Oracle® WebCenter Forms Recognition

Web Verifier User's Guide

11g Release 1 (11.1.1.8.0)

E50189-01

November 2013

ORACLE®

WebCenter Forms Recognition

11g Release 1 (11.1.1.8.0)

Contents

1 About Web Verifier

1.1 Verifier vs. Web Verifier

Web Verifier is a web-based extension of the Verifier thick client. It is an application module to allow users to verify documents with no software installed on the client side. The processing of activities primarily occurs on a central server allowing use of thin client devices running only a web browser.

This is an outstanding advantage for your business: keeping only a few servers busy and the workstations running the Web Verifier results in easier system management and lower costs. Furthermore, this solution allows for central management of storage and backup and provides for easier security.

Thus, the Web Verifier solution will contribute to considerable reduction of the cost and personnel-intensive maintenance of several workstations.

Another significant advantage is higher flexibility. As the only thing you need to work on your project is a workstation running a web browser and an Internet connection, you are able to access your project or the current batch list from everywhere. Furthermore, the system is able to remember the application state after the user's last session. When the user starts the application next time, the system will load the last application state. As this information is stored in the database on the central server (as user level project settings), access to it via authentication using your login information from every device irrespective of the operating system will be possible.

In addition, users of the Web Verifier module have the option to apply lockers to resources for exclusive use. Thus, no other module or user will be able to use this resource. However, in order to prevent dead locks, each locker is provided with a timeout. This means that upon exiting the application, the system will remove all user owned lockers.

Key Features

- Allows convenient correction of automatic classification results.
- Allows convenient correction of automatic extraction results.
- Allows manual indexing of documents.
- Allows semi-automatic indexing of documents by means of database lookups.
- Allows a final check of corrected documents before release.

Highlights

- The structured user interface makes the application easy to learn.
- Sophisticated status management and filter techniques show you only the documents you have to check, and nothing else.
- During the application design, the user interface can be configured, providing optimum display options for each document category.
- Through automatic locking, document batches can safely be processed by teams of operators.

2   About This Manual

2.1   Intended Audience

This manual is primarily for users of WebCenter Forms Recognition Web Verifier. To use this application and its documentation, you need no special skills. Application designers and system administrators will also find this manual useful.

2.2   Related Documentation

In addition to this manual, WebCenter Forms Recognition comes with the following documentation:

- WebCenter Forms Recognition Verifier User Guide

- WebCenter Forms Recognition Installation Guide: Explains how to install WebCenter Forms Recognition.

- WebCenter Forms Recognition Designer User Guide: Explains how to use the WebCenter Forms Recognition Designer component to create custom application.

- WebCenter Forms Recognition Runtime Server User Guide: Explains how to use the WebCenter Forms Recognition batch processing application.

- WebCenter Forms Recognition Scripting Guide: Explains how to write scripts to enhance WebCenter Forms Recognition projects.

3   Understanding How Web Verifier Works

3.1   Some Terms You Should Know

3.1.1.   What Is a Batch?

A batch is just a stack of documents. Usually, this stack is not sorted. In the context of WebCenter Forms Recognition, batches consist of electronic documents. The documents inside such a batch are typically paper-based documents that have been scanned to transform them into a digital format. Various documents are normally assigned to the same batch only because they have been received within the same time period. For example, all invoices received in the morning may be scanned until noon and therefore end up in the same batch.

3.1.2.   What Is a Folder?

In a business environment, folders are normally used to keep several documents together. WebCenter Forms Recognition does the same thing with folders. However, in the context of WebCenter Forms Recognition, a folder is always a structure inside a batch. This means that batches can either consist of document stacks, or they can consist of stacks of folders.

3.1.3.   What Is a Document?

A document is a piece of information that can serve as evidence of an event, situation, or business transaction. For example, a packing slip may provide evidence that an order has actually been shipped. Since people are used to working with paper, electronic documents strongly resemble paper-based documents. You will notice that WebCenter Forms Recognition documents consist of one or several pages, though the concept of a page is not really required for digital documents.

3.1.4.   What Is Classification?

Classification means taking an unsorted stack of documents and organizing them into smaller stacks so that each stack contains only documents belonging to the same category. In other words, you start with a mess and end up with an organized stack of invoices or other documents. Class and category are the same thing.

3.1.5.   What Is Indexing?
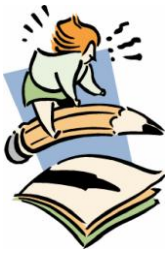
Imagine you have a homogeneous stack of invoices, and you start to write out the information contained in the documents. For each document in the stack, you will note the name of the supplier, the total sum to be paid, and the invoice number. This procedure is called indexing, and the information that was noted is the indexing information. Once you are finished, you file the invoices and use the indexing information to build your filing structure. Later, you will be able to search and identify the document with the help of the indexing information. In the context of WebCenter Forms Recognition, indexing information is applied to a set of fields associated with the document. For each document category, a different set of fields can be used.

3.1.6.   What Is Extraction?

If you take the stack of invoices and again write out the name of the supplier, the total sum to be paid, and the invoice number, but this time automatically, the procedure is called extraction. Extraction is a means for automatic document indexing. Extraction is context-sensitive; that is, the extracted information depends on the document category.

3.1.7.   What Is a State?

A state is a number that tells you how far the processing of a document has progressed. If the entire procedure of document processing consists of single steps, then the state increases with each step that has been completed. The state also indicates whether a step has been completed successfully, or whether there have been problems. In WebCenter Forms Recognition, states are determined hierarchically from the bottom up: If anything is wrong with a document, then there is also something wrong with the batch it belongs to.

3.1.8.   What is Verification?

Verification is a task related to quality assurance. It involves taking a document that has been processed or partially processed, checking the processing results, and correcting any errors.

3.1.9.   What Is Validation?

Validation is another task related to quality assurance. Validation means confirming that a processing result is correct. This can be done at several levels: for the class or a field associated to a document, for the document as a whole, or for an entire batch.

3.1.10. What Is a Learnset?

In classification, a Learnset is a group of documents whose classification is specified by a user. For each view and each class, the user must provide a sufficient number of representative documents. Similarly, in extraction, a Learnset is a set of documents whose field contents are selected by the user from a set of candidates.

3.2 WebCenter Forms Recognition and Quality Assurance

To properly ensure the quality of automatically processed documents, there are two things you need to understand:

Note: *Batches are the basic entity on which WebCenter Forms Recognition works*

WebCenter Forms Recognition works on batches. Tasks consist of processing steps that must always be completed for an entire batch before the next task can start.

For example, if Batches #9, #10, and #11 are waiting to be classified, the application will first classify all documents in Batch #9. If this is done, the state of Batch #9 is incremented. The next task may be to classify all documents in Batch #10, or it may be to extract data from all documents in Batch #11.

What the application will not do is to classify some documents from Batch #9, then some documents from Batch #10, and then go back to Batch #9 to classify the remaining documents.

If you find this hard to understand, just imagine that you are a mechanic repairing cars in a garage. Your customers always have to leave their entire car, not just the engine and the front seats. In addition, to avoid confusion, the mechanics at your garage have been instructed to complete one job before they start the next one. At your particular garage, the mechanics are allowed to repair the engine of one car and then the brakes of the next car, but they are not allowed to repair half of the engine, interrupt this for another job, and then go back to the engine again.

If batches are the basic entities, then entire batches need to be verified and approved before they are routed to subsequent systems where other users or processes work with them.

Note: *A batch is only valid if all of its parts are valid*

Imagine your car has broken down because a single critical component in the engine has failed. If the engine doesn't work properly, you cannot drive your car.

Similarly, a WebCenter Forms Recognition batch is valid only if all documents and processing results associated with the batch are valid. Because we are dealing with information and data, we do not use the terms "working" or "damaged." Instead, we use the terms "valid" or "invalid."

Like cars, WebCenter Forms Recognition batches consist of a restricted number of "parts" with well-defined relationships. Therefore, we can easily see why a batch can be invalid.

A batch is invalid if:

- One or more folders inside the batch are invalid.

   A folder is invalid if:

- One or more documents inside the folder are invalid.

   A document is invalid if:

- It has been classified automatically, but the classification result is invalid, or

- Data has been extracted automatically from it, but at least one or more fields are invalid.

  A classification result is invalid if:

- No matching class could be found, or

- The class has been changed manually and not yet validated.

  A field is invalid if:

- The field could not be filled,

- The field content does not comply with validation rules that have been defined, or

- The field content has been changed manually and not yet validated.

  Field validation rules may be violated for a number of reasons:

- The set of allowed characters may be restricted.

- Only uppercase characters may be allowed.

- There may be restrictions on the number of characters the field can contain.

- Your WebCenter Forms Recognition application may enforce that characters which could not be certainly identified during the OCR must be checked. These questionable results are indicated in red and are underlined.

- Besides these formal validation rules, all kinds of custom rules are possible. For example, if the contents of Field 3 do not equal the sum of Field 1 and Field 2, then Field 3 may be invalid. Such a rule will typically be applied for invoices.

The application will normally tell you why a field is invalid.

3.3   WebCenter Forms Recognition and Teamwork

In WebCenter Forms Recognition, the flow of incoming documents follows a sequence of standard processing steps. Some steps can be skipped, but the order of steps is fixed.

Automatic steps are executed by WebCenter Forms Recognition Runtime Server and include document import with batch creation, OCR and layout analysis, classification, extraction, export, and cleanup[1]. These automatic steps are completed with two manual verification steps that ensure only high-quality output is produced: verification of the classification and extraction steps.

If WebCenter Forms Recognition Runtime Server has completed an automatic step and the batch contains only valid results, the next automatic step can be accomplished without human intervention.

However, if WebCenter Forms Recognition Runtime Server detects that the batch contains invalid results, the batch must be routed to a verification station, where you, the operator, can analyze and resolve the problem using Verifier or Web Verifier. Invalid batches are presented to you in a task list, the so-called Batch View. (See section 8.1 THE BATCH VIEW) You will have to resolve each problem and validate each correction before you can release the batch. Subsequent automatic steps can be carried out only after release. Finally, when WebCenter Forms Recognition has finished processing a batch, the documents will be sent to their actual recipients.

One of the objectives of a WebCenter Forms Recognition application is to get documents to their recipients as quickly as possible. On the machine side, automatic steps can be distributed to several computers to ensure that no delays occur. They can simultaneously

---

[1]   This step deletes files from the hard drive that are no longer required.

perform the same or different tasks. Similarly, on the human side, WebCenter Forms Recognition supports a variety of task distributions in a team. For example, there can be specialized workstations, where one station's operator is only in charge of classification results and the other station's operator verifies the extraction results. This can be realized by configuring WebCenter Forms Recognition Verifier or Web Verifier accordingly. (See section 7.3 SETTINGS – WORKFLOW) In addition, several operators can carry out the same task at the same time, but on different batches. This is possible via a locking mechanism that avoids conflicting results by making sure that several people cannot change a batch at the same time.

As a practical example, consider two servers with WebCenter Forms Recognition Runtime Server, called Runtime_Alpha and Runtime_Beta, which share the time-consuming task of Optical Character Recognition (OCR). A third server, Runtime_Gamma, is in charge of the remaining automatic steps. If Runtime_Gamma generates an invalid classification result, the corresponding batch is routed to a member of the QA team named Miller. Miller corrects these results using the WebCenter Forms Recognition Verifier or Web Verifier instance running on Verifier_Miller.

Normally, correcting invalid extraction results requires more effort than correcting invalid classification results. Therefore, three other members of the QA team — Barnes, Hill, and Dawson — share a common worklist containing batches with invalid extraction results. If either of them starts processing a batch, this batch will be locked for the others. WebCenter Forms Recognition Verifier sets a corresponding marker in the worklist.
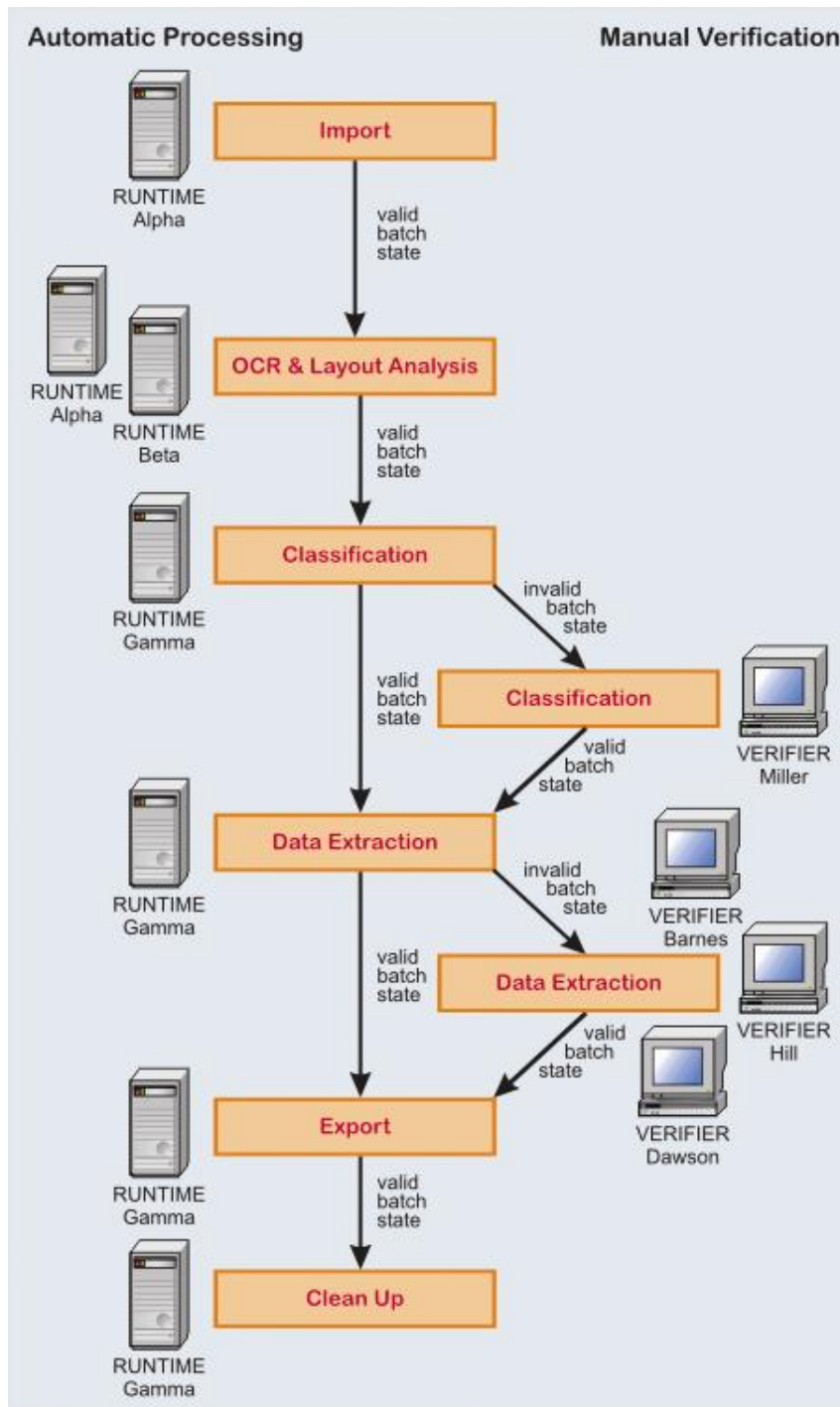
Figure 3-1: Sample workflow showing how processing steps can be distributed.
This configuration involves several servers running WebCenter Forms
Recognition Runtime Server, and a number of workstations with WebCenter
Forms Recognition Verifier that are used by a quality assurance team.

4   Installation

If your workstation is intended to run only the Web Verifier, no installation steps at your workstation are necessary. All you need is a web browser, an Internet connection and your login information.

For information about supported browsers please refer to the Installation Guide.

