

# Kofax Transformation Toolkit

## Getting Started Guide

Version: 6.3.0

Date: 2019-04-01



© 2006-2019 Kofax, 15211 Laguna Canyon Road, Irvine, California 92618, U.S.A. All right reserved.  
Portions © 2002-2006 Kofax Development GmbH. Portions © 1997-2006 Kofax U.K. Ltd. All Rights Reserved. Use is subject to license terms.

Third-party software is copyrighted and licensed from Kofax's suppliers.

This product is protected by U.S. Patent No. 5,159,667.

THIS SOFTWARE CONTAINS CONFIDENTIAL INFORMATION AND TRADE SECRETS OF KOFAX  
USE, DISCLOSURE OR REPRODUCTION IS PROHIBITED WITHOUT THE PRIOR EXPRESS  
WRITTEN PERMISSION OF KOFAX

Kofax, the Kofax logo, Kofax Transformation Toolkit, Ascent Xtrata Pro, INDICIUS, Xtrata, Ascent Capture, Kofax Capture, VirtualReScan, the "VRS VirtualReScan" logo, and VRS are trademarks or registered trademarks of Kofax or its affiliates in the U.S. and other countries. All other trademarks are the trademarks or registered trademarks of their respective owners.

U.S. Government Rights Commercial software. Government users are subject to the Kofax standard license agreement and applicable provisions of the FAR and its supplements.

You agree that you do not intend to and will not, directly or indirectly, export or transmit the Software or related documentation and technical data to any country to which such export or transmission is restricted by any applicable U.S. regulation or statute, without the prior written consent, if required, of the Bureau of Export Administration of the U.S. Department of Commerce, or such other governmental entity as may have jurisdiction over such export or transmission. You represent and warrant that you are not located in, under the control of, or a national or resident of any such country.

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

# Table of Contents

<b>Preface</b> .....	<b>5</b>
Related documentation.....	5
Training.....	5
Getting help for Kofax products.....	5
<b>Chapter 1: Overview</b> .....	<b>7</b>
<b>Chapter 2: Install Kofax Reporting on client site</b> .....	<b>8</b>
Command line installation.....	9
<b>Chapter 3: Reported statistical data</b> .....	<b>11</b>
Server Processing.....	11
Document Review.....	12
Correction.....	13
Validation.....	14
Verification.....	15
Knowledge Base Learning Server Processing Step.....	16
<b>Chapter 4: Predefined views</b> .....	<b>17</b>
View v_batch_completion.....	18
View v_tsf_user_perf.....	19
View v_tsf_user_perf_by_mod.....	19
View v_batch_session.....	20
View v_tsf_class_vs_review.....	20
View v_field_acc.....	20
View v_field_confidence.....	21
View v_tsf_statistics.....	21
View v_recog_acc_by_batch.....	22
View v_recog_acc_grouped.....	22
View v_recog_timing_by_batch.....	22
View v_doc_volume_by_batch.....	23
View v_doc_volume_by_grouped.....	23
View v_recog_acc_by_field.....	23
View v_recog_timing_grouped.....	24
View v_recog_acc_by_batch_grouped.....	24
View v_field_recog_acc_grouped.....	25
View v_tsf_class_acc.....	25
View v_tsf_sep_acc_by_batch_class.....	25

View v\_tsf\_extr\_acc\_grouped.....26  
View v\_sla\_monitoring.....26

# Preface

This guide contains information about installing and configuring the Kofax Reporting platform for Kofax Transformation Toolkit.

## Related documentation

In addition to this guide, Kofax Transformation Toolkit includes the following documentation.

### Kofax Transformation Toolkit Developer's Guide

The developer's guide contains installation information and provides details about the provided merge modules and code samples as well as detailed information concerning integration.

### **Help**

Help for Kofax Transformation Toolkit is available from the applications as follows:

- From any of the Kofax Transformation Toolkit components, press F1, click Help from the toolbar, or select Contents from the Help menu.

**Note** There is no Help available for the Kofax Transformation Toolkit sample applications.

### **Help for Scripting**

Information about scripting is available from the Help menu of any Project Builder interface that enables you to write or access scripts.

