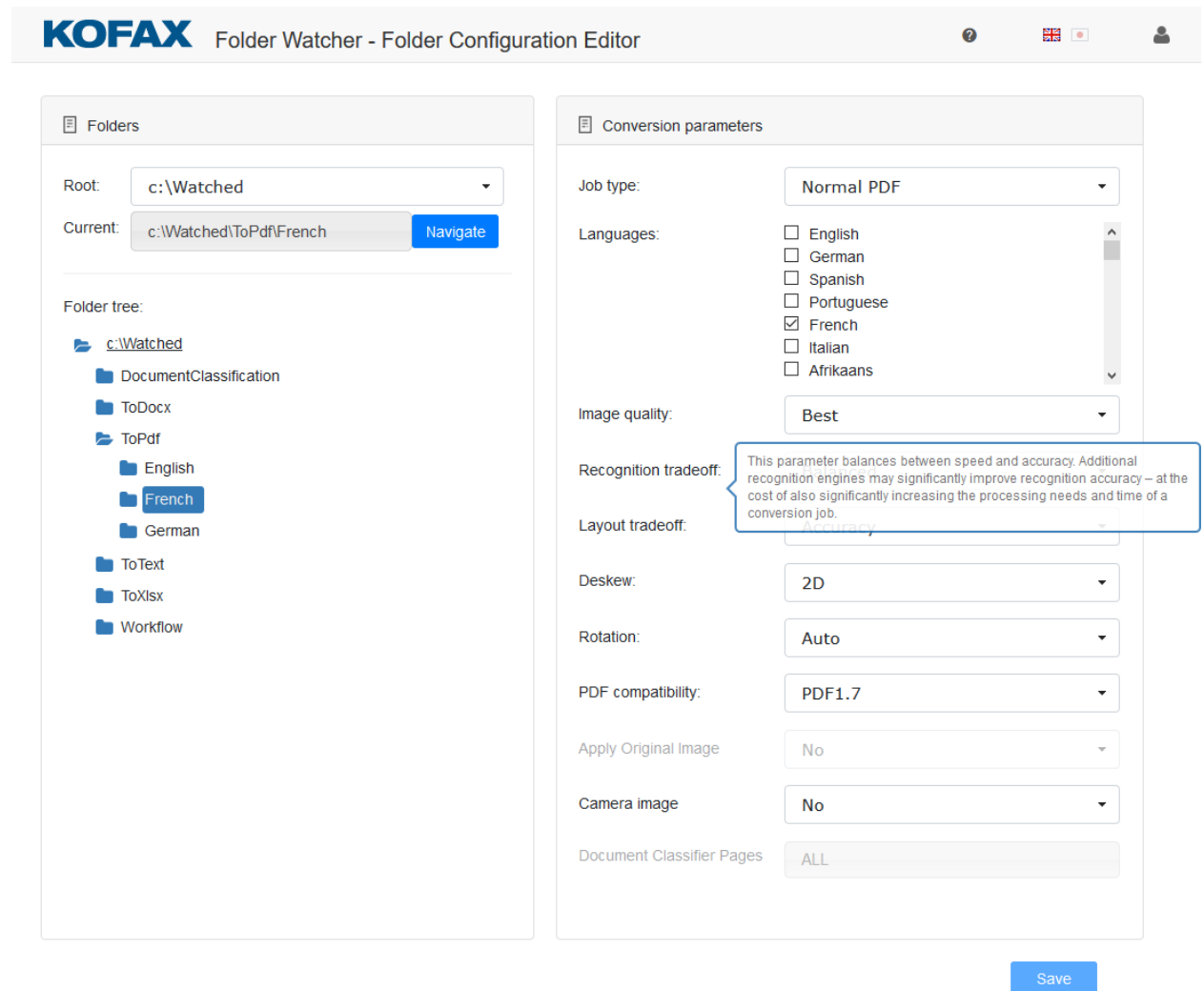
Folder Configuration Editor

The Folder Configuration Editor is used for editing the watched folder related conversion configuration files. The related URL is the following:

<http://{host}/folderconfigurationeditor>

The application can read the folder structure related to the root watched folder, and modify the configuration file (opserverfolderwatcher.config) on behalf of the logged in user, due to the included Windows impersonation.

Function of Folder Configuration Editor



After folder selection in the left side folder tree, the folder-related conversion configuration file is opened. Different conversion parameters can be set on the right panel. The configuration file can be created by clicking the **Create** button. After clicking the **Save** button, the created or modified conversion configuration file is saved.