Your administrator will create a user account for you for the WebCenter Forms Recognition context using the Designer application, and will assign at least one role to you. (Please refer to 6 ABOUT USERS, GROUPS, AND ROLES)

In context of the Web Verifier, all work associated with installation, user assignment, project design and database creation is done at a central server. This should be done by an experienced user. For instructions, please refer to the Installation Guide.

4.1   Supporting Languages

The Web Verifier application implements full support of CJKT. Among other advantages this means that the user can review and edit/ validate the extraction results using multiple non-western languages at a time.

Searching via the Associative Search engine, for example, can be applied to a pool that contains multi-language entries. In this case, the user can even search for a mixed text (with a couple of non-western words or characters typed together within one search query), but in this case the "hierogliphycal" pool items may obtain better weighting compared with the other languages. This is however specific for Associative Search Brainware V3 based processing.

This applies to the following workflow steps in WebCenter Forms Recognition:

- Classification
- Extraction
- Validation
- Learning

5   Starting and Exiting Web Verifier

5.1   Starting

If Web Verifier was installed as recommended by the setup program, you can launch it from your web browser via the server/application address given to you by your administrator. After startup and login, the application displays a window where you can select a project and the respective job (See 8 GETTING FAMILIAR WITH THE USER INTERFACE).

Note: *If you start the WebCenter Forms Recognition Web Verifier on a machine for the first time, it might happen that not all of the controls are displayed. Then, you have first to add the Web Verifier site to the trusted sites list of your Internet browser.*

Note: *From version 11.1.1.8.0, the Sax Basic scripting engine is no longer supported. If your application is configured to open a project that uses this engine for custom scripts, you will be presented with an error message:*

```
Project XXX.sdp is using an unsupported Sax Basic engine and
cannot be used. Please open this project in the Designer
application to upgrade to the newer engine for custom scripting
(WinWrap Basic).
```

Click *OK.* This will close the Verifier application.

You will need to open the project in the Designer application and to migrate it to use the WinWrap scripting engine.

5.2   Logging In

To log into the Web Verifier, start an Internet browser and open the Login page of the Web Verifier application. The project administrator will provide this address.

The login for Web Verifier is configured by the project administrator using either the Web Verifier Login or Windows Authentication.

5.2.1.   Web Verifier User Login

When you log in to an existing project in Web Verifier, supply your WebCenter Forms Recognition user name and password. Your user name and password were assigned to you by your project administrator.

After a user first login a dialog window may open for a password change if it has been required by the administrator. The user is obliged to change his password before he can use the application.

The user name and password are used for the login to the Web Verifier platform. The user name and password also set the user rights for the active session. The current users name is displayed in the top-right corner of the Web Verifier window.

Note: *If you have questions or problems with your user name or password, please contact your project administrator. If you forget your password, your administrator can reset it for you.*

Your project administrator can give you the option to remember your username and password between logons. This has been enabled if the *Remember me next time* checkbox appears on the logon form. To remember your username and password between logons, fill in your username and password and select *Remember me next time* before clicking *Log In*. Next time you log on with the same computer, the system will fill in the username and password automatically. Select the language for the application from the dropdown list on the bottom right corner.
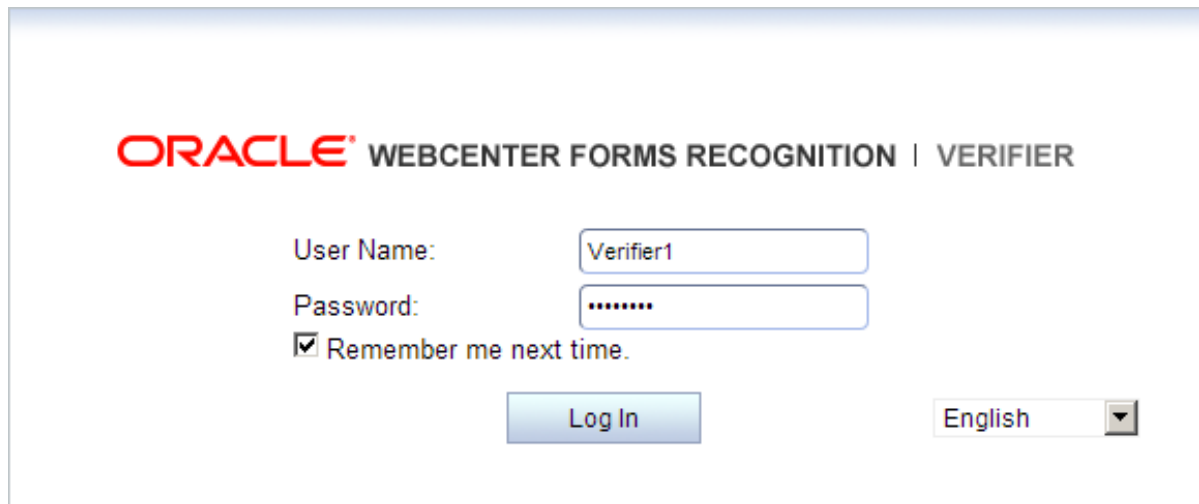
Figure 5-1: Web Verifier logon with password remember feature

For more information about user accounts and roles, see 6 ABOUT USERS, GROUPS, AND ROLES.

5.2.2. Windows Authentication

Note: In order to use Windows Authentication, the Web Verifier service needs to be configured by the administrator to allow Windows Authentication. Please check the Installation Guide for information on how to do this.

To Login using Windows Authentication:

1. Open the start page of the Web Verifier application. If the current computer user is configured for Web Verifier Windows Authentication by the project administrator, the Web Verifier interface should open. The Windows Authentication Form may appear.

2. Fill in your Windows credentials if the Windows Authentication Form appears.

When logged in with Windows Authentication, the Verifier window shows the Domain and Computer name in the top right corner.

Windows Authentication Login for Web Verifier shows following special behavior:

- A cancelled or failed authentication opens an error page. Press the reload button to open the Login window again.

- To switch to another user, e.g. the Administrator, open the File menu and click Re-login. The standard Web Verifier Login Window opens (see section 5.2.1).

- The Windows Authentication mode prevents the change of the password.

- Logout link terminates the session on the server. Reloading a Web Verifier page in the same browser instance will create a new session on the server for the current user.

- In order to completely logout from the system, it is necessary to close all open browser sessions. Otherwise, returning to starting page would automatically restore the last session. A message screen at the end of the logout step will remind you of this:

  Thank you for using Web Verifier.

  To complete the logout, you may now exit all opened browser windows.

5.2.3.   Login issues

5.2.3.1. Session timeout

To maintain the workflow within several Verifier users a Web Verifier session ends automatically after 20 minutes of no activity. All batches are released and lose their changes that were applied during the last session. Any action in the Web Verifier after a timeout will open the login screen with following information:

> *Your session is currently expired due to either a log out, or a timeout of 20 minutes. Please log on once more.*

5.3   Exiting the Application

To quit WebCenter Forms Recognition Web Verifier:

- Click *Logout* on the top right corner

- Close your web browser.

6   About Users, Groups, and Roles

To load a project, you must log in with your user name and password.

The user name/password combination not only lets you into a project, it governs what you can do once you get there.

User name/password combinations were set up by your project administrator.

Your administrator also set up user groups and assigned you to one or more of these groups. In turn, the administrator assigned one or more roles to each user group. User role is always more prominent than group role.

There are six roles: Administrator, Learnset Manager, Supervised Learning Verifier, Verifier, Verifier Settings and Verifier Filtering:

- **Administrator:** The Administrator role is to manage users, groups, and user-to-group assignments. Administrators install the system, configure applications, and manage data. They also design and maintain projects. This role is the most powerful of the six roles, because it encompasses the permissions for all other roles.

- **Learnset Manager:** The Learnset Manager role is to define, modify, and maintain the Learnset.

- **Supervised Learning Verifier:** The Supervised Learning Verifier role is to collect and manage local training data. Supervised Learning Verifiers are subject-matter experts who can propose Learnset candidates to improve system performance.

- **Verifier:** The role of the Verifier group is to verify documents that could not be automatically processed.

- **Verifier Settings:** The role of the Verifier Settings group is to allow the WebCenter Forms Recognition Verifier configuration to be changed.

- **Verifier Filtering:** The FLT role is to allow Verifier user to configure custom filtering of batches. By application design, FLT users would be able to use the filtering feature even if they do not have the SET role.

6.1   Changing Your Password

1) Select *Change Password* from the Options menu.

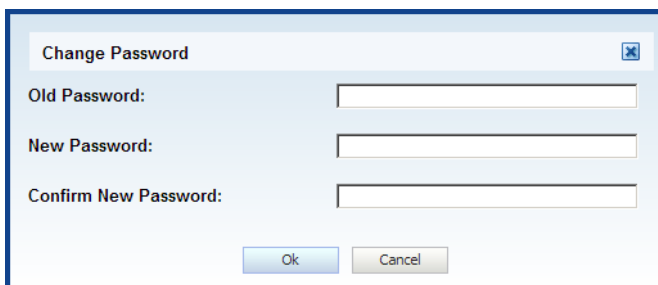*Figure 6-1: "Change Password" dialog*

2) Type in your existing password in *Old Password*.

3) Enter a new password.

4) Re-enter the new password for verification.

5) Click *OK*.

For users logged in via Windows Authentication, the *Change Password* option is greyed out.

7   Configuring Web Verifier

Configuring WebCenter Forms Recognition Web Verifier require specifying which batches of documents will be processed at a given station. This includes:

- The WebCenter Forms Recognition Designer project file that contains the settings used to process the documents. The name of the job that contains the batches to be verified.

- The processing steps that you want to verify: classification, extraction, document separation or all of them.

- The status of batches before and after processing.

Note: You can only work with WebCenter Forms Recognition Web Verifier after these settings are established. Only experienced users should change the settings.

To configure WebCenter Forms Recognition Web Verifier:

- On the Options menu, select *Settings*. Or

- Click *Settings* [icon] in the toolbar.

This displays the WebCenter Forms Recognition Web Verifier Properties.

7.1   Settings – General

For general settings such as referenced project and job, select the General tab. You will be presented with the *My Settings* view. When the Administrator creates users, he usually assigns them to one of the available groups with specific roles assigned. Depending on the group the current user belongs to, the following *Settings* views are possible:

- For a default **Administrator user** with all of the available roles assigned, the *Settings* window will look as follows:
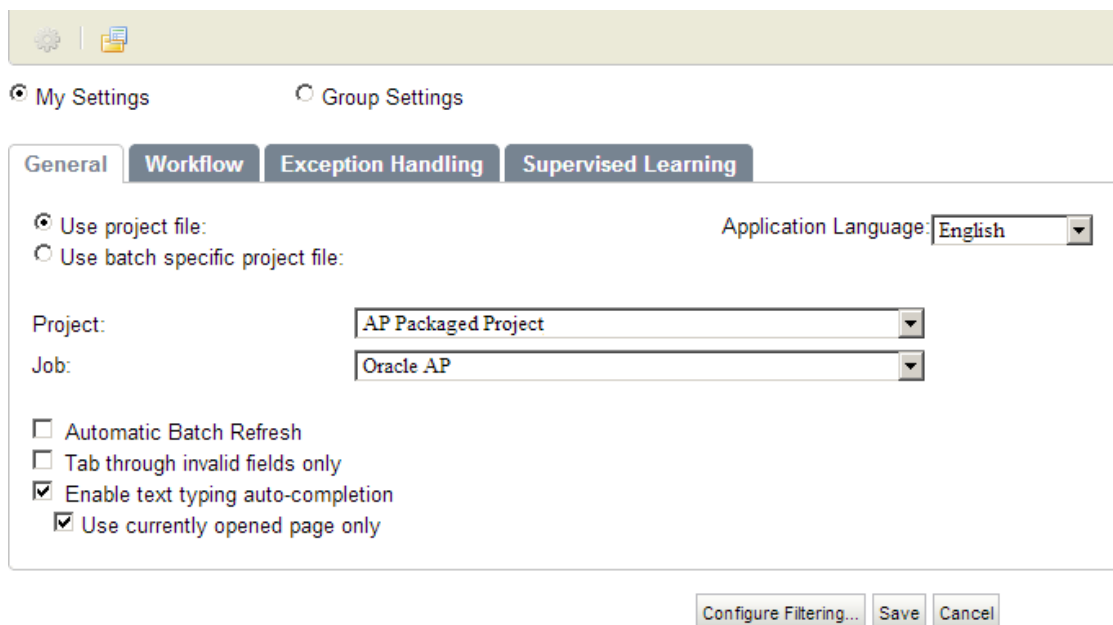
Figure 7-1: General Tab – view for Administrator users

Please refer to section 7.2 for further information.

- Learnset Manager users with the **SLM and SLV roles** assigned, but without the ADM role will be presented with a *Settings* view like shown below. The Group Settings are

not accessible, but the Supervised Learning tab is available. Please refer to section SETTINGS – SUPERVISED LEARNING for further information.
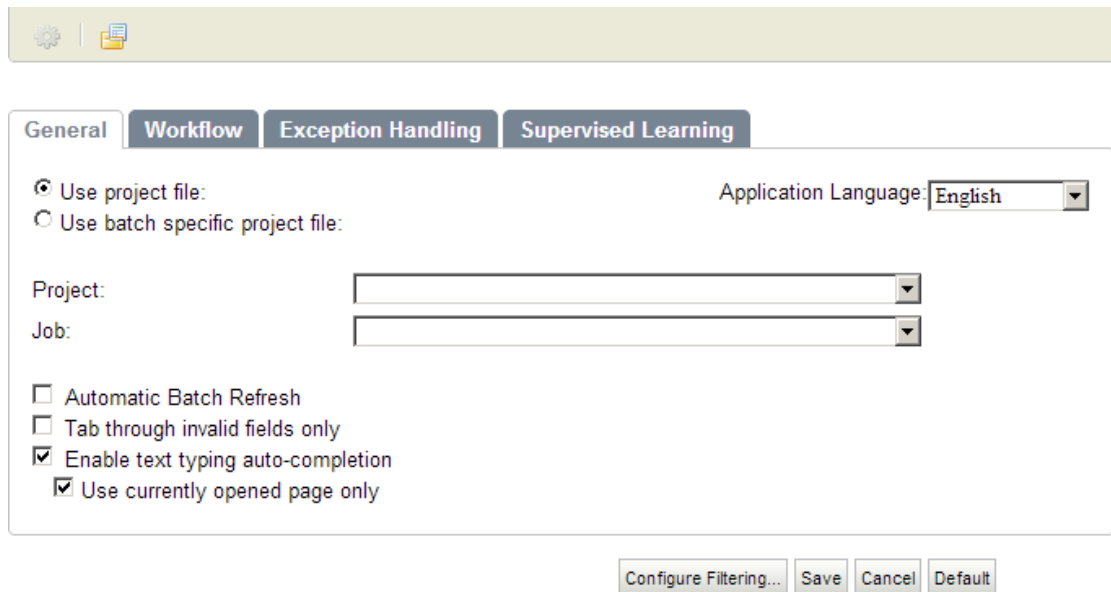
Figure 7-2: General Tab – view for SLM users

- For a default **Verifier user**, with the VER, SET, and FLT roles assigned, the *Settings* window will look as shown below. No Supervised Learning tab is available, and they are not able to edit Group Settings.
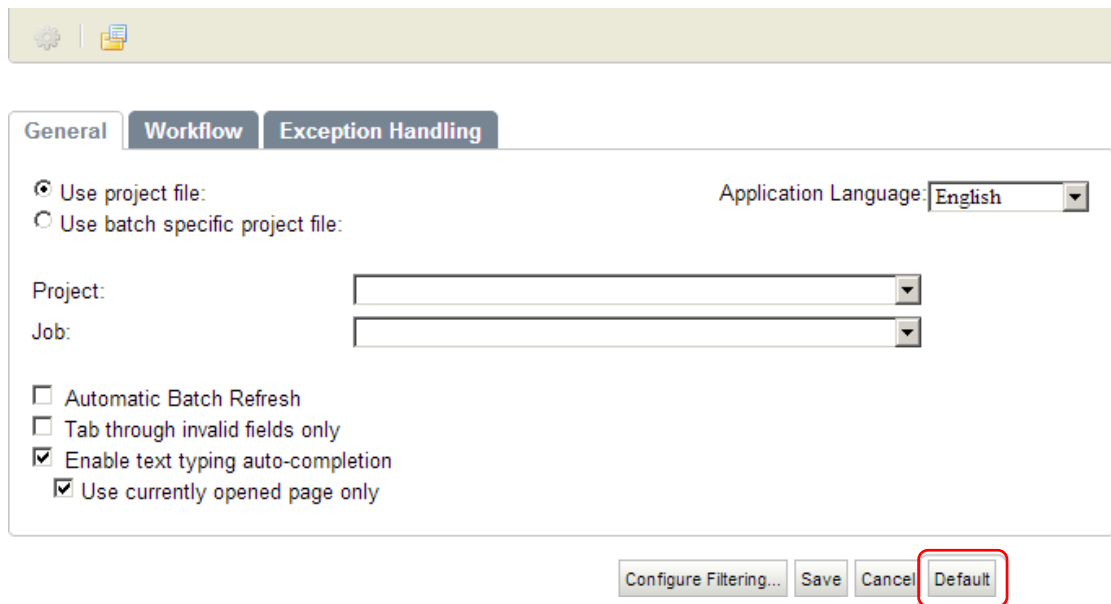
Figure 7-3: General Tab – view for Verifier users after Default button applied

Usually, the Administrator will already have configured the Group settings with the project selected. To adopt all settings for your group, click on the *Default* button (please see screenshot above). The fields in all available tabs will be populated automatically.