**Note** This Help is available in English only.

## Training

Kofax offers both classroom and computer-based training that help you make the most of your Kofax solution. Visit the Kofax Web site at [www.kofax.com](http://www.kofax.com) for complete details about the available training options and schedules.

## Getting help for Kofax products

Kofax regularly updates the Kofax Support site with the latest information about Kofax products.

To access some resources, you must have a valid Support Agreement with an authorized Kofax Reseller/ Partner or with Kofax directly.

Go to <http://www.kofax.com/support/> for:

- Access to product knowledge bases.  
Click **KNOWLEDGE Base**.
- Access to the Kofax Customer Portal (for eligible customers).  
Click **Kofax Customer Portal** and log in.

To optimize your use of the portal, go to the Kofax Customer Portal login page and click the link to open the *Guide to the Kofax Support Portal*. The guide describes how to access the portal, what to do before contacting the support team, how to open a new case or view an open case, and what information to collect before opening a case.

- Product information and release news  
Click **Transformation > Kofax Transformation Toolkit**.
- Downloadable product documentation  
Click **Transformation > Kofax Transformation Toolkit > Documentation** and select a document.
- Access to support tools  
Click **Tools** and select the tool to use.
- Information about the support commitment for Kofax products.  
For more information on the support commitment, click **Learn More**.

## Chapter 1

# Overview

Kofax Transformation Toolkit supports the Kofax Reporting platform to report statistical data during batch processing in order to have detailed information about the processing time and accuracy as well as other information to control your business process. You can create various kinds of reports for the gathered data by the help of [predefined database views](#).

The Kofax Reporting platform is a server application that can handling reported data from various Kofax products. For the communication between the client site that gathers statistical data and the server site that stores this data in a reporting database and provides several database views to generate reports either a web service architecture or an enterprise service bus is needs to be configured. The following documentation provides the information about how to set up the communication using the web service architecture (WSA) in case your are using an enterprise bus service, refer to the Kofax Reporting documentation for further information.

## Chapter 2

# Install Kofax Reporting on client site

In order to report statistical data to the Kofax Reporting platform you need to configure a web service architecture by installing the WSA Sender on each workstation that is supposed to gather data.

A standard installation uses the web service architecture. Therefore the WSA Sender installation and configuration steps are described in detail. Instead of the web service architecture an enterprise service bus (ESB), such as Sonic, can be used. For more details about its installation and usage refer to the *Kofax Reporting Administrator's Guide*.

In addition to installing the WSA Sender you need to ensure that the reporting components for Kofax Transformation Toolkit are installed and the Kofax Reporting platform at server site are installed and configured before you can report any statistical data.

**Note** You can configure the URL to the WSA Receiver during the Kofax Transformation Toolkit installation even if the receiver is not installed yet. Of course, the Kofax Reporting platform, including the WSA Sender at client site, must be fully installed and configured before any reported data can be written to the reporting database. For information how to set up and configure the server, refer to the *Kofax Reporting Administrator's Guide*.

Especially, if you are installing the Kofax Transformation Toolkit software on several client sites using the silent install it is best practice to set the `REPORTINGURL` and `REPORTING` properties. Otherwise, you have to run the configuration tool to set the URL manually on every client computer.

For more information about the Kofax Transformation Toolkit Kofax Transformation Toolkit and how to perform a toolkit silent install, refer to the *Kofax Transformation Toolkit Developer's Guide*.

**Note** You can install the WSA Sender using a [quiet mode installation](#). Several parameters can be set to define the configuration settings. You can allow use the default settings and start the WSA configuration utility at a later point of time to change the current settings.

1. Start Setup.exe from the Kofax Reporting installation ISO.  
The setup splash screen is displayed that allows selecting different .
2. Click **Install WSA sender (several instances...)** and then click **Install WSA sender Instance 1** to install the WSA Sender application.
3. Specify the destination folder. Click **Next**.
4. Configure the WSA sender by performing the following steps.

