

Kofax Transformation Modules

6.4.0

Product overview
& new features

KOFAX

Product overview

Kofax Transformation Modules streamlines the transformation of business documents into structured electronic information by automating the process of document classification, separation and extraction.

Whenever incoming documents drive transactions, organizations will benefit from:

- ◆ Reduced operating costs
- ◆ Increased productivity
- ◆ Better data quality
- ◆ Improved compliance

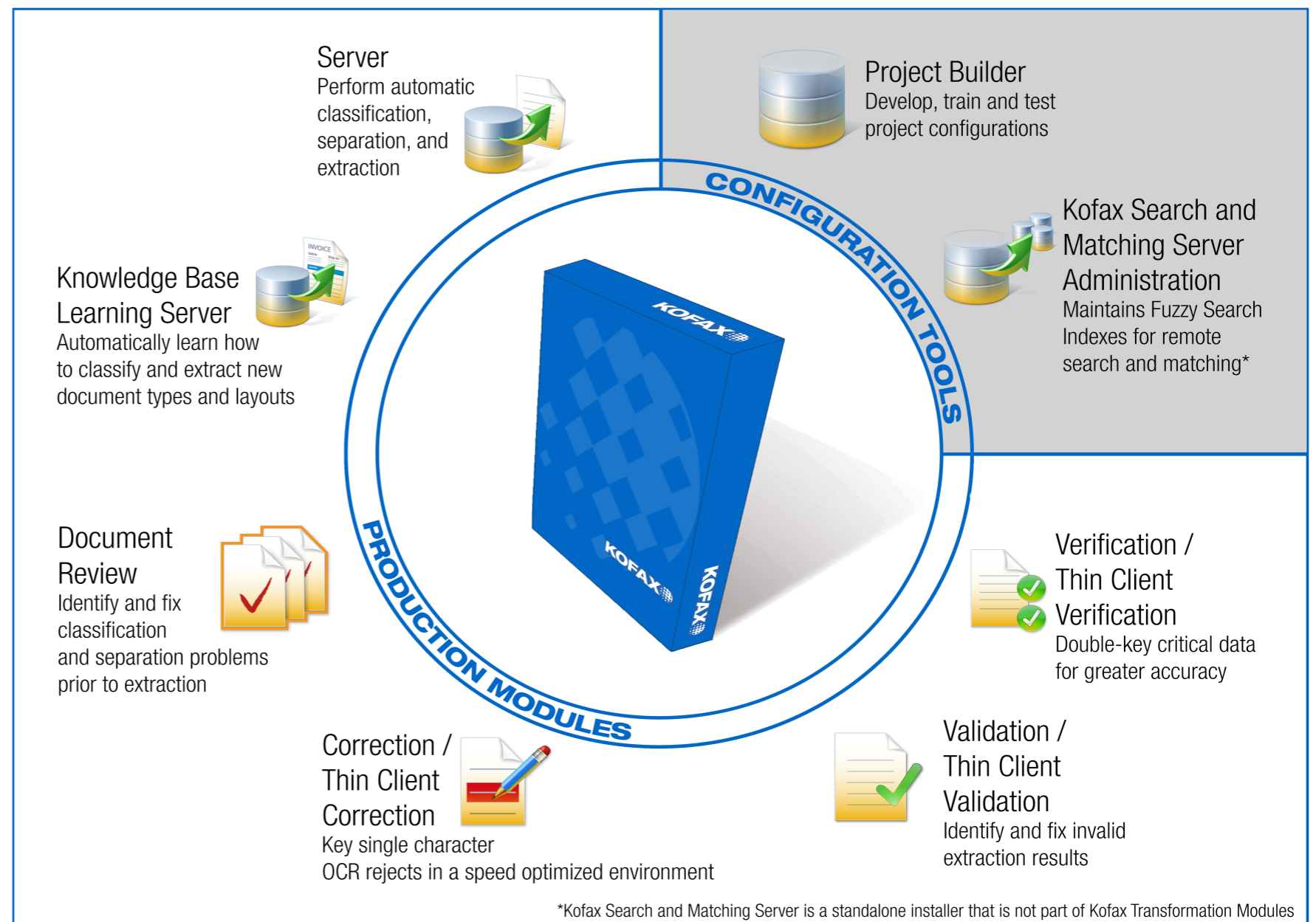
Kofax Transformation Modules is the most complete and versatile document transformation offering on the market, enabling the processing of forms, invoices, correspondence and other document types on a single platform. Using learn-by-example techniques for document classification, separation and extraction, solutions can be configured and optimized quickly and effectively.