If you want to configure your own settings, please find the descriptions for the single configuration options as follows:

Project File:

- **Use Project File:** that is used to select the file name of the WebCenter Forms Recognition project used to process the documents; this contains the design of the indexing form that you will use to verify the extraction.

The following scenarios are possible for the project loading procedure:

1. When you log into Web Verifier for the first time, and the project is already configured for you, you will be prompted that the project is loading.

2. When you select a new project and press *Save* to save the settings made, the project selected will be loaded after you returned to the Batch View and press *Show Batches*.

3. If you are using the setting *Use Batch Specific Project File*, the project will be loaded as soon as you attempt to open a batch (and each time a batch is opened for a different project).

Application Language:

The application language can be changed here.

Web Verifier supports an extended list of languages:

Chinese Simplified, Chinese Traditional, Danish, Dutch, English, French, Finnish, German, Italian, Japanese, Korean, Norwegian, Polish, Portuguese, Romanian, Russian, Spanish, Swedish, Turkish.

Note: By default, the Chinese language selection points to Traditional Chinese. In case Simplified Chinese language has to be used, please contact your system administrator.

Automatic Batch Refresh:

- If this option is checked, the Batch View automatically shows newly generated batches with matching states. If you do not want the automatic update, you can clear the checkbox. This leaves you the option to refresh the Batch View (Click *Refresh* under the Batch list) when you need up-to-date information.

Tab through invalid fields only:

- When the user presses TAB in Document Verification mode, the system tabs through invalid fields only. Similarly, when the user presses TAB inside of a table control, the system tabs through invalid table cells only.

Enable text typing auto-completion:

- User can enable this option to speed up typing. When you start to type, auto-text completes the word, suggesting the best match among all of the words or candidates available after OCR and Format Analysis. For example, you can type the first two characters of a 20-character invoice.

- User can also enable this function for currently opened page only; this check box is only available when the auto-completion option is checked.

7.2   Settings - Group Settings

Figure 7-4: Group Settings

This part of settings is only available for users who have the ADM role assigned (i.e. Administrator). Other users configure their own settings under *My Settings*. The *Group Settings* option will not be accessible for them.

The Administrator can configure settings of different groups in here by choosing the specific group from the drop down list.

Please refer to previous sections in 7 about *My Settings* on how to configure project files, process workflow, exception handling, and supervised learning.

7.3   Settings – Workflow

WebCenter Forms Recognition Runtime assigns pre-defined output states to batches after each processing step. Different states are used to distinguish successful steps from failures. If the state indicates a failure of the latest processing step, or if a step is to be carried out manually, the corresponding batch is forwarded to WebCenter Forms Recognition Verifier, where you correct the errors and supply missing results. Therefore, you need to know the output states used by your Runtime Server installation. They determine the input states used by the WebCenter Forms Recognition Verifier stations.

After verification, each batch must be returned to Runtime Server. Again, the output state of the verification step and the input state of the following step must match.

Example

Let's assume that Runtime Server uses the input and output states depicted below to process batches.
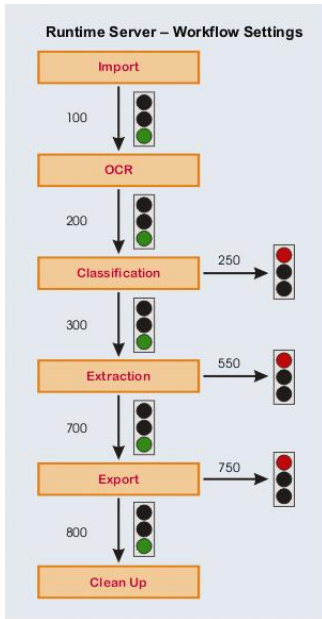
Figure 7-5: Sample Runtime server workflow configuration

In this case, a Verifier station conducting classification, extraction verification or document separation should use the settings shown to verify whether this is correct.
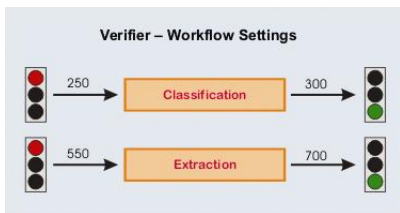
Figure 7-6: A Verifier workflow configuration matching the Runtime server settings above

This produces the following combined workflow:
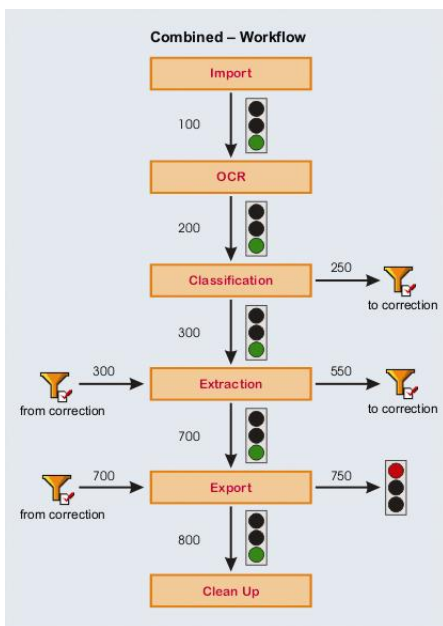
Figure 7-7: Combined workflow resulting from matching Runtime server and Verifier workflow configurations above.

Normally, the workflow should be set up like this:

- Manual indexing / extraction correction after manual classification[2]

- Export after manual indexing / extraction correction

However, there is no need to have everything done at one workstation. Tasks can be distributed among multiple workstations.

The export step is normally defined in custom scripts. Therefore, there is no default mechanism to handle export failures. However, there is the option of implementing a custom routine for this purpose. This should be done by an experienced user.

7.3.1.   Configuring Tasks to Perform at the Workstation

To specify the tasks that are to be carried out at the current WebCenter Forms Recognition Web Verifier station, select the Workflow tab.

Figure 7-8: Workflow tab

- To configure classification verification at this workstation, click *Classification Verification*. You can also configure extraction verification and document separation by clicking this button. One or all of the steps can be performed at the workstation.

- After you have selected the steps to perform, establish values for input and output states. To add an input value, right click on the *Input* option box and select *Add State* on the shortcut menu. (You can also change states and remove states this way.) To set an output value, select it from the dropdown box to the right of the document separation, classification or extraction button.

- *Perform automatic extraction after manual classification*: Forces WebCenter Forms Recognition to attempt to automatically extract data after the Verifier operator manually classifies the document. To select this option, the output state of the Classification

---

[2]   If a document could not be classified, automatic extraction probably will not work.

Verification workflow step must be entered as an input state for the Extraction Verification input step.

- *Disable navigation to valid documents*: When users click "*next document*" button, it will bring the users to the next invalid document and skip the validated ones. This option is checked by default.

- *Hide batch release dialog* enables users to go directly to the next invalid batch without the confirmation release dialog.

- *When entering document verification mode, by default open next available invalid batch:* This option is very advantageous for projects with large amount of batches and simultaneous Verifier users. When this option is checked, the behavior of the application changes when a user completes batch verification. When a complete batch is verified to completion, the user is typically prompted to *Verify Next Invalid Batch.* This setting determines the next invalid batch that will be presented to the user. *Checked* – The next invalid batch will be presented to the user, succeeding the current batch number. *Unchecked* – The first invalid batch at the top of the queue will be presented to the user.

7.4   Settings – Exception Handling

To specify what to do if the verification cannot be finished normally, select the Exception Handling tab.

A document with an unexpected error may not be suitable for Verification. Without a mechanism to handle unexpected failures, operators will not be able to remove the batch with this document from their task list. This is why WebCenter Forms Recognition Verifier applications incorporate an exception handling mechanism. It allows operators to manually assign special states to documents with unexpected errors.

The corresponding documents can be forwarded to verification stations that specialize in collecting exceptions.
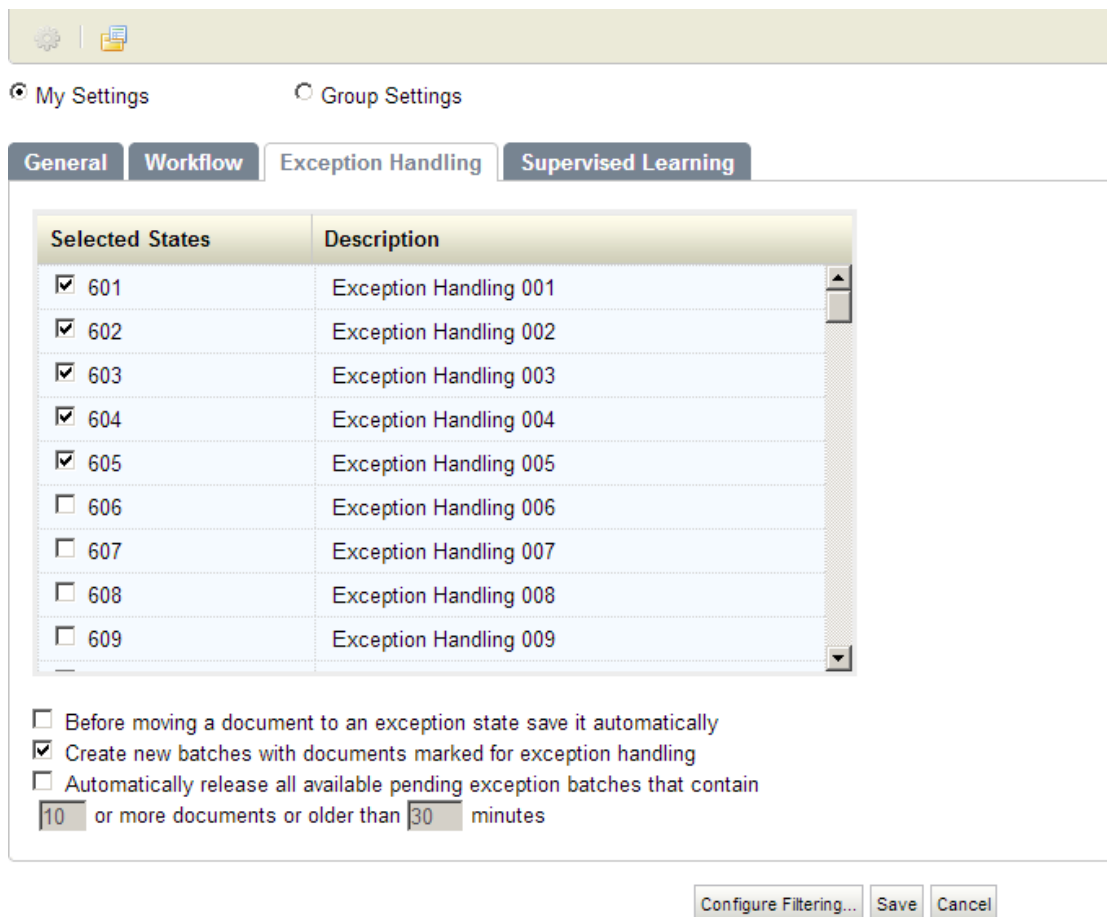
Figure 7-9: Exception Handling tab

7.4.1.   Selecting States

To select a state, mark the corresponding check box. For each selected state, a menu command will be available in the Verification View. The menu commands allow for case-specific handling of various types of unforeseeable errors.

The available exception states cover the range from 601 to 699. Remember that a batch state corresponds to the lowest document state within the batch. Routing batches using their exception state is only possible if the state for successful verification is greater than the one used for exceptions.

The description represents the menu command's label. To set the label, right click on the existing label and select *New description*. Then type the label into the corresponding field and confirm. The 'New Description' field allows 128 characters.

7.4.2.   Settings

- *Before moving a document to an exception state, save it automatically*: Saves a document automatically before moving it to an exception state.

- *Create new batches with documents marked for exception handling*: When this option is selected, the documents that are marked for exception handling will be moved to an *exception batch*. A batch is created for each exception code. The new batch receives a new batch ID. Documents from all verified batches are moved to the same *exception batch* in the Batch View. These batches can be released manually or automatically. (See section 8.1.1 MENU COMMANDS AND KEYBOARD SHORTCUTS -- BATCH VIEW)

- *Create new batches with documents marked for exception handling: When this option is turned off, documents marked for exception handling stay in their batches. These batches keep their batch ID but are renamed according to the state description.*

- *Automatically release all available pending exception batches that contain N or more documents or older than M minutes*: When this option is selected, an exception batch is released once it contains more than N documents or is older than M minutes. This allows critical exception documents to be processed without waiting for manual intervention. Exception batches will also be released when user has exited the application and logged in again.

7.5   Settings – Supervised Learning

It allows users to nominate documents to the knowledge base job. Only users with the SLV role assigned will be able to access this *Settings* tab.

Figure 7-10: Supervised Learning

- *Activate supervised learning workflow:* It allows user to copy documents from current job to knowledge base job.

- *Knowledge base job:* The job name which is used to keep track of the documents added to the knowledge base.

- *Always automatically re-extract data for poorly processed documents:* Global extraction will be applied on documents with low extraction rate.

  Copy document to global knowledge base:

- *Only if activated by the user:* Copying of the document to the global knowledge base has to be triggered by the user manually by clicking in the Document View page.

- *Automatically if unclassified or more than... % of fields are invalid:* Documents will be copied automatically to the global knowledge base if they meet the configured criteria.

