Kofax (OCR) Integration

By CMiC

CM*i*C
Computer Methods
international Corp.

Proprietary Notice

The contents of the CMiC software product, including both this manual and the program components, are proprietary to Computer Methods International Corp. (CMiC), are copyright protected and are considered Confidential Information by CMiC. All rights are reserved by CMiC. A copy of the manual and the program has been provided to the original licensee under license with CMiC strictly for their own use under the terms of the license. All copies of this manual whether in print or electronic format must contain a copy of this Proprietary Notice. Any selling, licensing or other distribution of the contents of either this manual or the program components, whether for profit or not, is unlawful and may subject the violator and the original licensee to termination of license, criminal charges, civil action, or any combination of these.

Copyright © 2019

Computer Methods International Corp.
4850 Keele Street
Toronto, Ontario M3J 3K1
Canada

Risk of Use Notice

The CMiC software product, including both this manual and the program components, is licensed on an "AS IS" basis. The entire risk as to the results of its use is with the licensee. Except in those jurisdictions which impose certain warranties by statute which may not be waived by one or more of the parties, and only to that extent, Computer Methods International Corp. (CMiC) makes no warranties whatsoever, either expressed or implied, with respect to the quality, performance, merchantability or fitness for any particular purpose of any or all components of this software product, except as provided in the licensee's license agreement. The licensee (and not CMiC or its agents) will be solely responsible for the costs of all service, or of any defect in this software product and any incidental or consequential damages caused or alleged to be caused either directly or indirectly by the software product to the licensee or any other person, including, but not limited to, any interruption of service, or loss of business or anticipatory profits, even if CMiC has been advised of the possibility of such damages.

Contents

Kofax (OCR) Integration

Overview

CMiC created an Optical Character Recognition (OCR) integration with Kofax to facilitate the extraction and exchange of digital documents and information to generate Imaging objects like AP Registered Invoices and Subcontract Request for Payments.

Kofax provides the scanning, OCR and data release infrastructure. Operators scan invoices into an OCR workflow where invoice attributes such as invoice number, job number, and invoice date are extracted and validated using CMiC's Kofax schema to enforce accuracy and data integrity.

The Kofax OCR process extracts and writes the invoice information to an XML file and copies the scanned document associated with the XML file to a CMiC Imaging Source folder for registration.

The Imaging registration process scans the Source folders, and when it detects a pair of XML and PDF files, the image is registered with the Imaging system and uses the data from the XML file to populate the flysheets for the document type to complete the import.

Kofax places invoices with missing attributes or with failed validation into a validation queue. An operator then checks the queue to rectify the problem. After manually adjusting the extraction, the release process copies the invoice and the associated XML data into an Imaging Source folder for registration.

CMiC Validation

The Kofax adapter provides a schema of database functions that assist with real-time validation. As an example, there's a database function that notifies Kofax if the invoice being OCR'd is a duplicate. This is important since the extraction will be correct, but the import will fail during the registration process since the invoice number is assigned to an existing invoice for the vendor.

In addition, the Kofax adapter comes with a provision to export Job and Vendor files nightly to support Kofax fuzzy look-up logic.

CMiC's DBAs enable the Kofax schema and initialize the export for the Job and Vendor files after purchasing the CMiC Kofax adapter.

Invoice Attributes

Operators prepare and scan invoices into the Kofax data acquisition process. Kofax's capture software will extract at a minimum the mandatory invoice attributes.

The table below shows the mandatory invoice attributes obtained during the OCR process along with values the Imaging registration process will populate:

Invoice Attributes	Mandatory - Y/N
1. Job company code	Y, provided by the OCR process

2.	Job code	Y, for Job charges – provided by the OCR process
3.	Department code	Y, for Overhead charges – provided by the OCR process
4.	Vendor code	Y - provided by the OCR process
5.	Registration group	Y - provided by the OCR process
6.	Invoice number	Y - provided by the OCR process
7.	Invoice date	Y - provided by the OCR process
8.	Invoice amount	Y - provided by the OCR process
9.	Discount amount	N - provided by the OCR process
10.	Description	N - provided by the OCR process

XML Interface

1. XML Format

After creating the [header] flysheet for the document type, create an XML file based on the flysheet table columns. Kofax will use the XML format to structure the data for the Imaging registration process to read and import.