Kofax Transformation Modules provides a seamless integration with Kofax Capture, giving access to the widest range of document scanners and back-end storage solutions, as well as benefiting from its distributed capture, high availability and enterprise capabilities.

Paper-based and electronic documents can be scanned or imported, creating a series of scanned image files. Kofax Capture then routes these image files through Kofax Transformation Modules where a batch is separated into documents, classified, and then extracted.

These classification, separation, and extraction results are presented for review by users of the Document Review, Correction or Thin Client Correction, Validation or Thin Client Validation, and Verification or the Thin Client Verification user modules, depending on the project configuration.

There is no limit to the number of Validation, Verification, Document Review, or Correction workstations that can be deployed.



Once all documents are successfully processed, the accurate and fully-validated data and images are exported to a back-end system using Kofax Capture.

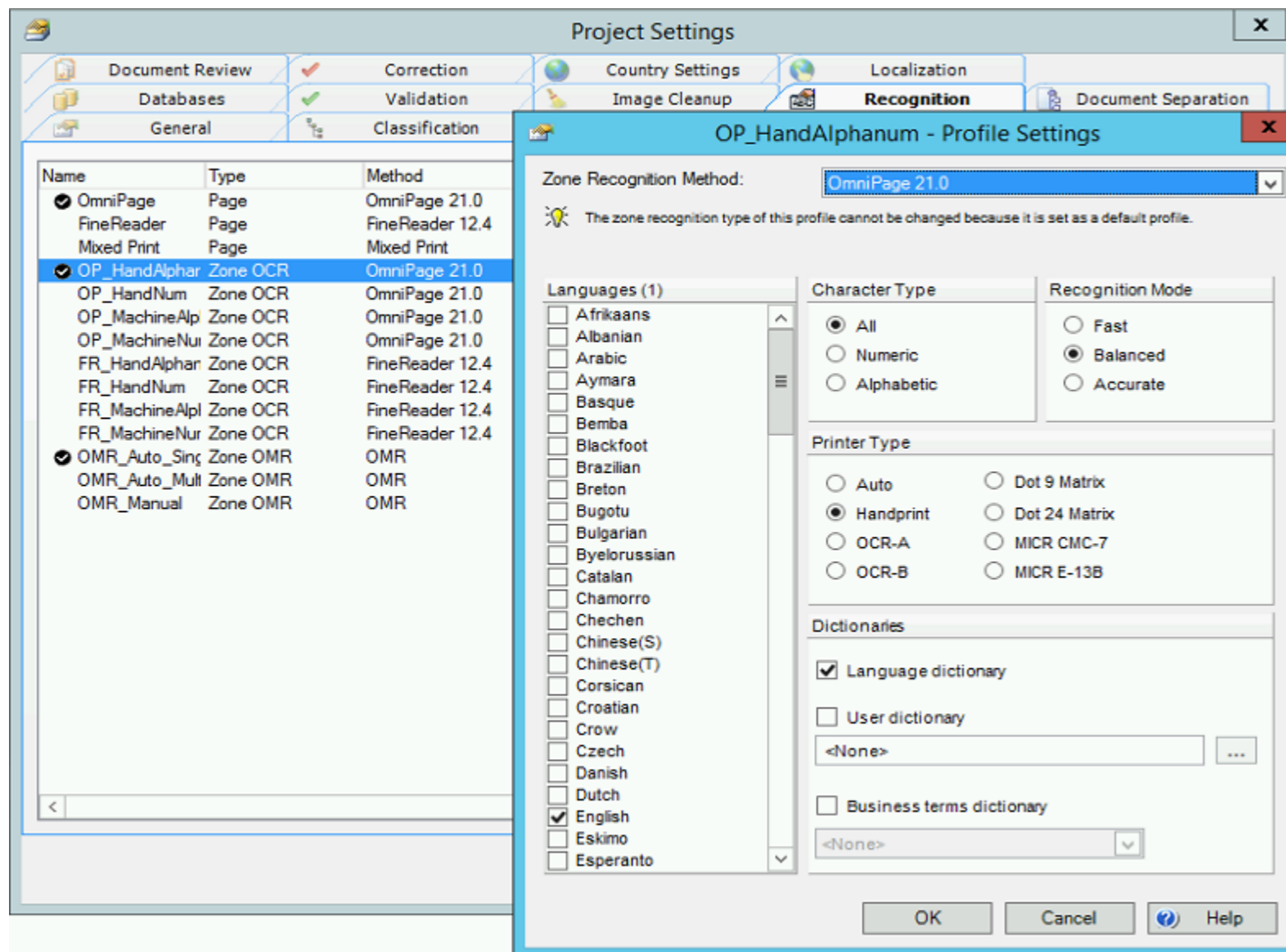
OmniPage recognition engine

Kofax OmniPage is a new recognition engine that recognizes full pages or individual zones, depending on the recognition profile. This engine supports a wide variety of languages and includes some business term dictionaries that may improve results. OmniPage, like other recognition engines, supports various print types, such as hand writing, matrix fonts, etc.

The OmniPage recognition engine is a replacement for the RecoStar recognition engine that is now deprecated.