*Note: If **Only if activated by the user** is not selected, and the document does not meet the criteria to be copied automatically, you can still copy a document to global knowledge base by clicking in the Document View page.*
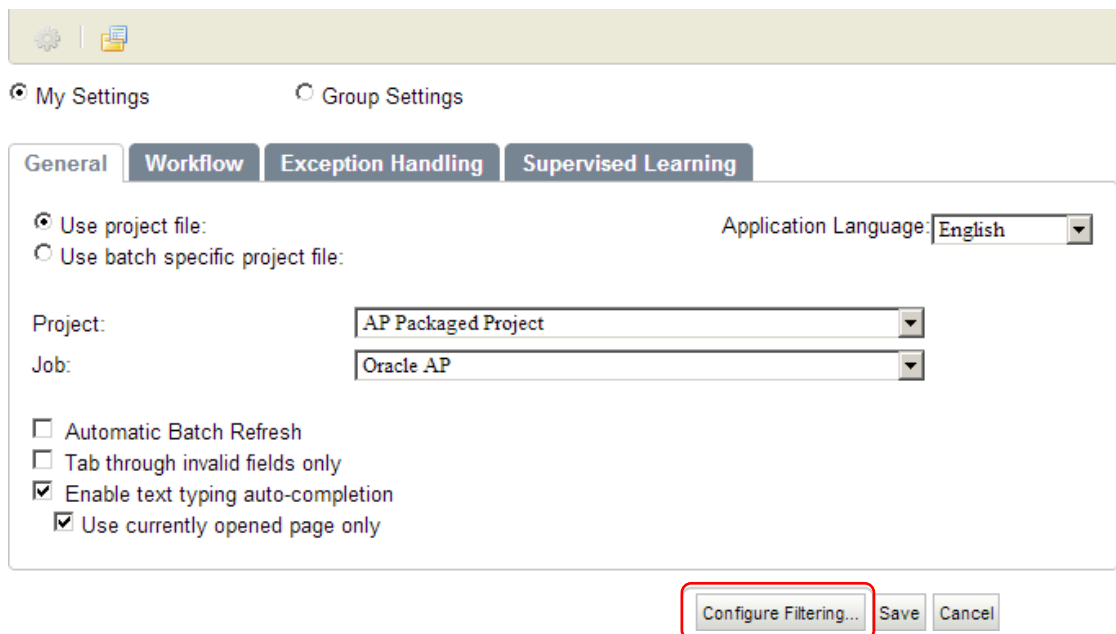
7.6   Batch Filter

The Batch Filter function enables you to specify filter conditions on which batches should be displayed. This is useful if you want to find a subset of batches in a huge job or to limit Verifier user activities.

The filtering dialog window is accessible outside the settings dialog so that users without a SET role but with the FLT role are able to filter batches. Only users having the FLT role assigned will be able to configure filter conditions.

To configure your filter settings:

- Select *Filtering* from the list under Options, or click on the *Batch filter* icon.

- Users with the SET role assigned can access the filtering window also from within the *Settings* dialog window:

Figure 7-11: Batch Filtering button within the Settings dialog window

- This will bring up the Filtering configuration window:
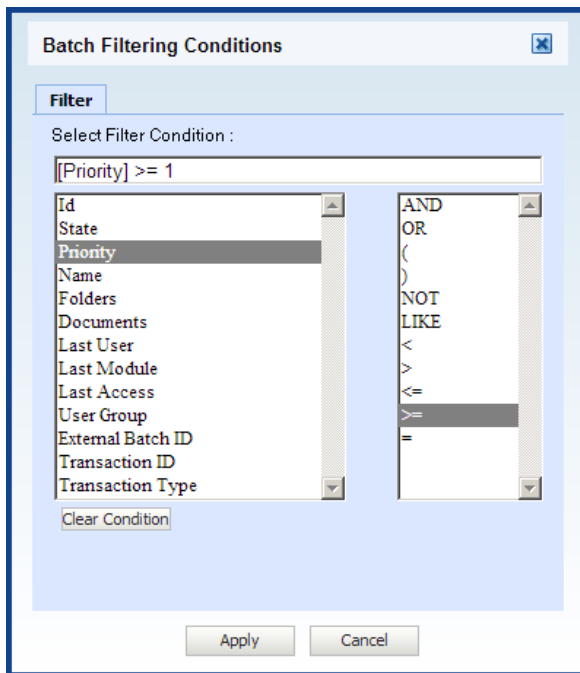
Figure 7-12: Batch Filtering Properties

- Double click on an entry in the left pane to select a batch attribute, and then double click on a filter condition in the right pane. E.g. in the example above, the batches with *Priority* equal or greater than 1 would be displayed.
- On the status bar, the currently applied filter conditions are displayed.
- You can clear the filter condition settings by clicking on *Clear Condition* within the Filtering dialog.

Filter condition	Format
Batch ID	string, numeric: see note below
State	numeric
Priority	numeric
Name	string
Folders	numeric
Documents	numeric
Client	string
Last User	string
Last Module	string
Last Access	date

Note: All string related operations (LIKE, '%', etc) can be used with [Batch ID] value. Numeric queries can also be used in the case batches were not imported from the file system and the IDs were not changed through script.

The filtering conditions are combined with the 'Show batches' condition using logical AND.

Within the *Group Settings* section, Administrators can easily define filter conditions for the several user groups.

1. Select a group from the dropdown list.

2. Click the *Configure Filtering…* button.

3. Establish the filtering conditions and confirm by clicking the *Apply* button.

4. Press *Save* to save your settings.

Users assigned to this group will now be presented with a subset of batches matching the filter conditions.

In Web Verifier, the filtering settings are saved in the database.

8   Getting Familiar with the User Interface

8.1   The Batch View

The first window displayed after starting WebCenter Forms Recognition Web Verifier is called the Batch View because it shows a list of batches. This is your worklist.

Click the following icon for viewing the Batch View:

Note: If WebCenter Forms Recognition Verifier is not yet configured, the list of batches will show empty. (See *7 CONFIGURING WEB VERIFIER* for details)

The Batch View looks like this:
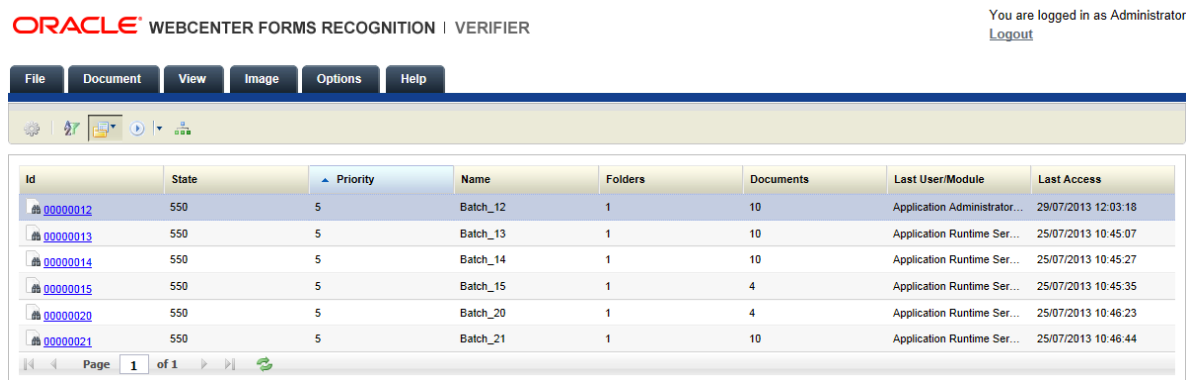
Figure 8-1: Batch View

Menu Bar (See section 8.1.1)

Toolbar (See section 8.1.2)

Table of Batches (See section 8.1.3)

8.1.1.   Menu Commands and Keyboard Shortcuts -- Batch View

Via the menu bar, the following commands can be accessed:

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
File				
	Print Setup…		n/a	The print output can be adjusted here. (See section PRINTING)
	Print…	[Ctrl] + [0]	n/a	Not available in Batch View

Table 8-1: Batch view menu commands (File)

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
Document				
	No commands are available in the Batch View context.			
View				

	Toolbar		n/a	Switch that can be used to show or hide the toolbar.
	Company Logo		n/a	Switch that can be used to show or hide the company logo.
	Show Batches	[Ctrl] + [1]	n/a	Not available in Batch View
	Verify Batch		n/a	Starts the verification of the selected, next or first invalid batch. Depending on the status of the batch, the Verification View is either displayed in Classification Mode (See section 8.3) or in Indexing Mode. (See section 8.4)
			Verify selected/ next batch	Starts the verification of the selected batch. Depending on the status of the batch, the Verification View is either displayed in Classification Mode (See section 8.3) or in Indexing Mode. (See section 8.4)
			Verify first invalid batch	Starts the verification of the first invalid batch on the list.
	Show Selected Batch		n/a	Displays the Document View of the currently selected batch. The Document View provides an overview of the documents within the batch.
	Batch Filter		n/a	Filters the table of batches as specified below.
			All Batches	All batches from the batch root directory.
			Batches to Verify, Classification Only	Only batches waiting for classification verification.
			Batches to Verify, Indexing Only	Only batches waiting for indexing verification.
			Batches to Verify	Only batches waiting for verification.
	Document Filter		n/a	Not available in Batch View

Table 8-2: Batch view menu commands (Document & View)

Menu	Submenu/ Command	Command	Description
Image			
	No commands are available in the Batch View context.		
Options			
	Settings	n/a	Displays a dialog box where you can configure WebCenter Forms Recognition Web Verifier. (See 7 CONFIGURING WEB VERIFIER)
	Filtering	n/a	Brings up a dialog window to configure Filter conditions for batch filtering. (See section BATCH FILTER)

	Reclassify Document Manually	n/a	Not available in Batch View
	Move Document to Exception State	n/a	Not available in Batch View
	Release Exception Batches		Manually releases all batches that were created when documents are marked as exceptions. This menu item is only activated, if any documents have already been moved to exception state. See section 7.4 SETTINGS – EXCEPTION HANDLING
	Apply global extraction		Not available in Batch View
	Copy to global knowledge base		Not available in Batch View
	Correct Table		Not available in Batch View
	Switch Table Highlighting		Not available in Batch View
	Change Password		Enables you to change your password.

Table 8-3: Batch view menu commands (Image & Options)

8.1.2.   Toolbar Buttons

The toolbar provides quick access to some frequently used commands:

Button	Description
	Displays a property sheet where you can configure WebCenter Forms Recognition Web Verifier. (See 7 CONFIGURING WEB VERIFIER)
	Displays a dialog box where you can configure the batch filtering conditions properties.
	If you click on the arrow to the right of this button, the available filters for the list of batches are displayed. You can select either: All batches Batches to Verify, Classification Only Batches to Verify, Indexing Only Batches to Verify
	Starts the verification of the currently selected batch. Depending on the batch state, the batch is either displayed in the classification window (Section 8.3 THE VERIFICATION VIEW – CLASSIFICATION WINDOW) or in the indexing window. (Section 8.4 THE VERIFICATION VIEW - INDEXING WINDOW) A dropdown list allows users to open selected batch or open the next invalid batch.
	Displays the batch structure of the currently selected batch. Selecting a document shows the Document View which provides an overview of the documents inside the batch. (Section 8.2 THE DOCUMENT VIEW)

Table 8-4: Batch view controls

8.1.3. Table of Batches

In the table of batches, a batch is represented by a single row. In front of each batch, a small symbol is displayed having the following meaning:

Symbol	Description
	Batch is finished and ready for export.
	Batch requires a correction of the classification results.
	Batch requires a correction of the extraction results.
	Batch is locked and displayed dimmed as it is in use by another application. Therefore it cannot be opened for correction.
	Batch contains documents with exception statuses. When it is dimmed, it needs to be released before you can work on it again.

Table 8-5: Table of batches – symbols

When no icon is shown, the batch state is out of workflow. Select another batch or change the settings for the workflow. (See section 7.3 SETTINGS – WORKFLOW)

The batch list can be sorted by each column. The table columns display the following information about the batch:

- **Batch ID:**
  A number that can be used to uniquely identify the batch. This is similar to a Social Security number, which uniquely identifies a person.

- **State:**
  An integer between 0 and 999 that indicates the progress of batch processing. The state also indicates whether the batch is ready for verification.

- **Priority:**
  An integer between 1 and 9 that indicates how urgent it is that a job be finished. 1 is the highest priority (very urgent) 9 the lowest.

- **Name:**
  An arbitrary name that is easier to read than the batch ID. Because the name is optional, it might be missing.

- **Folders:**
  Documents in a batch can be grouped in structures called folders. The value in this column indicates the number of folders inside the batch.

- **Document:**
  The value in this column indicates the number of documents inside the batch.

- **Last user/Module:**

  The computer name of the operator who has processed the batch before, and the name of the application that most recently processed the batch.

- **Last Access:**

    Displays the date when the batch was last processed.

- **External Group ID:**

    The Group ID which has been assigned to a batch is relating to security. Batches can be assigned to user group via a unique ID.

- **External Batch ID:**

    The name of the Batch Group. It can be used to represent any piece of information you would like to associate with batch. For example, external system ID, storage box ID, etc.

- **Transaction ID:**

    The Transaction ID assigned to a batch. It allows the developer to synchronize a newly created batch of documents with another external system. It can be used to identify originators of batch of documents.

- **Transaction Type:**

    The Transaction Type assigned to a batch. It allows the developer to synchronize a newly created batch of documents with another external system. It can be used to identify the types of documents (Invoices, Claim forms etc.) in batches or source of document (Email, Scanned etc.)

Note: The four table columns External group ID, External Batch ID, Transaction ID and Transaction Type are not displayed by default. See Installation Guide on how to activate these columns.

8.1.4.   Sorting and Navigating in the Batch View

You can sort any column in the Batch View. To sort any item, click on the title of that column. Batches will sort according to their position on the list. If you select the first batch, and then click the Batch column label, it will move to the bottom of the list.

For other items, the numbers will toggle between ascending and descending order, whether numerical or alphabetical.

In the table of batches you can select a batch by clicking on it. Depending on the status of the batch, either the classification or the verification view will open.

8.2   The Document View

The Document View can be used to investigate the documents in a selected batch.

To select this view, click *Show selected batch*:

The Document View looks like this:
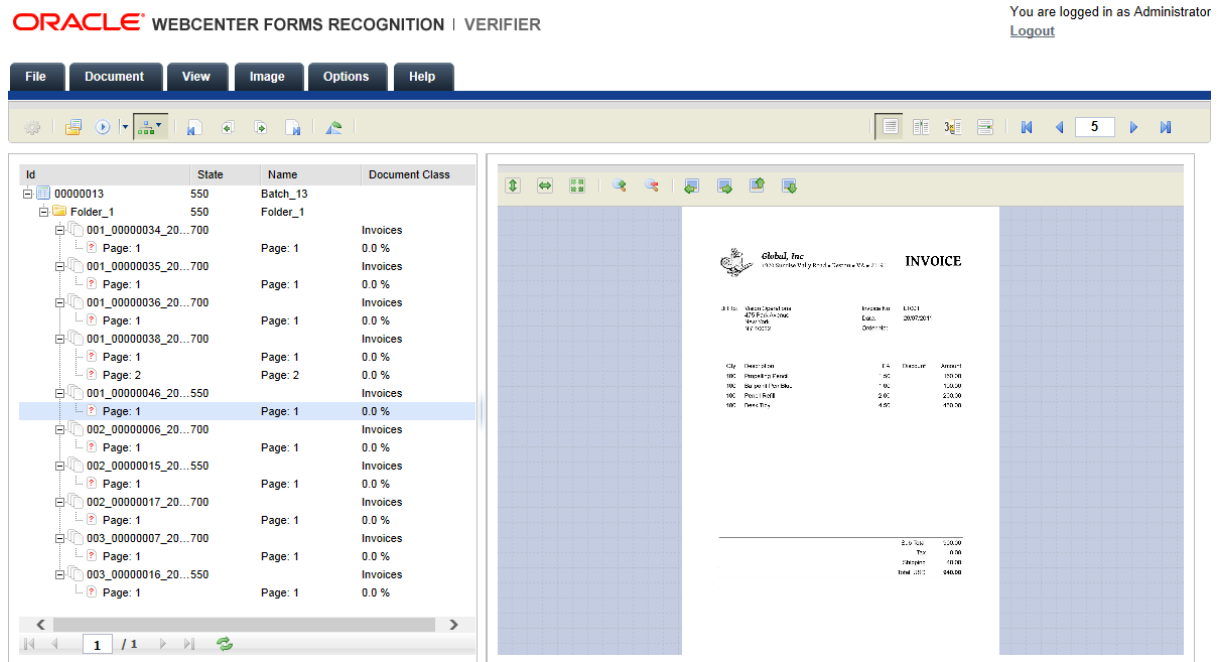
Figure 8-2: The Document View

Menu Bar (See section 8.2.1)

Main Toolbar (See section 8.2.2)

Batch Structure (See section 8.2.3)

Viewer Toolbar (See section 8.2.4)

Document (See section 8.2.5)

8.2.1.   Menu Commands and Keyboard Shortcuts – Document view

The following commands can be accessed via the menu bar:

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
File				
	Print Setup…		n/a	The print output can be adjusted here. (See section PRINTING)
	Print…	[Ctrl] + [0]	n/a	Starts printing dialog.
Document				
	First Document	[Ctrl] + [Alt] + [Home]	n/a	Moves to the first document.
	Previous Document	[Ctrl] + [Alt] + [Page Down]	n/a	Moves to the previous document.
	Next Document	[Ctrl] + [Alt] + [Page Up]	n/a	Moves to the next document.
	Last Document	[Ctrl] + [Alt] + [End]	n/a	Moves to the last document.