Conversion Parameters

The following document conversion parameters can be set:

- **Job type:** target document format after file conversion
- **Languages:** the OCR conversion adapter supports many OCR languages, but to provide an easy usage of the system, only the most frequent languages can be selected by default. If necessary, the list can be modified by editing a setting xml file, located in the “Common Application Data”. The location of the setting file typically is the following:

c:\ProgramData\Kofax\FolderWatcher\Settings.xml

To switch on or off the selected language, the related Enabled attribute value must be changed to True or False:

```
<Settings>
  <Languages>
    <Language Name="{name}" Code="{code}" Enabled="true">
      ....
    </Languages>
</Settings>
```

- **Image quality:** OPS supports three image quality settings: Good, Better, Best.
- **Recognition tradeoff:** OPS follows a 'good by default' principle but some use cases may require more than what is already good with digital or scanned image inputs. Camera inputs are often taken with much distortion or with bad illumination and therefore additional preprocessing steps or execution of additional recognition engines may significantly improve recognition accuracy – at the cost of also significantly increasing the processing needs and time of a conversion job. The available settings for the TradeOff element are: Speed, Balanced, and Accuracy. The default selection is Balanced, which already means a near-maximal recognition accuracy with most use cases and is the recommended selection.
- **Layout tradeoff:** with the help of the LayoutTradeOff parameter, it is possible to balance between the speed and accuracy of the recognition. The value of the settings influences which type of auto-zoning and preprocessing algorithm is selected by the OCR engine. The following values are available: Speed, Balanced, and Accuracy. The default value is Accuracy.
- **Deskew:** a technology used to compensate for the minor geometric distortions of input images. Uneven rotations and curved edges are common distortions but in case of camera images it is also common to encounter perspective distortions. Enabling deskew can compensate for these to some extent and can drastically improve recognition accuracy. The available settings for the Deskew element are: AUTO, AUTO3D, 2D, 3D, NO. The default selection is 2D.
- **Rotation:** OPS supports automatic recognition of fixed rotation settings. The default is to use automatic recognition of orientation and rotate outputs accordingly; however, in certain scenarios it might be desirable to use custom rotation settings. The available settings for the Rotation element are: Right, Down, Left, No, Auto. The default value is: Auto.
- **PDF compatibility:** This setting controls compatibility level of the result PDF file. The following values are available: PDF1.3, PDF1.4, PDF1.5, PDF1.6, PDF1.7, PDF/A-1a, PDF/A-1b, PDF/A-2b, PDF/A-2u, PDF/A-2a, PDF/A-3b, PDF/A-3u, PDF/A-3a. The default selection is PDF1.6.
- **Apply Original Image:** This setting can force to save the original image data (without encoding the images) into the result PDF document if it is possible. It has effect in case of Image Only PDF, Image Only PDF/A, Searchable Image PDF, Searchable Image PDF/A.
- **Camera Image:** The digital photos require other processing methods than the scanned images. Select 'Yes' only if the image was made with a digital camera. The recognition module tries to determine if the source image file is captured by a digital camera or a

smart phone, since camera images require different processing algorithms. If the image is already preprocessed, then the camera information (EXIF or resolution information) can be missing, therefore this setting helps to make sure the image is treated as if taken by a digital camera.

- **Document Classifier Pages:** Index collection of pages to classify or a predefined constant: ALL, EVEN, ODD, LAST. The default value is the ALL. For example, 0,1 will categorize the document on the first two pages.

For more information about conversion parameters consult the **OmniPage Server - API Reference Guide**.

Authentication

Folder Configuration Editor can work in two different authentication modes.

Domain environment

In Windows domain environment, the application is installed using Windows authentication settings. In this case, the application authenticates the user and checks the access rights before the user tries to modify the configuration file related to the personal folder of the user.

Workgroup environment

In workgroup environment, **no** authentication is defined, so the application does not try to identify users; therefore, the folder-related user right is not checked. All file system functions are performed on behalf of the system user.

Settings

Folder Configuration Editor receives the configuration file from Configuration Service hosted by Folder Watcher Server. By default, the communication port is 8080. You can change it by modifying the related web.config file. For proper communication, the specified port must be the same as the ServicePort in the folder watcher service configuration file.

```
<appSettings>
  <add key="ConfigurationServicePort" value ="8080"/>
</appSettings>
```


Workflow Job Type

Workflow job type

The workflow job differs from other type of job because it requires a special file called “workflow xml”. The workflow document can be created with the Workflow Designer tool. The Workflow Designer and its documentation is part of the install package, it can be found in the Tools\Workflow Designer folder.

Creation Workflow

These steps can be considered as a programming task, so it requires a thorough knowledge of OmniPage Capture SDK.

1. Copy the Workflow Designer locally from the install package, installation process is not required.
2. Configure the installed OPS in the Workflow Designer. (It is not required but if it is done then the workflow can be tested from the Designer through the OPS.)
 - a. Set the **Executor** to OmniPageServer
 - b. Set up the **EndPoint**
 - c. Set up the credential If you use domain authentication to access the OPS.

The screenshot displays the Workflow Designer interface. At the top, a yellow header bar contains the text "Workflow Sample". Below this, there are several configuration fields:

- Executor:** A dropdown menu with "OmniPageService" selected.
- EndPoint:** A text box containing "http://localhost/OmniPage.Server.Service/api".
- User Name:** A text box with the placeholder text "Enter the User Name".
- Password:** A text box with the placeholder text "Enter the Password".
- Domain:** A text box with the placeholder text "Enter the Domain".