**Note** You can keep the default values while installing and configure WSA at a later point of time. For more details about how to configure the WSA Sender refer to *Kofax Reporting Administrator's Guide - Configuring WSA*.

- Specify the Service name. Usually there is little reason to modify the default.



- Select the **Transport mode**: HTTP, MSMQ, IBM WebSphere MQ.
  - Specify the **Local server** name that means the name of the computer where the component is installed.
  - Specify the **Local port** that is used for the communication.
  - Specify the **Central URL** (URL of the WSA receiver on the central site).
  - Specify the "config.file" (path and file name) on client site. WSA sender saves the URL of its web service interface to this configuration file. Kofax applications read this file and use it for connecting to the web service interface.
5. Click **Next**, then click **Install**.
  6. Click **Unblock** if a firewall warning is displayed.
  7. Click **Finish** when the WSA sender is installed.
  8. Click **Exit** to close the installation.

## Command line installation

Alternatively, you can install the web service architecture using a command line.

- You can use a silent install with default settings:

```
msiexec.exe /i D:\WSASender\Kofax.Reporting.WSA.Sender.msi /quiet
```

- or preconfigure values for silent install:

```
msiexec.exe /i D:\WSASender\Kofax.Reporting.WSA.Sender.msi INSTALLDIR="C:\WSA\WSA-Sender" TRANSPORT_MODE=3 SERVICE_NAME="Kofax WSA Sender" /quiet
```

Use the following switches and parameters to control the installation:

- To modify the installation directory, use the parameter `INSTALLDIR`  
`INSTALLDIR=%ProgramFiles%\Kofax/Reporting/WSA-Sender`
- To modify the name of the Windows service used by WSA, use the parameter `SERVICE_NAME`  
`SERVICE_NAME="Kofax Reporting WSA Sender"`
- To modify the transport mode, use `TRANSPORT_MODE`  
`TRANSPORT_MODE=1`  
Supported values are:
  - 1: HTTP
  - 2: MSMQ
  - 3: IBM MQ
- To modify the name of the local server, use for example  
`LOCAL_SERVER_NAME="localhost"`
- To modify local port of the web service interface, use for example  
`LOCAL_PORT="25480"`
- To specify the central queue for MSMQ, use for example  
`CENTRAL_QUEUE_MSMQ="DIRECT=OS:.\private$\reporting"`
- To specify the central queue for IBM MQ, use for example  
`CENTRAL_QUEUE_IBMMQ="REPORTING.QUEUE.MANAGER:2REPORTING.QUEUE"`

- To specify the URL of the receiver WSA in the central site, use for example  
`CENTRAL_URL="http://localhost:25481"`
- To use a quiet mode installation, use `/quiet`

## Chapter 3

# Reported statistical data

Kofax Transformation Toolkit reports, depending on the module, various information to the Kofax Reporting platform. Some data, for example, how long a user utilizes a module and how long a batch is processed by a module, is gathered for every module. Other data, such as the number of keystrokes is gathered for all user interactive modules such as Validation or Verification and only, if defined in your workflow. In addition, some data is related to a specific module, such as the number of correct, missed and incorrect splits that are reported for Document Review in order to evaluate the document separation results.

### **Open Session**

- Appld
- MachineName
- ModuleId
- ModuleName
- SiteId
- SiteName
- Station
- UserId
- UserName

### **Create/Open Batch**

- Name
- Priority
- Class
- Description
- ClassDescription

The following topics give more details about important data that is gathered during processing used for a set of [predefined views](#).

## Server Processing

During Server processing the following data is gathered and reported to the reporting platform. For later reports a total sum of those counters per batch can also be provided.

### **Classification time**

The classification time is the time needed to determine the class a document and its pages belong to. In case document separation is configured the classification time is the time span to classify **and** separate the document.

### **Recognition (OCR) time**

The recognition time is the sum of all time spans for performing full text OCR on the pages of a document. Note that full text OCR can be performed in different processing steps, for example on all pages in Server step 1 and additionally OCR on demand in Server step 2.

### **Extraction time**

The extraction time is the time span to extract data, format field results and run validation rules for all fields on a document and its pages. Extract means that the configured locator and user-defined script extraction methods are performed. OCR recognition time is not considered.