	Append Document	[Ctrl] + [8]	n/a	Appends the current document to the previous one. As a result you will get a single document including pages of the previous and the current document.
	Cut Document	[Ctrl] + [9]	n/a	If applicable, splits the current document to two separate documents.
	Accept/Reject next unsure page	[Ctrl] + [Enter]	n/a	Enables change to current page status.
	Select next unsure page	[Ctrl] + [Space]	n/a	Focus moves to next page with "unsure separation".

Table 8-6: Document view menu commands (File & Document)

<u>Note:</u> The availability of the Document menu options depends on whether the document itself or one of its pages is selected, and whether the first or a subsequent page is selected.

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
View				
	Toolbar		n/a	Switch that can be used to show or hide the toolbar.
	Company Logo		n/a	Switch that can be used to show or hide the company logo.
	Show Batches	[Ctrl] + [1]	n/a	Displays the Batch View. (See section 8.1 THE BATCH VIEW)
	Verify Batch		n/a	Starts the verification of the selected batch. Depending on the status of the batch, the Verification View is either displayed in Classification Mode (See section 8.3 THE VERIFICATION VIEW – CLASSIFICATION WINDOW) or in Indexing Mode. (See section 8.4 THE VERIFICATION VIEW - INDEXING WINDOW)
			Verify selected/ next batch	Starts the verification of the selected batch. Depending on the status of the batch, the Verification View is either displayed in Classification Mode (See section 8.3) or in Indexing Mode. (See section 8.4)
			Verify first invalid batch	To verify the first invalid batch.
	Show Selected Batch		n/a	Not available in Document View
	Batch Filter		n/a	Not available in Document View
	Document Filter		n/a	Filters the structure of the selected document. When a filter is applied, you can only see the corresponding subset of documents. Folders that contain such documents will be open. The

				remaining folders will be closed, and you cannot open them.
			All documents	No filter is applied. All of the documents are listed.
			Documents to classify or index	Shows documents that have to bbe classified or indexed.
			Documents to classify	Shows documents that require classification.
			Documents to index	Shows documents that require indexing.
	Multi-Page View			Display options for multi-page documents.
			Single Page	Only one page is displayed. Use the navigation buttons to browse through further pages.
			Two Pages Horizontally	Two pages of the document are placed side by side.
			Three Pages Horizontally	Three pages of the document are placed side by side.
			Two Pages Vertically	Two pages of the document are placed one below the other.

Table 8-7: Document view menu commands (View)

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
Image				
	Zoom In	[Shift] + [Add]	n/a	Zooms in.
	Zoom Out	[Shift] + [Subtract]	n/a	Zooms out.
	Move Image to Left	[Ctrl] + [Left]	n/a	Scrolls the document area to the left.
	Move Image to Right	[Ctrl] + [Right]	n/a	Scrolls the document area to the right.
	Move Image Upwards	[Ctrl] + [Up]	n/a	Scrolls the document area upward.
	Move Image Downwards	[Ctrl] + [Down]	n/a	Scrolls the document area downward.
	Rotate	[Ctrl] + [R]	n/a	Rotates image 90 degrees to the right.
	First Page in Document	[Ctrl] + [Home]	n/a	Moves to the first page of the document.
	Previous Page	[Ctrl] + [Page Down]	n/a	Moves to the preceding page of the document.

Menu	Submenu/Command	Command	Description
	in Document			
	Next Page in Document	[Ctrl] + [Page Up]	n/a	Moves to the following page of the document.
	Last Page in Document	[Ctrl] + [End]	n/a	Moves to the final page of the document.
	Fit to Height	[Ctrl] + [H]	n/a	Fits the document to the height of the viewer.
	Fit to Width	[Ctrl] + [W]	n/a	Fits the document to the width of the viewer.
	Best Fit	[Ctrl] + [F]	n/a	Forces the document to fit into the viewer so all of it is displayed.
	Keep Focus on Field		n/a	Not available in Document View
	Keep Zoom		n/a	Not available in Document View
	Increase image area	[Ctrl] + [J]	n/a	Not available in Document View
	Decrease image area	[Ctrl] + [K]	Two Pages Vertically	Not available in Document View

Table 8-8: Document view menu commands (Image)

Menu	Submenu/Command	Command	Description
Options			
	Settings	n/a	Not available in Document View
	Filtering	n/a	Not available in Document View
	Reclassify Document Manually	n/a	Not available in Document View
	Move Document to Exception State	n/a	Not available in Document View
	Release Exception Batches	n/a	Releases all exception batches. This menu item is only activated, if any documents have already been moved to exception state.
	Apply global extraction	n/a	Not available in Document View
	Copy to global knowledge base	n/a	Not available in Document View
	Correct Table	n/a	Not available in Document View
	Switch Table Highlighting	n/a	Not available in Document View

	Change Password	n/a	Allows you to change your password.

Table 8-9: Document view menu commands (Options)

8.2.2.   Main Toolbar Buttons

The toolbar provides quick access to some frequently used commands:

Button	Description
⚙	Not available in Document View
🗔	Displays the Batch View.
▶ ▾	Starts the verification of the currently selected batch. Depending on the batch state, the batch is either displayed in the classification window (Section 8.3 THE VERIFICATION VIEW – CLASSIFICATION WINDOW) or in the indexing window. (Section 8.4 THE VERIFICATION VIEW - INDEXING WINDOW) A dropdown list allows users to open selected batch or open the next invalid batch.
⛁ ▾	Displays the available filters for the batch structure. You can select from among the following options: • All documents • Documents to Classify or Index • Documents to Classify • Documents to Index
◁|	Displays the first page in document. This button is only enabled if the current document has more than one page.
◁	Displays the previous page in document.
▷	Displays the next page in document.
|▷	Displays the last page in document. This button is only enabled if the current document has more than one page.
↻	Rotates current page clockwise.
🗎	Shows single page.
🗎🗎	Shows two pages horizontally.
3×🗎	Shows three pages horizontally.
🗎	Shows two pages vertically.
◁|	Displays the first document in the batch.

◀	Displays the previous document in the batch.
▷	Displays the next document in the batch.
▶|	Displays the last document in the batch.

Table 8-10: Document view toolbar

8.2.3.   Batch Structure Area

In the batch structure, a hierarchical representation of the batch contents is displayed.
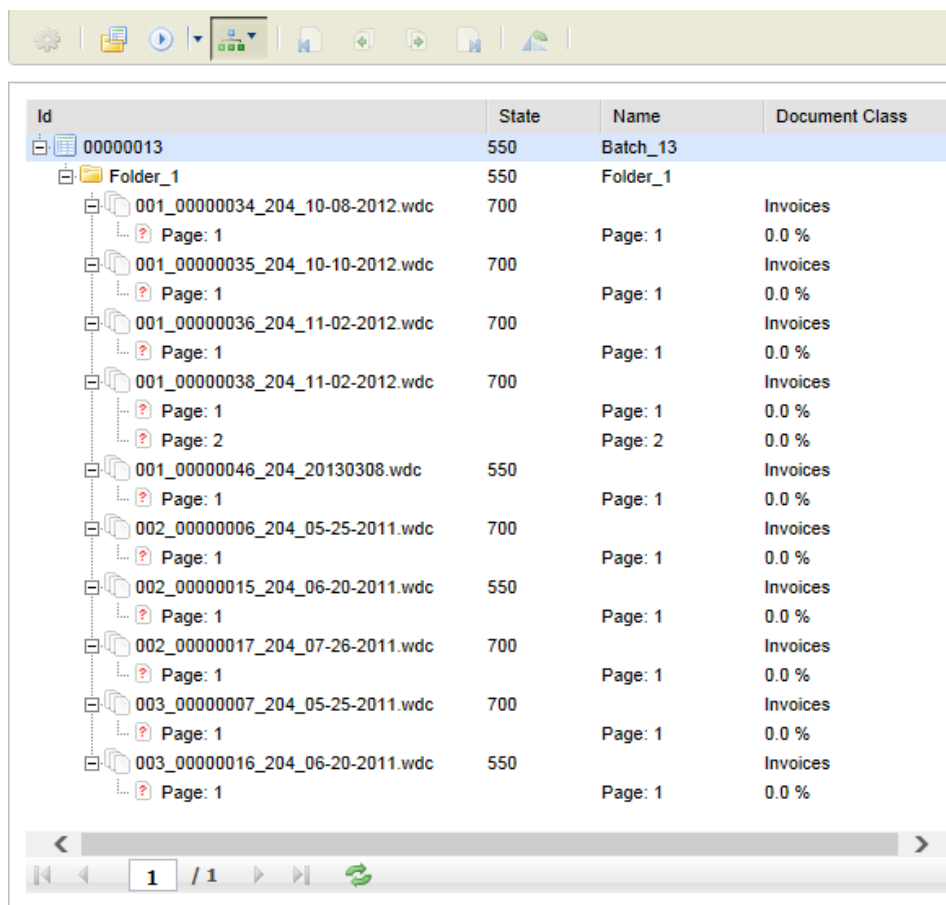
Figure 8-3: Batch structure

The levels of this hierarchy are:

- **Batch**

- **Folder**

- **Document**

   For each entry, the following information is provided:

- **ID:**
   A number that can be used to uniquely identify the batch, folder, or document. This is similar to the number on a Social Security card that can be used to uniquely identify the owner of the card.

- **State:**
  An integer value between 0 and 999 that indicates the progress of batch processing. The batch state is calculated from the states of its folders. It corresponds to the lowest value of all folder states. The folder state is in turn calculated from the states of the documents. It corresponds to the lowest value of all document states.

- **Name:**
  An arbitrary batch or folder name that is easier to read than the ID. Because the name is optional, it might be missing.

- **Document Class:**
  A document's classification result. This entry might be missing if the document has not been classified.

Note: Depending on your access rights, the document view might list documents which are inaccessible. This is due to their input states which are out of the workflow you have access to.

8.2.3.1. Sorting and Navigating in the Document View

You can sort any column in the Document View. To sort any item, click on the title of that column. Batches will sort according to their position on the list. If you select the first batch, and then click the Batch column label, it will move to the bottom of the list.

For other items, the numbers will toggle between ascending and descending order, whether it is numerical or alphabetical.

In the batch structure, you can select the documents by clicking on them.

To expand or collapse a folder, double click on it, or click the + or - sign next to it.

8.2.3.2. Splitting and Appending Documents

In the document list, you can split multipage documents into separate documents, with the exception of the first page of a document which cannot be split. You can also merge consecutive pages of documents into one with multiple pages. (See FIGURE 8- and FIGURE 8-)

Splitting Multipage Documents

To split a multipage document, do the following:

1) Select *View*, *Show Selected Batch*, or click *Show Selected Batch*.

2) In the document list, click on the desired multipage document. This document must have at least two pages.

3) Right click on the page.

4) Select *Cut pages into a new document*. The document is now split into two documents. The second document will have the same name as the first, but will also have an underscore and 1 after the number or name. For example, a document initially called "invoiceabc" would now be "invoiceabc" and "invoiceabc_1". Corresponding TIFFs will also be created.
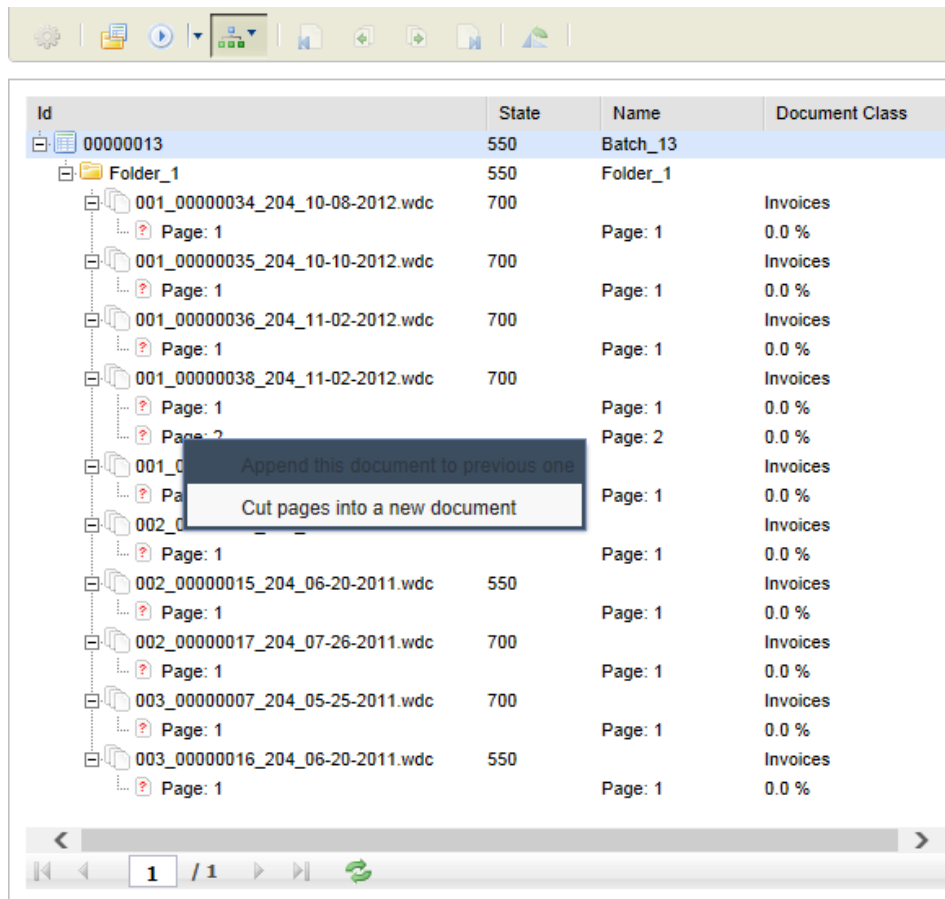
Figure 8-4: Splitting pages in a new document

Appending a Document

To append a document to another:

1) Select *View*, *Show Selected Batch*, *All Documents from the Main Menu*, or click *Show Selected Batch*, *All Documents*.

2) Select the document to append to the previous document.

3) Right click on the document.

4) Select *Append this document to previous one*. The document will now appear in the list as a multi-page document. The document name of the previous document is kept for the new multipage document.