Any other OCR provider can use this XML interface to integrate with CMiC's Imaging application.

2. XML Data

Kofax's release process writes the extracted data, along with the image filename into an XML file. The format of the XML file is the same for other invoice types such as Subcontract Request for Payment and PO invoices, only the names of the tags differ.

Here's what a sample of the XML file would look like for an *AP Register Invoice*:

```
<Rows>
 <Row num="1">
  <Apri_Comp_Code>COMPANY CODE</Apri_Comp_Code>
  <Apri_Job_Code>JOB CODE</Apri_Job_Code>
  <Apri_Exp_Dept_Code>DEPARTMENT CODE</Apri_Exp_Dept_Code>
  <Apri_Ven_Code>VENDOR CODE</Apri_Ven_Code>
  <Apri_Inv_Code>INVOICE NUMBER</Apri_Inv_Code>
  <Apri_Inv_Date>INVOICE DATE</Apri_Inv_Date>
  <Apri_Amt>INVOICE AMOUNT</Apri_Amt>
  <Apri_Disc_Amt>DISCOUNT AMOUNT</Apri_Disc_Amt>
  <Apri_Desc>INVOICE DESCRIPTION</Apri_Desc>
  <Image_Name>FILE NAME</Image_Name>
 </Row>
<Rows>
```

Data Exchange and Imaging Registration

After Kofax scans and extracts the data, the Release module creates the XML file according to the file specification and then copies scanned document files along with the XML data-file into the Imaging Source path for the document type.

The Imaging registration process checks the Source folder at a specified frequency and if any new files are detected, it:

1. Opens the XML file in the source folder.

2. Parses the XML file and creates an *AP Registered Invoice* or the targeted Imaging document type.

3. The scanned image then appears in the Image Manager application with an *un-submitted* status, along with the header flysheet populated with the data from the XML file.

   The auto-submit to workflow feature will kick in if the registration process creates the image and associated header flysheet successfully. The image is submitted automatically to the workflow approval process assigned to the document type to trigger the approval process.

Kofax (OCR) Adapter Installation

Overview

This section explains how to initialize the Kofax (OCR) database components and schedule the export-data-files for data look-ups and validation.

Assumptions

All files are in the standard CMiC SQL directory.

NOTE: If needed, ODBC drivers are available from the Oracle database client installation media.

Installation Scripts and Scheduled Tasks

1. **kofax_main_script.sql**

   This script installs the *Kofax* (database) user, and grants access to the stored procedures used by the Kofax system for look-ups and validation. Run the script as a privileged user like SYS from the command prompt path.

   NOTE: It's important that you set the SQL path to the *CMiC SQL* directory for successful execution.

2. **kofax_flat_files.bat**

   Required parameters (for each environment):
   - *connectstr- kofax* - connection string for target database
   - *environment* – test or prod
   - *destination* – destination folder for CSV export files

3. **Modify and add this batch file to a scheduled task.**

   Each environment (TEST and PROD) needs a separate scheduled task. Configure the scheduled task to call the *kofax_flat_files.bat* file from a specified folder. Examples of these would be something like D:\cm\v10\prod\sql\kofax_flat_files.bat or D:\cm\v10\test\sql\kofax_flat_files.bat.

   The batch file executes the *JobFlatFile.sql* and *VendorFlatFile.sql* scripts, which stores the CMiC Job and Vendor data into the *JobFlatFile.csv* and *VendorFlatFile.csv* files – then copies them to the destination folders specified in the *kofax_flat_files.bat* batch file.

Notes

At the time of creating this manual, the Kofax adapter supports the following Imaging document types:

1. AP Register Invoice

2. Register Pay Request

3. Voucher-based Subcontract Request for Payment

4. PO Invoices

5. User-defined Imaging Document Types and its associated header flysheet

CMiC is always enhancing our Kofax adapter to support more Imaging document types and quality updates. For the latest, contact your account manager.

Index