### **Batch restructuring**

Any changes to the structure of the batch, such as deleting a document, adding a page or merging two documents that are applied via script are tracked and reported.

## Document Review

Document Review allows correcting document separation and classification results. Field data may be available, but is reported at this point of time.

For document separation the number of correct, missing and incorrect splits per document and batch are evaluated for the document separation results performed by the Server. After the separation results are validated in Document Review the counters are reported. Any structural changes for a batch, for example, splitting a document in Validation or other modules, do not change the value of those counters. The sum of all correct, missing and incorrect splits is equal to the number of all pages of all documents in the batch. Those counters are reported only if you perform document separation.

### **Correct splits**

A correct split is counted for a page that is either a correct first page that means it is recognized as first page by document separation and is still a first page after the review. Or, a page that is not a first page, which are all pages of a documents except the first one.

### **Missing splits**

A missing split is counted for a page that is not recognized as first page by document separation, but the document is split by the user during the review.

### **Incorrect splits**

An incorrect split is counted for a page that is recognized as first page, but which is no longer a first page as the corresponding document is merged to another document by the user during the review.

In Document Review you can define document review rules for a batch or at class level. For example, a batch must have 10 documents or documents that are assigned to a specific class must have at least 2 pages. In addition to the rules that can easily be configured or script your own rule by defining document review methods. During Server processing the rules and methods are checked and any failures are listed as problems to the Document Review user.

### **Batch problems**

This option is set to TRUE for a batch that has at least one batch rule failure.

### **Batch rule failure**

This counter returns the number of all failed batch rules for a batch. A batch rule failure is encountered in case a document review rule that is defined on project level is not passed. Those so-called batch rules can either be scripted methods or predefined rules.

### **Document problems**

This counter returns the number of documents that have failed at least one document review method (document method) or document review rule of the class.

### **Document rule failure**

This counter returns the number of all failed document review rules plus the number of documents that are unclassified. A document review rule is defined on class level and can either be a predefined rule or a scripted method that is defined in the project settings. A failure is encountered when either a document is unclassified or a document review rule for the class the document is classified to is not passed. Note a class may have several document review rules defined.

### **Document unclassified**

This counter returns the number of all unclassified documents in the batch. A documents stays unclassified if no classification result can be determined or the confidence is lower than the minimal confidence defined in the project settings in Project Builder.

**Note** The counters for folder problems and folder rule failures are currently not supported.

You can use Document Review also to correct the structure of a batch.

### **Batch editing and batch restructuring**

Any changes to the structure of the batch, such as deleting a document, rotating a page or merging two documents, are tracked and reported. These changes can be applied by the user or may be implemented and applied from the script.

### **Keystrokes**

Each key a user presses is counted. However, a combination using Ctrl, Shift and Alt keys are counted as one keystroke.

## Correction

Correction allows correcting fields that are extracted by an Advanced Zone Locator. The Advanced Zone Locator recognizes the characters that are found for the zones that are defined for a class on the document and applies a recognition status for each recognized character of the zone.

If the recognition confidence is equal to or higher than the defined threshold for a recognized character its recognition status is set to "confident" and if the recognition confidence less than the defined threshold for a recognized character its recognition status is set to "unconfident".

For each document the number of characters for the following statuses is reported.

**Confident and correct characters**

A recognized character is counted as "confident and correct" if it is confident, and the value is not changed during correction.

**Confident and incorrect characters**

A recognized character is counted as "confident and incorrect" if it is confident, but the value is changed during correction.

**Inconfident and correct characters**

A recognized character is counted as "inconfident and correct" if it is unconfident, but the value is not changed during correction.

**Inconfident and incorrect characters**

A recognized character is counted as "inconfident and incorrect" if it is unconfident and the value is changed during correction.

Additional counters for Correction are reported. The time span is gathered for each document whereas keystrokes are counted per field as well as per document.

**Time span**

The reported time span is the sum of time spent to correct each field of that document.