Splitting the appended document back into a new document will restore the old document name of the formerly appended document. This also works if more than two documents have been merged.
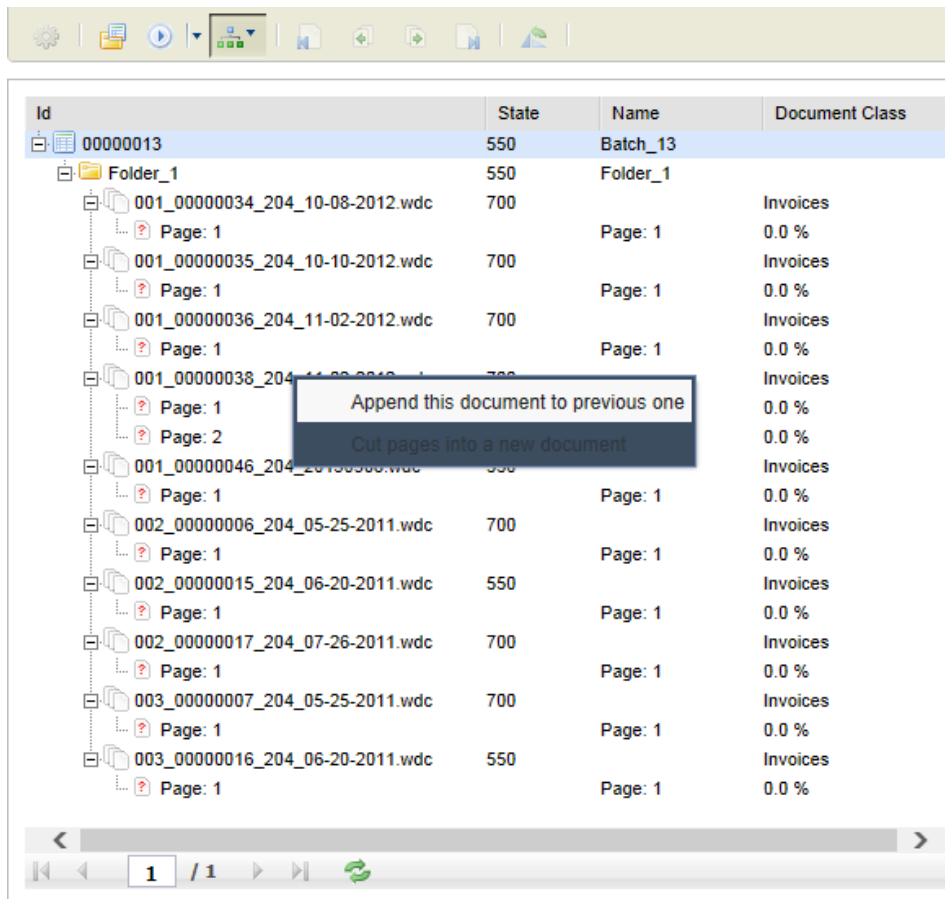
Figure 8-5: Appending pages to a document

The original page properties of merged/splitted pages are saved in the OriginalDocumentFileName DocPage property.

8.2.4.   Viewer Toolbar Buttons

The viewer toolbar allows you to adjust the magnification used to display documents via the following commands:

Button	Description
	Fits the document to window height.
	Fits the document to window width.
	Best fit.
	Zooms in.
	Zooms out.
	Moves image to left.

	Moves image to right.
	Moves image upwards.
	Moves image downwards.

Table 8-11: Viewer Toolbar Controls

8.2.5. Document Area

This area shows the first page of the document that has been selected in the batch structure.

8.3   The Verification View – Classification Window

To display the Verification View, select a batch from the list that requires verification. Clicking on it will open the first document in verification view. If you want to verify a batch which is

already highlighted as selected, click on the toolbar button *Verify selected batch* ⊙ or select the appropriate option from the *View* menu.

When you open the Verification View, the classification window is displayed automatically if the next document that is to be verified requires a correction of the classification result. Whether this is the case depends on the document's state.

The Classification window looks like this:
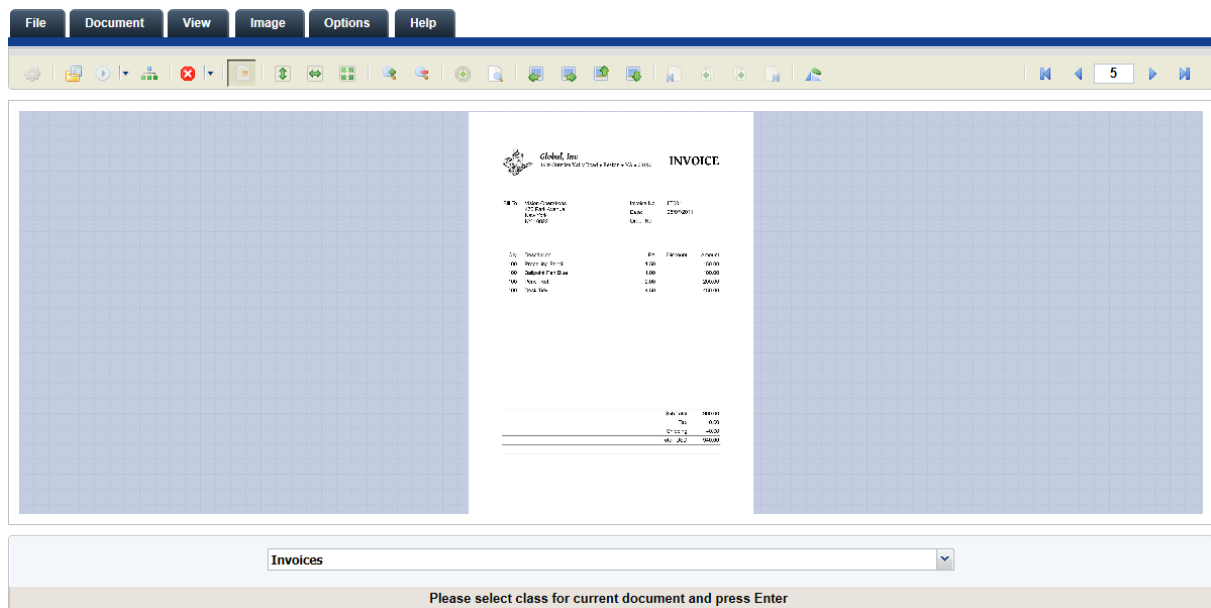
Figure 8-6: The Classification window

Menu Bar (See section 8.3.1 MENU COMMANDS AND KEYBOARD SHORTCUTS)

Toolbar (See section 8.3.2 TOOLBAR BUTTONS)

Document (See section 8.3.3 DOCUMENT AREA)

Class Selection (See section 8.3.4 CLASS SELECTION LIST)

8.3.1.   Menu Commands and Keyboard Shortcuts – Classification Window

Via the menu bar, the following commands can be accessed:

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
File				
	Print Setup…		n/a	Adjusts print output
	Print…	[Ctrl] + [0]	n/a	Starts printing dialog.
Document				
	First Document	[Ctrl] + [Alt] + [Home]	n/a	Moves to the first document.
	Previous Document	[Ctrl] + [Alt] + [Page Down]	n/a	Moves to the previous document.
	Next Document	[Ctrl] + [Alt] + [Page Up]	n/a	Moves to the next document.
	Last Document	[Ctrl] + [Alt] + [End]	n/a	Moves to the last document.
	Append Document	[Ctrl] + [8]	n/a	Not available in Classification Mode.
	Cut Document	[Ctrl] + [9]	n/a	Not available in Classification Mode.
	Accept/Reject next unsure page	[Ctrl] + [Enter]	n/a	Not available in Classification Mode.
	Select next unsure page	[Ctrl] + [Space]	n/a	Not available in Classification Mode.

Table 8-12: Verification View - Classification Mode menu commands (File & Document)

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
View				
	Toolbar		n/a	Switch that can be used to show or hide the toolbar.
	Company Logo		n/a	Switch that can be used to show or hide the company logo.
	Show Batches	[Ctrl] + [1]	n/a	Displays the Batch View. (See section 8.1 THE BATCH VIEW)
	Verify Batch		n/a	Not available in Classification Mode.
	Show Selected Batch		n/a	Not available in Classification Mode.
	Batch Filter		n/a	Not available in Classification Mode.

	Document Filter		n/a	Not available in Classification Mode.

Table 8-13: Verification View - Classification Mode menu commands (View)

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
Image				
	Zoom In	[Shift] + [Add]	n/a	Zooms in.
	Zoom Out	[Shift] + [Subtract]	n/a	Zooms out.
	Move Image to Left	[Ctrl] + [Left]	n/a	Scrolls the document area to the left.
	Move Image to Right	[Ctrl] + [Right]	n/a	Scrolls the document area to the right.
	Move Image Upwards	[Ctrl] + [Up]	n/a	Scrolls the document area upward.
	Move Image Downwards	[Ctrl] + [Down]	n/a	Scrolls the document area downward.
	Rotate	[Ctrl] + [R]	n/a	Rotates image 90 degrees to the right.
	First Page in Document	[Ctrl] + [Home]	n/a	Moves to the first document page of multipage document. This option is only enabled if the current document contains more than one page.
	Previous Page in Document	[Ctrl] + [Page Down]	n/a	Moves to the previous document page of multipage document. This option is only enabled if the current document contains more than one page.
	Next Page in Document	[Ctrl] + [Page Up]	n/a	Moves to the next document page of multipage document. This option is only enabled if the current document contains more than one page.
	Last Page in Document	[Ctrl] + [End]	n/a	Moves to the last document page of multipage document. This option is only enabled if the current document contains more than one page.
	Fit to Height	[Ctrl] + [H]	n/a	Fits the document to the height of the viewer.
	Fit to Width	[Ctrl] + [W]	n/a	Fits the document to the width of the viewer.
	Best Fit	[Ctrl] + [F]	n/a	Forces the document to fit into the viewer so all of it is displayed.
	Keep Focus on Field		n/a	Keeps the focus on the same field for each document you view in the batch. Not available in classification view.
	Keep Zoom		n/a	Keeps the established zoom settings on each document you view in the batch.

	Increase image area	[Ctrl] + [J]	n/a	Not available in Classification Mode.
	Decrease image area	[Ctrl] + [K]	Two Pages Vertically	Not available in Classification Mode.

Table 8-14: Classification view menu commands (Image)

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
Options				
	Settings		n/a	Not available in Classification Mode
	Filtering		n/a	Not available in Classification Mode
	Reclassify Document Manually	[F7]	n/a	Not available in Classification Mode
	Move Document to Exception State	[F9]	n/a	Moves current document to default exception state. Accessibility: Classification view, Indexing view. *Note: The appropriate toolbar button provides a selection of different exception handling states to choose from. The selection made on the list applies for both, the toolbar button as well as the menu option.*
	Release Exception Batches	[Ctrl] + [E]	n/a	Releases all exception batches. This menu item is only activated, if any documents have already been moved to exception state.
	Apply global extraction		n/a	Not available in Classification Mode
	Copy to global knowledge base		n/a	Not available in Classification Mode
	Correct Table	[Ctrl] + [T]	n/a	Not available in Classification Mode
	Switch Table Highlighting	[Ctrl] + [Q]	n/a	Not available in Classification Mode
	Change Password		n/a	Allows you to change your password.

Table 8-15: Verification View - Classification Mode menu commands (Options)

8.3.2. Toolbar Buttons

The toolbar provides quick access to some frequently used commands:

Button	Description
⚙	Not available in classification view.

	Displays the Batch View.
	Not available in classification view.
	If you click this button, the Document View is displayed. (See section 8.2)
	Clicking the arrow next to this button displays a list of exceptions. You can use these exceptions if you cannot correct a document at all — for example because it belongs to none of the defined classes. Please check with your supervisor to determine which exceptions to use. *Note:* In order to avoid selection conflicts, only the toolbar button provides a list of exception handling states to choose from. The selection made here will also apply if you move a document to exception state by selecting the appropriate option within the Options menu.
	Highlight Candidates – Not available in classification view.
	Fits the current image to the height of the window.
	Fits the current image to the width of the window.
	Fits the current image to the width or height of the window so that maximum enlargement is obtained.
	Zooms in.
	Zooms out.
	Not available in classification view
	Keeps the established zoom settings on each document you view in the batch.
	Moves image to left.
	Moves image to right.
	Moves image upwards.
	Moves image downwards.
	Displays the first page in document. This button is only enabled if the current document has more than one page. If you click on it, the first page is displayed.
	Displays the previous page in document. This button is only enabled if the current document has more than one page. If you click on it, the previous page is displayed.

	Displays the next page in document. This button is only enabled if the current document has more than one page. If you click on it, the next page is displayed.
	Displays the last page in document. This button is only enabled if the current document has more than one page. If you click on it, the last page is displayed.
	Rotates current page clockwise.
	Displays the first document in the batch.
	Displays the previous document in the batch.
	Displays the next document in the batch.
	Displays the last document in the batch.

Table 8-16: Verification View - Classification Mode Controls

8.3.3. Document Area

This area shows the current document. All words that are recognized during OCR are highlighted by default.

8.3.4. Class Selection List

When filled, this box shows the classification result of the current document, or it will be empty if no result determined. If you open the list, you will see all available classes.

To set or change a classification result, make sure that you are not in Browsing Mode. Then either:

- Click on the arrow on the right side of the list box to open the list, and then select a class.
- Use the arrow keys to browse through the list of classes. The entries in the list are sorted alphabetically.
- If you know the correct class name, you may type its first characters and wait until the system automatically displays the full class name.

8.4   The Verification View - Indexing Window

To display the Verification View, select a batch from the list that requires verification. Clicking on a batch will immediately start the verification. If you want to verify a batch which is already highlighted as selected, click on the toolbar button *Verify Batch* . Make sure to select *Verify selected/next batch* on the associated dropdown menu.

The indexing window is displayed automatically if the next document that is to be processed requires a correction of the extraction result. Whether this is the case depends on the document's state.

Normally, the Indexing Mode looks like this:
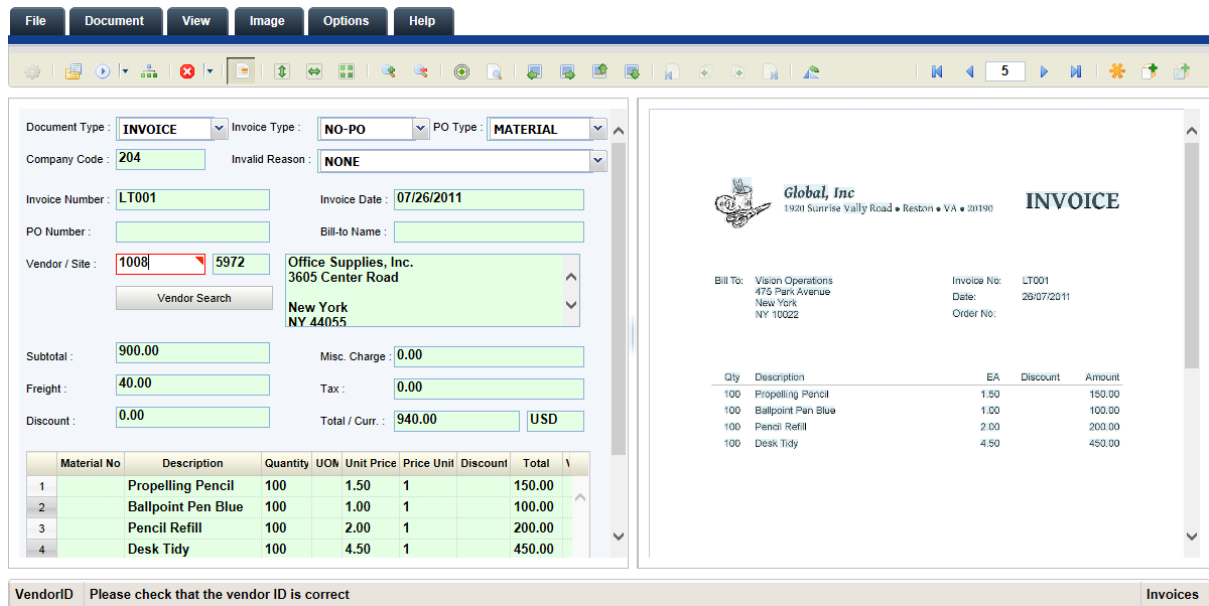
Figure 8-7: Indexing Mode

Your indexing window may look quite different than the one shown in FIGURE 8-7: INDEXING MODE. This is due to two main reasons:

- The fields and documents that are displayed are specific for your organization.

- The layout of the window can be customized by an application designer. Therefore, you might not be able to see all the window elements shown above. In addition, they may be arranged differently. For example, the document display could also be on the left side. Finally, there may be different window layouts for different document classes.

Regardless of these differences, the basic window elements always work the same way.