For all new projects, the OmniPage recognition engine is used as the default page profile on the Project Settings - Recognition tab. In addition, the default zone profile is an OmniPage profile that is specialized for hand written alphanumeric text. There are several other OmniPage zone profiles that are specialized for other zone configurations.

For upgraded projects, any existing RecoStar recognition profiles are replaced with comparable OmniPage recognition profiles.



Kofax recommends that a copy of an existing OmniPage profile is made rather than editing one of the default profiles.

This ensures that the default profiles are always available for reference.

Natural language processing

The Natural language processing engine is now available to extract named entities and the mood or sentiment of a document. This engine is installed separately, and supports several languages.

The following locator methods use the Natural language processing engine.

Named Entity Locator

This new locator method is used to assign extracted entities to fields. Choose to extract the entity name only by using a simple field, or use a table field to extract the entity name as well as the entity confidence, entity type, and the entity sentiment. If you want to extract more than one type of entity, a separate Named Entity Locator is required for each entity type.

Once entities are extracted, customize a script to interpret the Named Entity Locator results as needed.

Sentiment Locator

This new locator method extracts text sentiments from a document. This means that the Sentiment Locator is able to determine the overall mood or sentiments of a document, based on words or phrases found on that document.

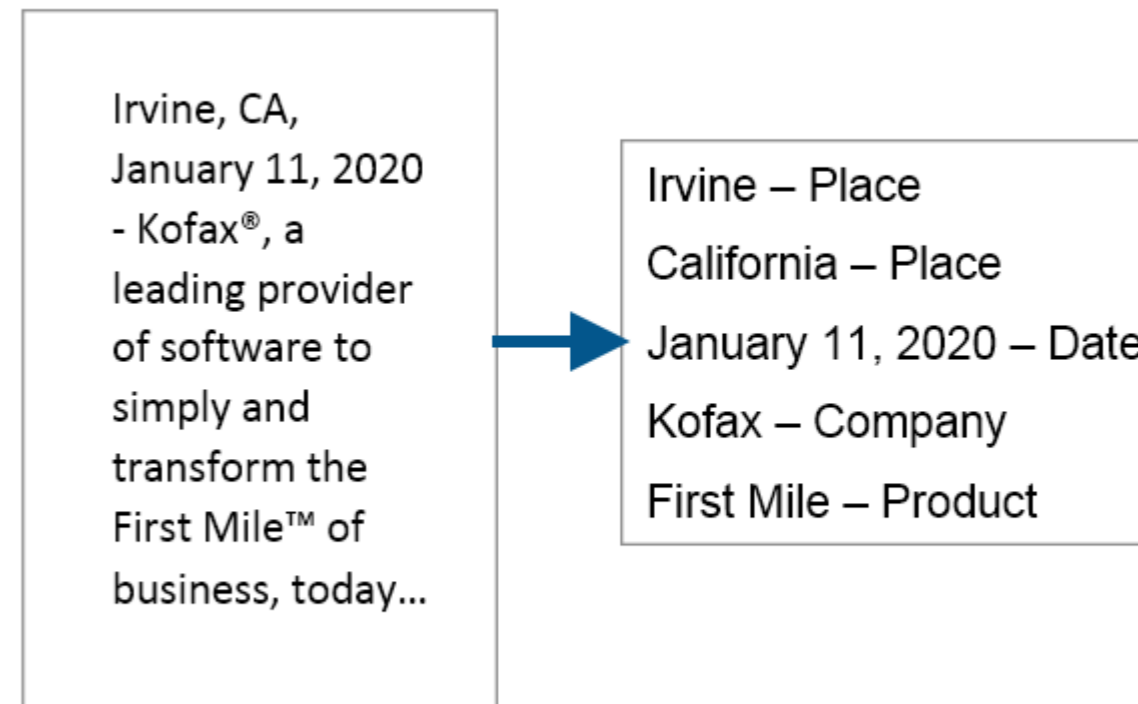
Once the sentiment is extracted, customize a script to interpret the Sentiment Locator results as needed.