**Note** The time span is not reported for Thin Clients modules.

**Keystrokes**

Each key a user presses is counted. However, a combination using Ctrl, Shift and Alt keys are counted as one keystroke.

## Validation

Validation allows editing the results for extracted fields on a document for the assigned class. Fields can be either simple fields or table fields. For tables fields the data is gathered for each cell.

A field status is "valid" if a field is defined as "always valid" or the confidence of the extracted value is higher or equal to a defined validation threshold and the distance to the second best alternative is higher or equal to a defined value. In addition, all defined formatting and validation rules must be passed.

A field status is "invalid" if either no value is extracted or the confidence of the extracted value is lower than a defined validation threshold or the distance to the second best alternative is less than a defined value or if any formatting or validation rules have failed.

**Valid and correct**

When a field status is valid and the original value of this field is not changed during the validation step then the field is counted as valid and correct.

**Valid and incorrect**

When a field status is valid and during the validation step the automatically extracted field value is considered as incorrect and is changed then the field is counted as valid and incorrect.

### **Invalid and correct**

When a field status is invalid and during the validation step the automatically extracted value is considered as correct and is not changed then the field is counted as invalid and correct.

### **Invalid and incorrect**

When a field status is invalid and during the validation step the automatically extracted value is considered as incorrect and is changed then the field is counted as invalid and incorrect.

In addition the following statistical data is gathered for each document.

### **Time span**

The reported time span reflects the time it took to validate a document. That means the time when a document is loaded until it is saved.

**Note** The time span is not reported for Thin Clients modules.

### **Keystrokes**

Each key a user presses is counted. However, a combination using Ctrl, Shift and Alt keys are counted as one keystroke.

During Validation a user can use the batch edit mode to change the structure of the batch.

### **Batch editing and batch restructuring**

Any changes to the structure of the batch such as deleting a document, rotating a page or merging two documents, are tracked and reported. These changes can be applied by the user or may be implemented and applied from the script.

**Note** Batch edit operations, for example, when you merge two documents or split a document, are reported. However, if you split a document this does not change the counters for correct, missed or wrong splits. Those statistics refer to the document separation and its evaluation and are reported after you have validated a batch in Document Review.

## Verification

Verification allows validating extraction results for special fields for a second time to double-check its correctness. Currently, only general statistics are gathered for Verification, for example how long a batch is processed or the number of keystrokes per document or batch.

In addition the following statistical data is gathered for each document.

### **Time span**

The reported time span reflects the time it took to verify a document. That means the time when a document is loaded, all fields are verified, until it is saved.

### **Keystrokes**

Each key a user presses is counted. However, a combination using Ctrl, Shift and Alt keys are counted as one keystroke.

## Knowledge Base Learning Server Processing Step

Knowledge Base Learning Server is an unattended module that copies the documents marked for Generic and Specific Online Learning to the specified online learning folder for the current generation and adds the documents to the dynamic knowledge base. For reporting only general statics are gathered how long a batch is processed with the Knowledge Base Learning Server.



## Chapter 4

# Predefined views

The following predefined views are provided, which can be used to create reports. For more information about the existing tables and views and how to create new views and reports, see the Kofax Reporting SDK documentation.

**Note** Kofax Transformation Toolkit always reports data during the batch processing, so no additional activation is needed. However, some views are currently defined to retrieve data only for "finished" batches, which means those that are exported. If needed, you can adjust the views.

The following predefined views are provided.

### [v\\_batch\\_completion](#)

This view shows the current status of every batch.

### [v\\_tsf\\_user\\_perf](#)

This view exposes the KTM user performance, such as the number of documents or key strokes and the duration of a module session.

### [v\\_tsf\\_user\\_perf\\_by\\_mod](#)

This view exposes the time a user has spent for each user interactive module.

### [v\\_batch\\_session](#)

The view exposes statistics for each completed batch session.

### [v\\_tsf\\_class\\_vs\\_review](#)

The view exposes statistics for each page in order to replace the "Classification versus Review" report from KC Advanced Reports.

### [v\\_field\\_acc](#)

The view exposes statistics for each exported batch and each field in order to replace the " KTM Field Accuracy" report from KC Advanced Reports.