8.4.1. Increasing/Decreasing Image Area

You can easily increase and decrease the two parts of the indexing window by dragging the vertical split bar between the image area and field area either to the right or left.

Furthermore, two menu items have been added under the *Image* menu, which allow the user to manually increase or decrease the image area.

This option is only available in the extraction verification view.

8.4.2. Menu Commands and Keyboard Shortcuts

Via the menu bar, the following commands can be accessed:

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
File				
	Print Setup…		n/a	Adjusts print output

	Print…	[Ctrl] + [0]	n/a	Starts printing dialog.
Document				
	First Document	[Ctrl] + [Alt] + [Home]	n/a	Moves to the first document.
	Previous Document	[Ctrl] + [Alt] + [Page Down]	n/a	Moves to the previous document.
	Next Document	[Ctrl] + [Alt] + [Page Up]	n/a	Moves to the next document.
	Last Document	[Ctrl] + [Alt] + [End]	n/a	Moves to the last document.
	Append Document	[Ctrl] + [8]	n/a	Not available in Indexing Mode.
	Cut Document	[Ctrl] + [9]	n/a	Not available in Indexing Mode.
	Accept/Reject next unsure page	[Ctrl] + [Enter]	n/a	Not available in Indexing Mode.
	Select next unsure page	[Ctrl] + [Space]	n/a	Not available in Indexing Mode.

Table 8-17: Verification View - Indexing Mode menu commands (File & Document)

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
View				
	Toolbar		n/a	Switch that can be used to show or hide the toolbar.
	Company Logo		n/a	Switch that can be used to show or hide the company logo.
	Show Batches	[Ctrl] + [1]	n/a	Displays the Batch View. (See section 8.1 THE BATCH VIEW)
	Verify Batch		n/a	Not available in Indexing Mode.
	Show Selected Batch		n/a	Displays the structure of the selected batch in Document View.
	Batch Filter		n/a	Not available in Indexing Mode.
	Document Filter		n/a	Not available in Indexing Mode.

Table 8-18: Verification View - Indexing Mode menu commands (View)

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
Image				
	Zoom In	[Shift] + [Add]	n/a	Zooms in.

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
	Zoom Out	[Shift] + [Subtract]	n/a	Zooms out.
	Move Image to Left	[Ctrl] + [Left]	n/a	Scrolls the document area to the left.
	Move Image to Right	[Ctrl] + [Right]	n/a	Scrolls the document area to the right.
	Move Image Upwards	[Ctrl] + [Up]	n/a	Scrolls the document area upward.
	Move Image Downwards	[Ctrl] + [Down]	n/a	Scrolls the document area downward.
	Rotate	[Ctrl] + [R]	n/a	Rotates image 90 degrees to the right.
	First Page in Document	[Ctrl] + [Home]	n/a	Moves to the first document page of multipage document. This option is only enabled if the current document contains more than one page.
	Previous Page in Document	[Ctrl] + [Page Down]	n/a	Moves to the previous document page of multipage document. This option is only enabled if the current document contains more than one page.
	Next Page in Document	[Ctrl] + [Page Up]	n/a	Moves to the next document page of multipage document. This option is only enabled if the current document contains more than one page.
	Last Page in Document	[Ctrl] + [End]	n/a	Moves to the last document page of multipage document. This option is only enabled if the current document contains more than one page.
	Fit to Height	[Ctrl] + [H]	n/a	Fits the document to the height of the viewer.
	Fit to Width	[Ctrl] + [W]	n/a	Fits the document to the width of the viewer.
	Best Fit	[Ctrl] + [F]	n/a	Forces the document to fit into the viewer so all of it is displayed.
	Keep Focus on Field		n/a	Keeps the focus on the same field for each document you view in the batch. Not available in classification view.
	Keep Zoom		n/a	Keeps the established zoom settings on each document you view in the batch.
	Increase image area	[Ctrl] + [J]	n/a	Increases the image area. (See section 8.4.1)
	Decrease image area	[Ctrl] + [K]	Two Pages Vertically	Decreases the image area. (See section 8.4.1)

Table 8-19: Verification View - Indexing Mode menu commands (Image)

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
Options				

Menu	Submenu/ Command	Keyboard Shortcut	Command	Description
	Settings		n/a	Not available in Indexing Mode w
	Filtering		n/a	Not available in Indexing Mode
	Reclassify Document Manually	[F7]	n/a	Opens the classification window (See section 8.3 THE VERIFICATION VIEW – CLASSIFICATION WINDOW) for the current document.
	Move Document to Exception State	[F9]	n/a	Moves current document to default exception state. Accessibility: Classification view, Indexing view. *Note:* The appropriate toolbar button provides a selection of different exception handling states to choose from. The selection made on the list applies for both, the toolbar button as well as the menu option.
	Release Exception Batches	[Ctrl] + [E]	n/a	Releases all exception batches. This menu item is only activated, if any documents have already been moved to exception state.
	Apply global extraction		n/a	Applies global extraction on the document. This menu option is only available if the Supervised Learning is configured appropriately (See section 7.5 SETTINGS – SUPERVISED LEARNING).
	Copy to global knowledge base		n/a	On completion of verification of the current document, copies the document to global knowledge base.
	Correct Table	[Ctrl] + [T]	n/a	This option is available if the cursor focus is on a table cell. Clicking on it starts table correction.
	Switch Table Highlighting	[Ctrl] + [Q]	n/a	This will change the line highlighting to show mapped lines and mapped columns.
	Change Password		n/a	Allows you to change your password.

Table 8-20: Verification View - Indexing Mode menu commands (Options)

8.4.3.   Toolbar Buttons

The toolbar provides quick access to some frequently used commands:

Button	Description
	Not available in Indexing Mode.
	Displays the Batch View.
	Not available in Indexing Mode.

	If you click this button, the Document View is displayed. (See section 8.2)
	Clicking the arrow next to this button displays a list of exceptions. You can use these exceptions if you cannot correct a document at all — for example because it belongs to none of the defined classes. Please check with your supervisor to determine which exceptions to use.
	Note: In order to avoid selection conflicts, only the toolbar button provides a list of exception handling states to choose from. The selection made here will also apply if you move a document to exception state by selecting the appropriate option within the Options menu.
	Highlight Candidates – Clicking on it highlights all areas in yellow that were taken into account to fill the current field.
	Fits the current image to the height of the window.
	Fits the current image to the width of the window.
	Fits the current image to the width or height of the window so that maximum enlargement is obtained.
	Zooms in.
	Zooms out.
	If this button appears pressed down, the application always displays the document area that is associated with the currently selected field.
	Keeps the established zoom settings on each document you view in the batch.
	Moves image to left.
	Moves image to right.
	Moves image upwards.
	Moves image downwards.
	Displays the first page in document.
	Displays the previous page in document.
	Displays the next page in document.

	Displays the last page in document.
	Rotates current page clockwise.
	Displays the first document in the batch, and the application switches to Browsing Mode.
	Displays the previous document in the batch, and the application switches to Browsing Mode.
	Displays the next document in the batch, and the application switches to Browsing Mode.
	Displays the last document in the batch, and the application switches to Browsing Mode.
	Apply global extraction.
	Copy to global knowledge base.
	Correct tables.

Table 8-21: Verification View - Indexing Mode Toolbar

8.4.4. Field Area

A form has three main elements: a label, a viewer, and a form field. A form field might be a text field, table field, checkbox, list box, or Yes/No field. A form may also contain buttons.

Figure 8-8: Field area example

- **Toggle Read only button**

  Enables/Disables the *Read only* mode for all fields. Fields that are designed as Read only for the project are not affected by this button.

- **Form fields**

  These are controls that are used to display and edit extracted data and to enter data during manual indexing. You can use form fields to create check boxes and combo boxes.

- **Check boxes**

  A toggle selection of data input, such as On/Off or Yes/No. Check boxes are derived from form fields. You can set up the caption with the text desired and select the default view.

- **List boxes**

  It contains a selection list to use when verifying an item on the document. Used during manual verification, this selection works with automatic completion.

- **Labels**

  They are captions that help users to identify form fields and – if desired – viewers and tables.

- **Viewer**

  It contains snippets of document areas, normally those that were extracted to fill fields or tables.

- **Buttons**

  They fire actions for a new script event.

- **Tables**

  Relevant when table extraction is configured. The Web Verifier form supports multiple tables. However, even if you defined multiple tables, you can only display the first table on the verification form. You can display different tables on different forms.. For more information on tables, please see chapter Working with Tables

In the field area, the following markers are used to indicate the nature of the field:

- The currently selected field is indicated by a white background with a red or green frame, depending on whether the field is invalid (red) or valid (green).
- Valid extracted fields have a green background and are framed blue.
- Fields that need to be validated because they were extracted with low confidence have a red background, and are marked by a red triangle in the right upper corner.
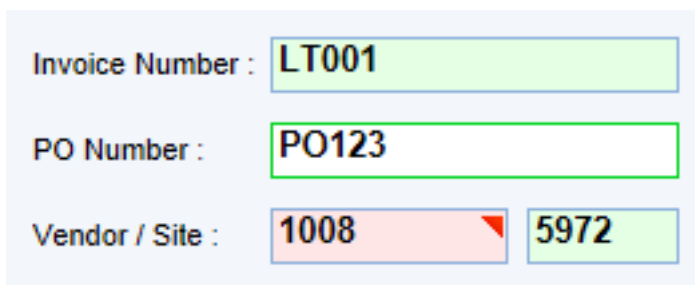
Figure 8-9: *Field markers*

Table fields provide additional highlighting options, as shown in FIGURE 8-.

Figure 8-10: Highlighting for table fields

- Clicking on a table cell will highlight the appropriate document area.

Navigation within the field area can be done using one of the following methods:

- With the mouse. This method does not affect the validation state of a field.

- By pressing the TAB key. This method gets you to the next field, but not to the next document. This method does not affect the validation state of a field.
  The order that the TAB key moves through the form is part of the form's design.

- By pressing the ENTER key. This method validates the entire field or the next invalid character within a field. Once the field is corrected, it is validated and the focus moves to the next field that requires correction. This field may also be within another document.

8.4.5.   Document Area

The document area shows the currently selected document or page along with highlights.

- Red framed areas indicate the selection for the currently selected field.

Figure 8-11: Document area – selected field highlighting

- Yellow areas were considered as candidates, but another candidate seemed more likely. If the extraction result is invalid or wrong, these areas may point to the correct indexing data.

- If you launch the table correction mode, the unlearned lines are shown in grey.

Figure 8-12: Document area – table correction, unlearned lines

- As soon as the lines are learned, the highlighting of the correctly recognized lines will turn to green (or blue if the line is selected with low confidence). Please see chapter WORKING WITH TABLES for further information.

Figure 8-13: Document area – table correction, learned lines

<u>Note:</u> *The highlighting colours may vary due to customization.*

8.5   Printing

The user can print the currently opened image document by selecting *File*, *Print….*

Web Verifier allows all the pages of the currently selected document to be printed by selecting the *All* button, or a desired page or page range to be specified in the *Pages* field of the Print dialog.

<u>Note:</u> *This function is available in all modes of Web Verifier (classification verification, extraction verification and document browsing mode) with the exception of batch browsing mode.*

8.6   Printing of Verified Data Content

8.6.1.   Description

The amount of a printed form's data can be configured from the *Print Setup* dialog in the *File \ Print Setup…* menu item of Web Verifier
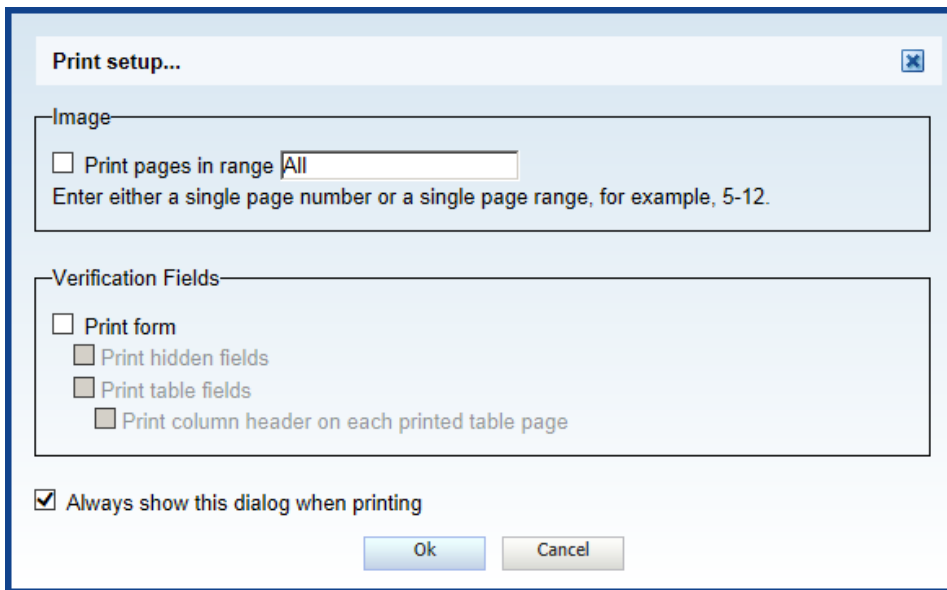
Figure 8-14: Print Setup

Here is the description of the available options:

Option Name	Description
Print image	It allows user to select the desired range of document file pages to be printed in addition to the form content (if enabled in user settings).
Print form	Activates printing of the verification form (turned on by default). The following options below are enabled only if the present one is activated.
Print hidden fields	When selected, Web Verifier prints not only the fields visible on the current verification form, but all the fields available in the loaded document.
Print table fields	When deselected, Web Verifier does not print table fields (this option might be useful for quick printing of documents with long tables).
Print column header on each printed table page	Enabled only if "Print table fields" option is turned on. When turned on (default), Web Verifier prints column header on each page (this option is useful for printing of long tables).
Always show this dialog when printing	If this check box is checked, the Print Setup dialog will show when users press Print.

Table 8-22: Print Setup – available options

Save your Web Verifier Settings if you would like your current *Print Setup* preferences to be restored next time you start WebCenter Forms Recognition Web Verifier application.

When you have configured your printing settings, Web Verifier application will print the desired field names (using display name property for each printed field) and the textual content of the fields.

The order the fields are going to be printed is defined by the custom fields' order configured in the *Form Design* mode of the WebCenter Forms Recognition Designer application.

In addition to the fields' content, Web Verifier also prints the document file name and currently assigned document class name in the header of the printed information.

9   Working with Web Verifier

This chapter provides step-by-step instructions for the main tasks that can be carried out with WebCenter Forms Recognition Web Verifier. We recommend that you read 8 before you start working with Chapter 9.

Consider the following:

- Documents that have not been classified are displayed in the classification window. Once you have assigned a class, the indexing window will be displayed. All fields are empty and you need to do the indexing manually.

- Documents that have been classified correctly, but which have invalid extraction results, are displayed in the indexing window. You need to correct the extraction results.

- Documents that have been classified incorrectly are displayed in the indexing window. Select *Classify Document Manually* from the Options menu to open the classification window. Correct the class and confirm by pressing ENTER. This displays the indexing window. Usually, the fields will be empty because documents belonging to different classes normally do not have the same set of fields. In most cases, you need to do the indexing manually.

