

KOFAX



Kofax OmniPage Server
Conversion Client
User Guide

Table of Contents

- Introduction3
- Concepts, Terms and Abbreviations3
- File Conversion3
- Document Conversion3
- Security3
- Functions of Conversion Client4
- Creating a conversion process4
- Downloading the result files5
- Creating new conversion5
- Setting conversion parameters5
- Settings7
- Workflow Job Type7
- Workflow job type7
- Creation Workflow7
- Use Workflow8
- Input file mapping8

Introduction

This document describes the Conversion Client which is a Web-based client application of the OmniPage Server.

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 Conversion Client application is developed for companies to provide its users to convert their files easily with a very simple web-based application.

Concepts, Terms and Abbreviations

The basic concept is that the Conversion Client can receive convertible files through a web interface and send those to the conversion server. At the completion of the conversion process, the new documents will be available for the users to download them.

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 information loss if the destination file format cannot represent everything represented by the source file format.

Document Conversion

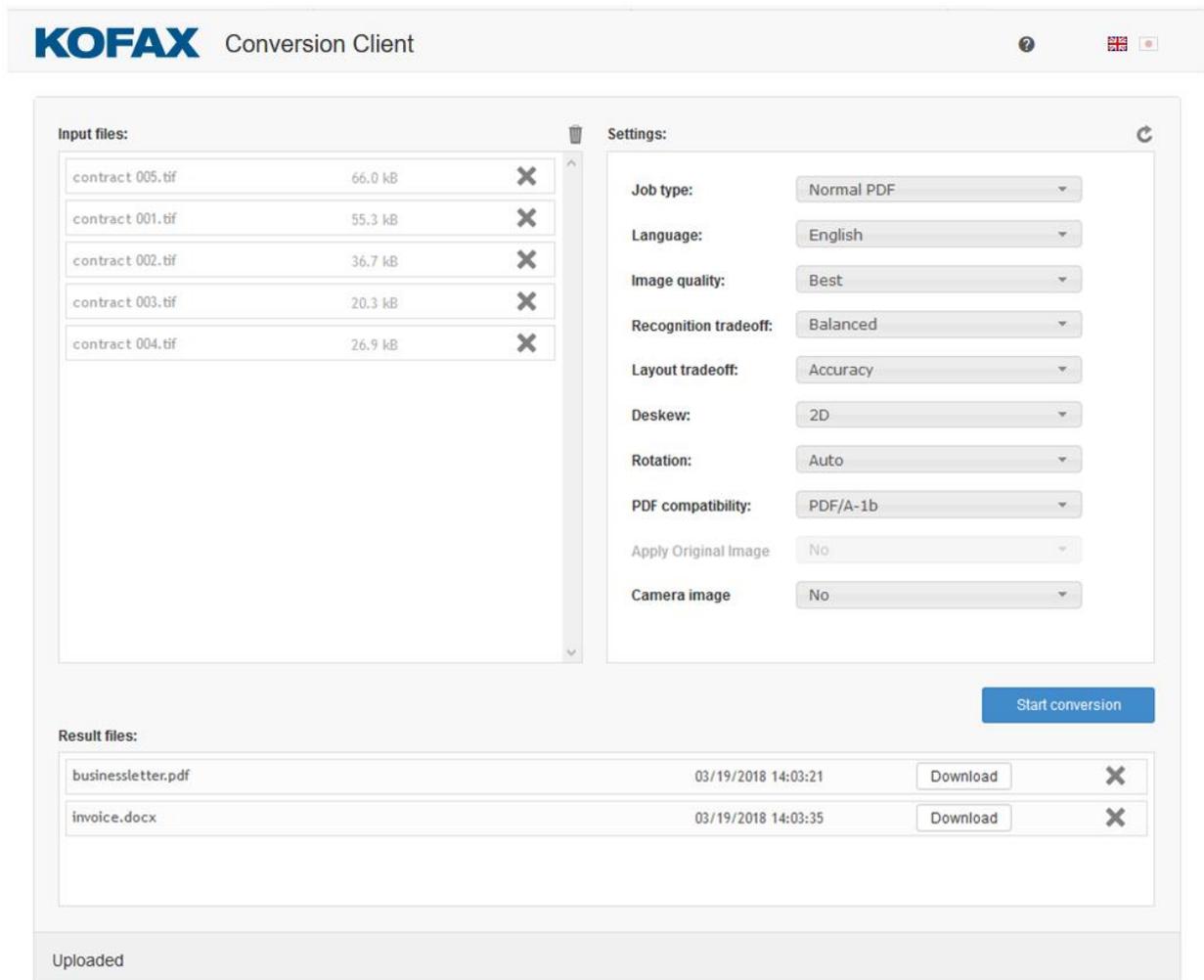
In this document, document conversion refers to the process of transforming a set of input documents into a set of output documents. The transformation is typically file conversion, OCR or both.