### [v\\_field\\_confidence](#)

The view exposes statistics for index field in order to replace the " KTM Field Confidence" report from KC Advanced Reports.

### [v\\_tsf\\_statistics](#)

The view exposes statistics for each document about field recognition accuracy in order to replace the " KTM Daily Statistics" report.

### [v\\_recog\\_acc\\_by\\_batch](#)

The view exposes statistics for a batch in order to replace the " KTM Recognition Timing by Batch" report.

#### **v\_recog\_acc\_grouped**

The view exposes statistics for a batch the field recognition accuracy grouped by group and field name order to replace the " KTM Recognition Accuracy Grouped" report.

#### **v\_recog\_timing\_by\_batch**

The view exposes statistics for each exported batch in order to replace the " KTM Recognition Accuracy Grouped" report.

#### **v\_doc\_volume\_by\_batch**

The view exposes statistics for each exported batch in order to replace the " KTM Document Volume by Batch" report.

#### **v\_doc\_volume\_grouped**

The view exposes statistics for each batch about the document volume grouped by the group value in order to replace the " KTM Document Volume Grouped" report.

#### **v\_recog\_acc\_by\_field**

The view exposes statistics about the recognition accuracy grouped by batch class and field name in order to replace the " KTM Recognition Accuracy by Field" report.

#### **v\_recog\_timing\_grouped**

The view exposes statistics for each exported document about the recognition time grouped per day or by group value in order to replace the " KTM Recognition Timing Grouped" and " KTM Recognition Timing per Day" reports.

#### **v\_recog\_acc\_by\_batch\_grouped**

The view exposes statistics about the recognition accuracy grouped by batch per day or by group value in order to replace the " KTM Recognition Accuracy Grouped" and " KTM Recognition Accuracy per Day" reports.

#### **v\_field\_recog\_acc\_grouped**

The view exposes statistics about the recognition accuracy grouped by batch, group value and field name in order to replace the " KTM Field Recognition Accuracy Grouped" and " KTM Field Recognition Accuracy per Day" reports.

#### **v\_tsf\_class\_acc**

The view exposes statistics for each exported document about the classification accuracy.

#### **v\_tsf\_sep\_acc\_by\_batch\_class**

The view exposes statistics for each exported batch about the document separation accuracy.

#### **v\_tsf\_extr\_acc\_grouped**

The view exposes statistics about the extraction accuracy grouped by batch, group value and field name.

#### **v\_sla\_monitoring**

The view exposes statistics about the total processing time for each batch.

## View v\_batch\_completion

This view shows the current status of every batch.

This predefined view returns the following information for a batch.

- Batch creation date and time (UTC, granularity: seconds)
- Batch name
- Batch ID, a unique number assigned to the batch by Kofax Capture
- Batch export status: 0, if the batch is not exported, otherwise 1
- Batch deleted status: 0, if the batch is not deleted, otherwise 1
- Batch state: can either be ready, suspended, closed, completed
- Last module that processed the batch
- Last processing date and time (UTC, granularity: second)

## View v\_tsf\_user\_perf

This view exposes the user performance, for example, the number of processed documents, key strokes and the session duration. It holds one record for each attended session on a document with index fields. Additional grouping and summarizing must be done by the report author.

- User name
- Session end date (UTC, granularity: seconds)
- Number of key strokes
- Unique document key
- Unique document ID
- Unique batch key
- Unique batch ID
- Batch name
- Number of index fields for this document
- Processing time in seconds

## View v\_tsf\_user\_perf\_by\_mod

This view exposes the time a user spent in every user interactive module. It holds one record per attended batch session. Additional grouping and summarizing must be done by the report author.

- User name
- Session end date (UTC, granularity: seconds)
- Module name
- Time spent in module (seconds)

## View v\_batch\_session

The view holds one record for each completed batch session. Additional grouping and summarizing must be done by the report author.