Themes Locator

This new locator method is used to extract the theme of a document. Using a table field, map columns for the theme, the confidence, and the sentiment of a theme.

Summary Locator

This new locator method provides a summary of a document, page, or region, depending on the locator settings. The results of this locator can be viewed so that the gist of a document can be seen without reading an entire document.



Paragraph detection

Paragraph detection is performed during recognition and takes advantage of character recognition. The detected paragraphs can then be used by a Classification Locator or in a script.

For example, an organization processes large legal contracts and they want to check whether these documents include an indemnity clause, an intellectual property clause, etc. Rather than reading the entire document it is possible to use paragraph detection and a Classification Locator to classify the paragraphs and identifying the recognized clauses. This can be done by populating a field with a single paragraph or by populating a table with several paragraphs. This enables an operator to quickly access a clause or to find documents where a clause was not detected.

Similarly, an organization receives a document that is issued by the government on a daily basis and contains a list of newly registered companies. When a document is processed, it is necessary to place each new company and its corresponding address into a database. The document is structured as a very dense 6-column PDF where each company is a small paragraph. It is possible to use paragraph detection and a script to populate the database.

Sentences are typically grouped into paragraphs by indenting the first sentence in the paragraph, by adding vertical white space between paragraphs, or by numbering paragraphs. Paragraph detection uses this typical behavior, it is also able to detect paragraphs in other situations, such as the following:

- ◆ When a document has multiple columns
- ◆ When a paragraph contains different fonts and text sizes
- ◆ When a paragraph contains text with various font effects, such as bold, italic, super and subscripts, text colors, and background text colors
- ◆ When there are images embedded in the text flow
- ◆ When a paragraph includes numbered or bullet lists

Paragraph detection works best when a document is text-based, without images, tables, or other content that breaks up the document. The results of paragraph detection decrease as the complexity of the document layout increases.

Right-to-left languages are not currently supported for paragraph detection.

34 Handelsregister

25.000 EUR. Geschäftsführer: Steven De Proost, Begijnendijk/Belgien; Koen Boriau, Antwerpen/Belgen. Die Gesellschafterversammlung vom 12.3. hat die Verlegung des Sitzes von Köln nach Bayreuth beschlossen. (10.4.)

S&P ImmoPlan UG (haftungsbeschränkt), Bodenseering 4, **Bayreuth** (Der Erwerb, die Planung, die Entwicklung und die Vermarktung von Grundstücken und Immobilien sowie alle damit verbundenen Tätigkeiten). Stammkapital: 3.000 EUR. Geschäftsführer: Michael Schönberger, Memmelsdorf; Robert Plastil, München. (11.4.)

Amatec PV 20 GmbH & Co. KG, An der Feuerwache 15, **Bayreuth** (Errichtung und Betrieb von Photovoltaikanlagen). Persönlich haftende Gesellschafterin. Firma Amatec Projects Management GmbH, Bayreuth. Ein Kommanditist. Der Sitz wurde von Hebertsfelden nach Bayreuth verlegt. (18.4.)

Amatec PV 21 GmbH & Co. KG, An der Feuerwache 15, **Bayreuth** (Errichtung und Betrieb von Photovoltaikanlagen). Persönlich haftende Gesellschafterin. Firma Amatec Projects Management GmbH, Bayreuth. Ein Kommanditist. Der Sitz wurde von Hebertsfelden nach Bayreuth verlegt. (18.4.)

Amatec PV 22 GmbH & Co. KG, An der Feuerwache 15, **Bayreuth** (Errichtung und Betrieb von Photovoltaikanlagen). Persönlich haftende Gesellschafterin. Firma Amatec

Pflasterbau Müller GmbH, Naabquellenweg 8 b, **Fichtelberg** (Pflaster- und Natursteinarbeiten, Steinfresh). Stammkapital: 25.000 EUR. Geschäftsführer: Rainer Markus Müller, Fichtelberg. (18.4.)