Below these fields is a downward-pointing triangle. Underneath, there is a section titled "Converts input jpg to pdf" with an upward-pointing triangle in the top right corner. This section contains the following configuration:

- Type:** A dropdown menu with "CWFF_LOADIMG, CWFF_RECOGNIZE, CWFF_EXPORTDOC" selected.
- File Names (Collection):** A text box with a browse button (three dots) to its right.
- Converter:** A dropdown menu with "Converters.Text.PDF" selected.
- Output:** A text box containing "%MYDOCUMENTS%\WorkflowSample.pdf" with a browse button (three dots) to its right.

At the bottom of this section is another downward-pointing triangle.

3. Create and parameterize a workflow in the Workflow designer.
4. Export a workflow as XML file.

For details see the „Users Guide” of the Workflow Designer which can be found in the bin folder of Workflow Designer.

Use Workflow

These steps do not require any special skills.

- The Workflow job type should be selected in the Folder Configuration Editor.
- The previously created workflow xml must be copied with the input files into the watcher folder.
- The conversion parameters come from the workflow xml file the ‘traditional parameters’ in the Settings pane are insensitive.
- The workflow job must contain exactly one workflow xml file, the .xml extension is important and at least one input image.
- The output of the workflow will be appeared in the result panel/folder. It includes the output files and the result.xml which contains the description of the workflow (workflow results files, errors/warnings during the workflow).
- At the end of the job, only the input images will be deleted from the input folder, not the workflow xml file..

Input file mapping

The OmniPage Server can match the input files with input file references contained by workflow xml if:

- There is only one <Inputs> node in the workflow xml. In this case all input files will be placed under this <Inputs> node.
- The count of input files is the same as the <Input> nodes. In this case, the inputs will be substituted in the order of upload into the xml.

Otherwise “File mapping error” will be occurred.

Document Classification job type

The Document Classification job requires a 'knowledge base' file which can be created with the export function of the 'Document Classifier Assistant' tool. The Document Classifier Assistant is not part of the OmniPage Server, to get this tool please contact your reseller.

Perform Document Classification

- In the Folder Configuration Editor, select the 'Document Classification' job type for a folder.
- The conversion parameters come from the 'knowledge base' file, the 'traditional parameters' in the Settings pane are insensitive.
- Copy the previously created 'knowledge base' file into the folder. The 'dcp' extension is important.
- Copy the input files to classify into this folder.
- The folder must contain exactly one 'knowledge base' file (.dcp), and at least one input image.
- The classified files will appear in the result folder. Each document will be put into a subfolder which has the same name as the class of document. There are three reserved folders:
 - unclassified – it contains the documents which could not be classified
 - unrecognized – it contains the documents which could not be processed, for example damaged or unsupported input image file
 - classification_result – the detailed description of each classification is put into this folder in xml format
- At the end of the classification job, the 'knowledge base' file will not be deleted from the folder.

Tracking jobs lifecycle

The lifecycle of the Folder Watcher job can be tracked by analyzing the content of the FolderWatcherStatus.info file. This file is generated in the log folder.

For example the following entries are generated during the processing of the sample.tif input file:

```
2018-03-26 11:24:28,123|sample.tif|picked up
2018-03-26 11:24:33,210|sample.tif|conversion started
2018-03-26 11:24:38,183|sample.tif|conversion completed
2018-03-26 11:24:43,131|sample.tif|done
```

Troubleshooting

Folder access issue

The FolderWatcher does not process the files dropped into the configured 'Watched Folder'.

In such cases, check the log files. If you find *"UnauthorizedAccessException: Access is denied."* log entries, then the folder or the share permissions are not set properly. The service user of Folder Watcher must have Read and Change rights to the 'Watched Folder'.

Solution: Give the correct permissions for the service user to the shared folder.

Service connection issue

The Folder Watcher hosts a WCF service on port 8080 by default. If this port is in use by another application, then the Folder Configuration Editor cannot connect to the service to access the configuration. In this case, you get the following error message in the FolderConfigurationEditor or in the SystemConfigurationEditor web application:

"No settings available: Unable to connect to the remote server"

Solution:

- Free up the 8080 port on your machine and restart the FolderWatcher Windows service.

Or

- Configure the FolderWatcher's services manually to use another port in the in a following way: In this example the port will be set to 8081. Use the following command to check if the port is in use:

```
netstat -anob | find "8081"
```

Change the ServicePort config value in the following file:

```
c:\Program Files\Kofax\Folder
Watcher\FolderWatcher\2.1\Omnipage.Server.FolderWatcher.exe.config
For example <add key="ServicePort" value="8081"></add>
```

and the ConfigurationServicePort config value in the following files:

```
c:\Program Files\Kofax\Folder Watcher\2.1\SystemConfigurationEditor\Web.config
c:\Program Files\Kofax\Folder Watcher\2.1\FolderConfigurationEditor\Web.config
For example <add key="ConfigurationServicePort" value="8081"></add>
```

If you use a notification service, then change the port of the notification service URL to 8081 in the FolderConfigurationEditor.

Restart the **FolderWatcher** service.