- Module name
- Batch session end date and time (UTC, granularity: seconds)
- Unique batch key
- Unique batch ID
- Batch name
- Number of completed documents
- Number of key strokes
- Batch processing time in seconds
- Number of pages
- Number of documents
- Milliseconds spent for classification and separation during Server processing
- User name
- Unique batch session key that can be used to join the view with v\_doc\_session

## View v\_tsf\_class\_vs\_review

The granularity of this view is a single index field (for table fields, there may be several records with the same field name).

- Batch name
- Document Review operator name
- Classification result from server processing
- Classification result after Document Review
- Review date (UTC, granularity: seconds)
- Classification date (UTC, granularity: seconds)
- Confidently classified (TRUE or FALSE)
- Classification confidence

## View v\_field\_acc

For every exported batch, the view provides one record per named field. It compares the result of automatic classification and extraction during server processing with the result of the validation process.

The number of single character changes is calculated via a Leveshtein distance algorithm implemented in function `fn_chars_changed`.

**Note** As this calculation must be done for every single field, the performance of this query may not be very good.

- Field name
- Batch export date and time (UTC, seconds granularity)
- Batch name
- Number of documents
- Number of fields
- Number of automatically recognized characters
- Number of confidently recognized characters
- Percentage of confidently recognized characters
- Number of validated characters
- Number of changed characters (calculated via `fn_chars_changed`)
- Percentage of changed characters (based on validated characters)

## View `v_field_confidence`

The granularity of this view is a single index field (for table fields, there may be several records with the same field name).

- Field name
- Recognition date and time (UTC, granularity: seconds)
- Batch class
- Document class
- Form type
- Percentage of confident characters

## View `v_tsf_statistics`

This view can be used to replace the Statistics Viewer "KTM Daily Statistics" report that provides a summary of document and page traffic and field recognition accuracy.

The view holds one record per document. Additional grouping and summarizing must be done by the report author.

This predefined view returns the following information for each document.

- Last processing date and time (UTC, granularity: second)
- KTM group value
- Number of pages
- Number of valid and correct fields
- Number of valid and incorrect fields

- Number of rejected and correct fields
- Number of rejected and incorrect fields
- Total number of fields

## View v\_recog\_acc\_by\_batch

This view can be used to replace the Statistics Viewer "KTM Recognition Accuracy by Batch" report that provides the average recognition accuracy of a field within a batch. The view is grouped by batch and field name. The view is based on field detail data.

**Important** Field details are only kept in the database for a limited time. By default, the value for the time is set to 5 days and can be changed in the ETL configuration.

- Batch name
- Field name
- Percentage of valid and correct fields
- Percentage of valid and incorrect fields
- Percentage of rejected and correct fields
- Percentage of rejected and incorrect fields
- Total number of fields

## View v\_recog\_acc\_grouped

This view can be used to replace the Statistics Viewer "KTM Recognition Accuracy Grouped" report from the Actual Reports group that provides the field recognition accuracy per group and field name.

The view is grouped by group value and field name.

- KTM group value
- Field name
- Number of valid and correct fields
- Number of valid and incorrect fields
- Number of rejected and correct fields
- Number of rejected and incorrect fields
- Total number of fields

## View v\_recog\_timing\_by\_batch

This view can be used to replace the Statistics Viewer "KTM Recognition Timing by Batch" report from the Actual Reports group.

It holds one record for each exported batch.

- Batch name
- Batch export date and time (UTC, granularity: seconds)
- Number of documents
- Number of pages
- Total seconds spent in OCR
- Total seconds spent in extraction

## View v\_doc\_volume\_by\_batch

The Statistics Viewer "KTM Document Volume by Batch" report from the Actual Reports group can be replaced by either the v\_recog\_timing\_by\_batch or from v\_doc\_volume\_by\_batch view.

This view holds one record for each exported batch.

- Batch name
- Export date and time (UTC, granularity: seconds)
- Number of documents
- Number of pages

## View v\_doc\_volume\_by\_grouped

This view can be used to replace the Statistics Viewer "KTM Document Volume Grouped" report from the Actual Reports group.

It shows the document volume grouped by group value for each batch.