Amatec PV 36 GmbH & Co. KG, An der Feuerwache 15, **Bayreuth** (Errichtung und Betrieb von Photovoltaikanlagen). Persönlich haftende Gesellschafterin. Firma Amatec Projects Management GmbH, Bayreuth. Ein Kommanditist. Der Sitz wurde von Hebertsfelden nach Bayreuth verlegt. (24.4.)

Amatec PV 37 GmbH & Co. KG, An der Feuerwache 15, **Bayreuth** (Errichtung und Betrieb von Photovoltaikanlagen). Persönlich haftende Gesellschafterin. Firma Amatec Projects Management GmbH, Bayreuth. Ein Kommanditist. Der Sitz wurde von Hebertsfelden nach Bayreuth verlegt. (24.4.)

picfol GmbH, Kleiner Johannes 21, **Pegnitz** (Der Handel mit und der Vertrieb von Folien aus Kunststoff, insbesondere nichttransparente und/oder bedruckte, selbsthaftende Kunststofffolien, Adhäsionsfolien, Haftfolien und Hart-PVC-Folien). Stammkapital: 25.000 EUR. Geschäftsführer: Thorsten Küspert, Prebitz. Die Gesellschafterversammlung vom 1.4. hat die Verlegung des Sitzes von Baring nach Pegnitz beschlossen. (26.4.)

Pompl-Forst UG (haftungsbeschränkt), Knechtel 2, **Mittelsdorf** (Frucht- und Gemüseerzeugung)

als Geschäftsführer aus: Maximilian Preiß, Bisch zum einzelvertretungsk Geschäftsführer bestellt

FWR Energie GmbH (Hans Vogtmann ist als ausgeschieden. Harald zum Geschäftsführer bestellt

KMT GmbH Kunststoff Maschinentechnik, G Preiß ist als Geschäftsf Die Gesellschaft ist auf Margit Preiß, Goldkron

Unternehmensberatung GmbH, Fichtelberg. A Geschäftsführerin aus Voigt, Fichtelberg, wurde bestellt. Die Prokura Ch erloschen. Einzelprokur Fichtelberg. (4.4.)

Carl Beutlhauser Kon & Co. KG, Kulmbach. Stahl ist erloschen. (5.4)

Elektro Schumacher (Gesellschafterversamm teilweise Änderung der Der Sitz wurde nach Bi 6, verlegt. (5.4.)

Hacker – Transport – Konrad Hacker ist als G

Thin Client Server

The Thin Client Server is built on a framework that has been updated to comply with current security standards. In addition, the user interface for each Thin Client module have been updated so that it matches the new Kofax web-based branding. This includes new toolbar icons and an updated logo.

KOFAX Transformation - Validation

? ⓘ John Smith



The overall layout of Thin Client modules has not changed. The same panes and fields are still available.

The communication between the Thin Client Server and the Thin Client modules is now more efficient. For example, thumbnails are loaded for a document only when that document becomes active. This ensures that there are fewer memory issues. This is especially true for large batches or a batch that has large multi-page documents.

Most of the Thin Clients behavior has remained the same. However, the following small changes have been made so that the Thin Client behavior is more in line with the Rich Client behavior.

- ◆ Common changes for all modules
 - ◆ A new top-level toolbar for common settings
 - ◆ A new Settings window to configure automatic batch opening
 - ◆ Filtering batches has been reorganized and behaves the same for all modules
 - ◆ URL containing filter options must be regenerated
 - ◆ New setting **Prompt before closing completed batch**
 - ◆ User informed about session timeouts via a pop-up message
 - ◆ By default, single sign-on is tried first
 - ◆ Bitonalization of PDF is now supported
- ◆ Thin Client Validation
 - ◆ Use Ctrl+Enter to force a field valid
 - ◆ Table-specific keyboard shortcuts are available and fully documented
 - ◆ The **Support scripted layout changes** project setting has been removed
 - ◆ Use Shift Arrow Up and Shift Arrow Down to navigate between editable fields
 - ◆ Modified fields are committed to the Server when a user closes or suspends a document
- ◆ Thin Client Verification
 - ◆ Use Shift Arrow Up and Shift Arrow Down to navigate between editable fields
- ◆ Thin Client Correction
 - ◆ Navigation between fields now uses the Arrow keys for better performance
 - ◆ Use Ctrl + Enter to override a field