Security

Conversion Client application uses Windows authentication or no authentication according to related Conversion Server authentication mode. In domain environment, the conversion happens on behalf of logged-in Windows user. In that case, the logged-in username is displayed in the application header.

If the system is installed in a Windows workgroup environment, the user initializing the conversion process cannot be identified, the conversion is performed on behalf of anonymous account.

Functions of Conversion Client



The Installed Conversion Client application can be reached through navigating to the following link: <http://{hostname}/conversionclient>. This way the user can launch a file conversion process comfortably.

Creating a conversion process

To launch a conversion process, users should drag and drop their files into the **Input Files** panel in the top left corner of the application. Conversion process is started by clicking the **'Start conversion'** button.

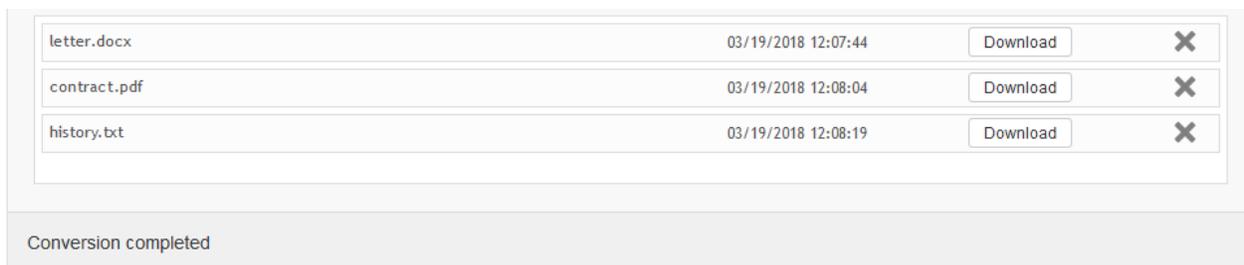


Figure 2: Input files handling

Dragged input files can be reordered by drag and drop. Input files can be removed from conversion process one by one, by clicking the related **X** button, or all of those by clicking the trash icon over the input files list.

Downloading the result files

After the conversion, the result files are received from conversion server and those can be downloaded on the client side by clicking the related **Download** button. A list of result files is available in **Result Files** box at the bottom pane of the application. Result files can be downloaded while those are available on the conversion server. The result file name is derived from the first input file name.



Creating new conversion

Conversions can be started continuously. After the conversion process finishes a new one can be started by dragging files to the input box again.

Setting conversion parameters

Conversion parameters can be set in the right pane of the application. After a parameter changes, the setting is saved to browser cookies to provide ease of use for users. The saved values can be reverted to the default ones by clicking the trash icon over the parameter box.

The following document conversion parameters can be set:

- **Job Type:** this setting determines the target document format after the OCR recognition.
- **Languages:** the document language to recognize.
- **Image quality:** this setting determines the quality of the images in the output file and the size of the file. There are 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 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 it 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 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. Enable this setting to compensate these to some extent and drastically improve recognition accuracy. The available settings for the Deskew element are: AUTO, AUTO3D, 2D, 3D, NO. The default value 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.

Settings

The OCR conversion adapter supports many OCR languages. 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 through editing a setting xml file, located in the “Common Application Data”. The location of the setting file is typically the following:

c:\ProgramData\Kofax\ConversionClient\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">
      ....
    </Language>
  </Languages>
</Settings>
```

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

Workflow Job Type

Workflow job type

The workflow job differs from other type of jobs 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.

- Copy the Workflow Designer locally from the install package, installation process is not required.
- Configure the installed OPS in the Workflow Designer. (This step is not required but if it is done, the workflow can be tested from the Designer through the OPS.)
- Create and parameterize a workflow in the Workflow designer.

- Export a workflow as XML file.

For details see the „Users Guide” of the Workflow designer.

Use Workflow

These steps do not require any special skills.

- The Workflow job type should be selected in the Conversion Client right panel.
- The previously created workflow xml must be dropped with the input files into the input panel.
- The conversion parameters come from the workflow xml file, the ‘traditional parameters’ in the Settings pane are insensitive.
- All the input files are processed with the given workflow xml document. The ‘.xml’ extension is required.
- The workflow job must contain exactly one workflow xml file and at least one input image.
- The output of the workflow appears in the result panel. It includes the output files and result.xml files which contain the description of the workflow (workflow results files, errors/warnings during the workflow).

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 are placed under this <Inputs> node.
- The count of input files is the same as the <Input> nodes. In this case, the inputs are substituted in the order of upload into the xml. Otherwise “File mapping error” will occur.