- KTM group value
- Batch name
- Export date and time (UTC, granularity: seconds)
- Number of documents
- Number of pages

## View v\_recog\_acc\_by\_field

This view can be used to design the historical report Statistics Viewer "KTM Recognition Accuracy by Field" that provides aggregated data grouped by batch class and field name.

- Batch class name
- Field name
- Percentage of valid and correct fields

- Percentage of valid and incorrect fields
- Percentage of rejected and correct fields
- Percentage of rejected and incorrect fields
- Total number of fields

## View v\_recog\_timing\_grouped

This view can be used to replace the historical report Statistics Viewer "KTM Document Recognition Timing Grouped" or "KTM Document Recognition Timing per Day " that holds one record for each exported document. Additional grouping and summarizing must be done by the report author.

- Batch class name
- KTM group value
- Batch export date and time (UTC, granularity: seconds)
- Unique document key
- Number of pages
- Total OCR time in seconds
- Total extraction time in seconds

## View v\_recog\_acc\_by\_batch\_grouped

This view can be used to replace the historical report Statistics Viewer "KTM Recognition Accuracy Grouped " and "KTM Recognition Accuracy per Day ". It holds one record for each combination of batch id and group value. Additional grouping and summarizing must be done by the report author.

- Batch class name
- Batch name
- KTM group value
- Export date and time (UTC, granularity: seconds)
- Number of valid and correct fields
- Number of valid and incorrect fields
- Number of rejected and correct fields
- Number of rejected and incorrect fields
- Number of fields
- Number of documents
- Number of pages



## View v\_field\_recog\_acc\_grouped

This view can be used to design the historical report Statistics Viewer "KTM Field Recognition Accuracy Grouped " and "KTM Field Recognition Accuracy per Day ". It is grouped by batch ID, group value and field name. Additional grouping and summarizing must be done by the report author.

- Batch name
- Batch class name
- KTM group value
- Field name
- Export date and time (UTC, granularity: seconds)
- Number of valid and correct fields
- Number of valid and incorrect fields
- Number of rejected and correct fields
- Number of rejected and incorrect fields
- Number of fields

## View v\_tsf\_class\_acc

This view exposes classification accuracy. It covers only documents that were successfully exported. Additional grouping and summarizing must be done by the report author.

- Batch class name
- KTM group value
- Classification date and time (UTC, granularity: seconds)
- Unique document key
- TRUE, if document correct and confidently classified, else FALSE
- TRUE, if document correct and not confidently classified, else FALSE
- TRUE, if document incorrect and confidently classified, else FALSE
- TRUE, if document incorrect and not confidently classified, else FALSE

## View v\_tsf\_sep\_acc\_by\_batch\_class

This view exposes document separation accuracy, grouped by batch class and month of batch creation. It holds one record for each successfully exported batch. Additional grouping and summarizing must be done by the report author.

- Batch class
- Batch creation date and time (UTC, granularity: seconds)
- Batch name
- Number of correct splits

- Number of wrong splits
- Number of missed splits
- TRUE, if batch has wrong or missed split points, else FALSE.

## View v\_tsf\_extr\_acc\_grouped

This view exposes extraction accuracy of a field grouped by batch class, group value and field name. Additional grouping and summarizing must be done by the report author.

- Batch class
- Unique batch ID
- KTM group value
- Field name
- Batch name
- Export date and time (UTC, granularity: seconds)
- Number of missed splits
- Field name
- Number of fields
- Number of characters
- Number of valid and correct fields
- Number of valid and incorrect fields
- Number of invalid and correct fields
- Number of invalid and incorrect fields
- Number of confident and correct characters
- Number of confident and incorrect characters
- Number of unconfident and correct characters
- Number of unconfident and incorrect characters

## View v\_sla\_monitoring

This view contains one record per batch. It shows the total processing time for not completed batches, batch completion status, number of documents, number of pages, batch class, station and operator from batch creation.

- Batch class
- Unique batch ID
- Batch name
- Creation time of oldest document (UTC, granularity: seconds)
- FALSE, for batches that are not exported; TRUE, for batches that are exported
- Numeric batch status
- Scan user

- Scan site
- Scan station
- Total processing time of this batch
- Number of documents
- Number of pages