9.1   Page Separation Workflow in Web Verifier

9.1.1.   Web Verifier Settings for "Document Separation"

The only entry pertaining to "Document Separation" is in the "Workflow" tab. Here, define the input state for "Document Separation". The final step is to activate this workflow step by clicking *Document Separation*.
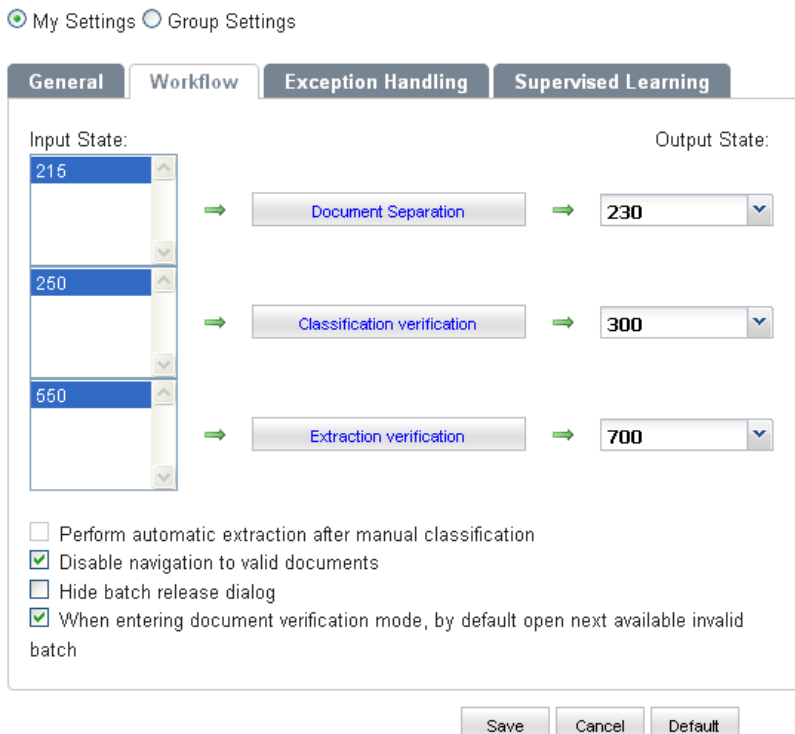
Figure 9-1: Settings for Document Separation

9.1.1.1. Manual Correction of Automatic Document Separation in WebCenter Forms Recognition Web Verifier

If during automatic document separation in Runtime Server there was at least one unsure page-level decision for a batch of documents, the whole batch gets the state "failed document separation". Such a batch is supposed to be manually reviewed and, if required, corrected in WebCenter Forms Recognition Web Verifier applications. It is labeled by the following icon:

The automatic document separation results can be corrected in Document Browsing mode of the Web Verifier application. When the next batch is opened, the system automatically displays the first uncertainly split / merged page:

Here, the user has the following options:

- Toggle the unsure status ("Accept / Reject Next Unsure Page" menu command). This command sets the page to "manually accepted" state or to "manually rejected" state respectively. There are 3 different states of page correction status: blue page icon for extracted with high confidence by the engine, blue page icon with a red question mark for extracted with low confidence by the engine (unsure), and blue page icon with green check sign for manually accepted / corrected by the Verifier user. These states remain the same after the user closes the batch in Web Verifier and can be reviewed by the other users. If all pages of a document become accepted (the pages extracted with high confidence are accepted by default), the document is redirected to successful document separation state (in the example above – "230"). Otherwise, if at least one of the document's pages becomes manually rejected, the whole document gets the lowest "page separation failed" state configured in Web Verifier settings (in the example above – "215").

- Split the document into two separate documents ("Cut Document" menu command). The "top document" receives all the pages above the currently selected one while the "bottom document" receives all the pages below, including the currently selected. In this case, the currently selected page and the preceding page automatically get "manually accepted" page correction status.

- Merge selected document with the previous one ("Append Document" menu command). In this case, the first page of the currently selected document and the last page of the preceding one automatically get "manually accepted" page correction status.

- Go to the next unsure page ("Select Next Unsure Page" menu command). This action selects the next unsure page to verify (the one with a red question mark) without changing any page states.

When verifying the correctness of automatic page separation in WebCenter Forms Recognition Web Verifier, the user can switch to a more convenient page view mode, for example having two consecutive pages displayed simultaneously:
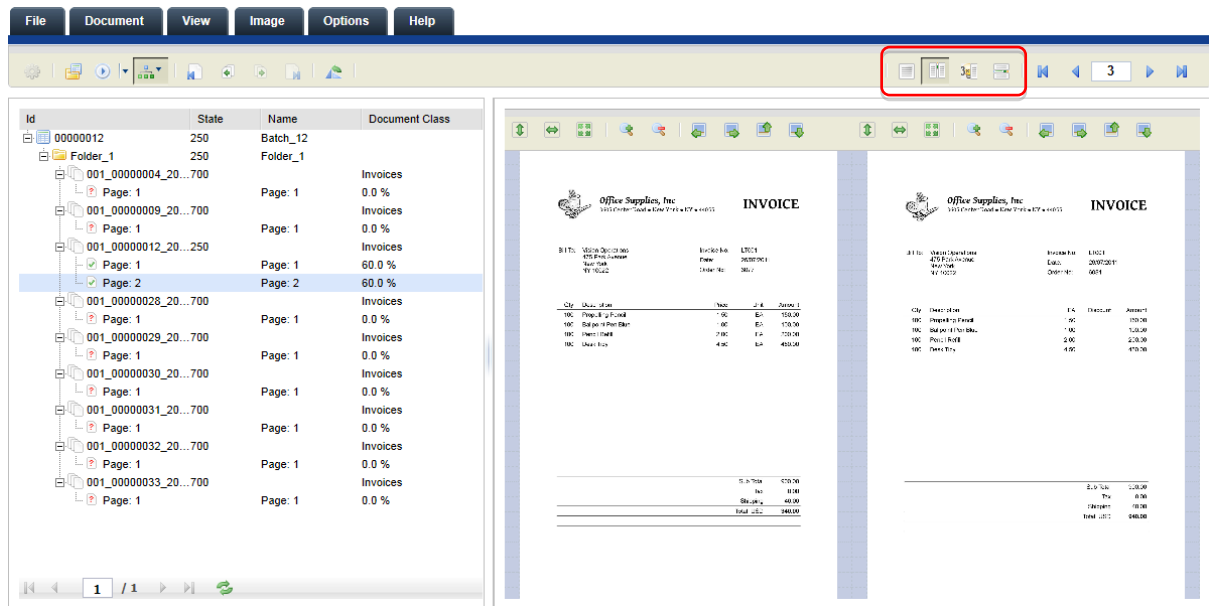
Figure 9-2: Verifying automatic page separation – View modes

There are 4 different page view modes available in WebCenter Forms Recognition Web Verifier:

- Single page (default)

- Two pages displayed horizontally

- Three pages displayed horizontally

- Two pages displayed vertically

  They can also be accessed via *View → Multi-Page View*

Page reordering

It is possible to reorder the pages. This can be done within the document tree by dragging the pages to the proper position.

To move a page to another position:

1. Perform a long left-click on the desired page. A tooltip appears.
2. Drag the page to the new position. A dynamic green position line indicates the actual position.
3. The document has to be reprocessed. Click *Yes* on the notification shown below to accomplish pages reordering.
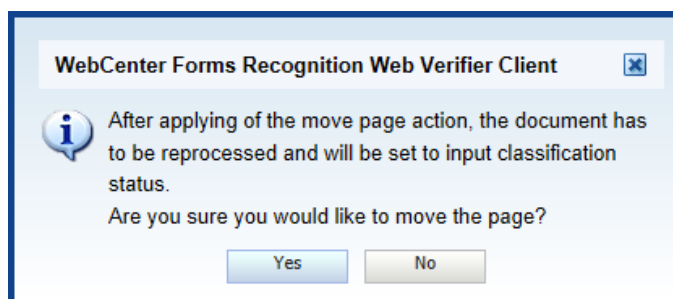
Figure 9-3: Page reordering – Notification for reprocessing

9.2   Manual Correction of Classification Results

Manual correction of classification results is done if the WebCenter Forms Recognition Web Verifier workstation is configured as follows:

- Classification verification is enabled.

- Extraction verification is disabled.

To determine your settings, check the workflow tab of the WebCenter Forms Recognition Web Verifier Settings dialog box. (See section 7.3 SETTINGS – WORKFLOW)

If you do this task regularly, you may want to apply the appropriate filter in the Batch View using the menu command View - Batch Filter - Batches to verify, classification only.

To correct invalid classification results:

1) In Batch View, check the state column to find a batch you can verify. Use the arrow keys to navigate and select a batch.

2) Once you select a batch, press ENTER to open the Verification View. The Verification View opens in Verify Mode, with the first invalid document being displayed. The cursor is already placed in the classification list box.

3) To select a class, either:

- Click on the arrow on the right side of the list box to open the list and then select a class.

- Use the arrow keys to browse the list of classes and make your selection. The entries in the list are sorted alphabetically.

- If you know the correct class name, type its first characters and wait until the system automatically displays the full class name.

4) To confirm your selection, press ENTER. The application validates this document and its state increases. The next document requiring verification is displayed automatically. Proceed as described above.

5) When all documents in the batch are validated, the application prompts you to select what you want to do next.
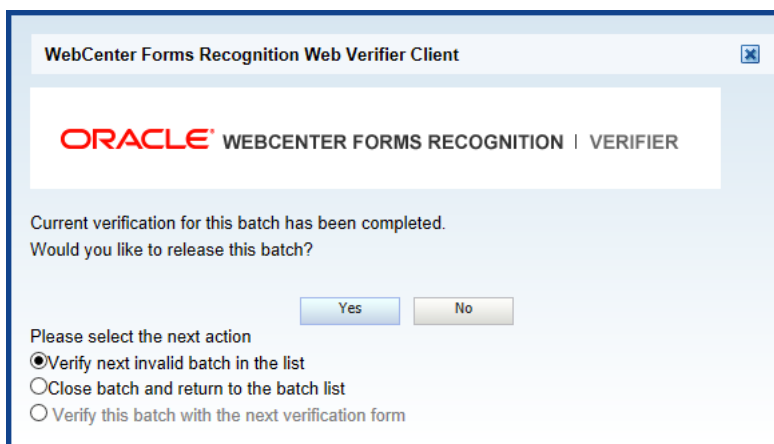
Figure 9-4: Finishing a batch

6) The following choices are displayed:

- Verify next invalid batch on the list: Releases the current batch and opens the next batch that needs verification.

- Close batch and return to the batch list: Releases the current batch and displays the Batch View where you can select the next batch.

– Verify this batch with the next verification form: Change verification forms.

– Confirm your choice by clicking *Yes* or *No*.

9.3 Processing of Documents Classified to No Longer Existing Document Classes

9.3.1.1. Description

The Web Verifier application is able to correctly process (open) documents classified to non-existing (previously removed) document classes using internally saved information about the former parent class assignment.

Note: The documents classified to non-existing document classes (can quite often be the case in context of supervised learning workflow) can only be processed if their former parent class still exists in the project the document is being processed with.

9.3.1.2. Usage

The present feature is very useful in context of supervised learning workflow (using Advanced Verifier), where the so-called "vendor" class is quite often deleted (or not inserted) from the global project's configuration.

9.4 Manual Correction of Extraction Results

Manual correction of extraction results is done if the WebCenter Forms Recognition Web Verifier workstation is configured as follows:

- Classification verification is disabled.
- Extraction verification is enabled.

To determine your settings, check the workflow tab of the WebCenter Forms Recognition Web Verifier Settings dialog box. (See section 7.3 SETTINGS – WORKFLOW)

Note: If you do this task regularly, you may want to apply the appropriate filter in the Batch View using the menu command View - Batch Filter - Batches to verify, extraction only.

9.4.1. Correcting Invalid Results

To correct invalid results, do the following:

1) In the Batch View, check the state column to find a batch you can verify. Clicking on it will immediately start the verification.

2) The Verification View opens in Verify Mode, with the first invalid document being displayed. The cursor is already placed in the first invalid field.

9.4.1.1. Form Elements and Field Types

A form has three main elements: a label, a viewer, and a form field. From a form field, you can select a text field or table field. Using a text field or table field, you can create check boxes or combo boxes. The field types for validation include Read Only, Multi-line, Combo Box, and Check Box. You can also add a button to a form to fire actions.

Elements of a form can include:

- Form fields display extracted data. You can also enter and edit data during manual indexing. You can use form fields to create check boxes and combo boxes.
- Labels identify form fields, viewers, and tables.

- Viewers are sections of document areas, normally those that were extracted to fill fields or tables.

- Buttons fire actions for a new script event.

- Tables extracted from documents.

  The following is a list of field types and their descriptions:

- **Read Only**: When selected, information on a field is dimmed and cannot be selected or edited.

- **Multi-line fields:** Required in the context of address analysis, but can also be useful in other cases. A multi-line field enables line wrap and displays a vertical scroll bar, if required.

- **List Box**: A drop-down box that lists predefined strings related to the verification document. It can either show the nearest values automatically or show only selected values.

- **Check Box**: A toggle selection for one of two choices of the data input for a field. Example: Yes/No.

9.4.1.2. Editing Text Fields

WebCenter Forms Recognition Verifier includes automated features for editing text fields that can speed up text entry and correction. Options for character changes include multi-line fields, combo boxes, and check boxes. You can also insert and replace text in cells and fields, either in single words or blocks of text, using drag and drop, or by double clicking on the selected text.

Multi-line fields are necessary for address analysis, but can also be useful in other cases. A multi-line field enables line wrap and displays a vertical scroll bar, if required.

A Combo Box lists predefined strings related to the verification document. To aid in verification, you can select from the list of strings.

The Check Box provides an either/or option that toggles table data entry choices on and off. For example, a Yes/No check box; checking Yes would bring up data entry related to the verification, and unchecked for No would hide them.

9.4.1.3. Auto-Completion

Auto-completion helps to speed up typing. When you start to type, auto-text completes the word, suggesting the best match among all of the words or candidates available after OCR and Format Analysis. For example, you can type the first two characters of a 20-character invoice. The auto-text feature finds the best matched candidate suggested by the Format Analysis engine and places it in that field.

The auto-selected text also appears highlighted in the original document. Select whether a single-line or a multi-line text field should be displayed. To override auto-completion, continue typing the desired text.

The auto-completion feature for a header field is supposed to automatically select the best candidate from the available ones, which works within *Highlight Candidates* mode. However, the viewer will be updated only if the candidate appears once in the document; otherwise, when the word for the field is inserted by auto-completion, the viewer will be blank.

Note: Auto-completion does not work on formatted text and characters that were incorrectly read by OCR.

9.4.1.4. Inserting Words in Fields

To speed up verification, you can insert words to replace or append text. The method for inserting words depends on the availability of candidates. A candidate is one that matches the learned words for that field. You can insert words in fields or table cells. You can append or insert words and use the mouse to append or replace the field.

Words with Candidates

If the word is a candidate for a field, you can copy it into the field box. A candidate is one that matches the learned selections for the field.

- To copy text to the field box, click on the desired text you want to copy. A box appears around the word. Double click on the box or right click in it and select *Copy to Current Field*.
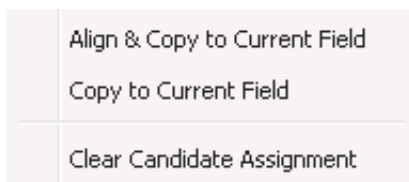
```
Align & Copy to Current Field

Copy to Current Field

Clear Candidate Assignment
```

Note: You can insert only one candidate per field per document verification session.

Make sure that this word fits the format analysis rules defined for that field.

Words Without Candidates

Even if the word does not belong to any candidates for the field, you can still use it to fill a field. For example, a field named "sales total" might be replaced by "invoice total."

- To use text for filling a field, drag a box around the desired word or click on it. Double click on the desired word in the box, or right click in the document and select *Align & Copy to Current Field*.
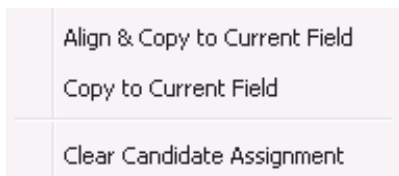
```
Align & Copy to Current Field

Copy to Current Field

Clear Candidate Assignment
```

Note: You can insert only one candidate per field per document verification session.

Make sure that this word fits the format analysis rules defined for that field.

9.4.1.5. Inserting Blocks of Text

Inserting large blocks of text with minimal mouse movement is helpful when you have multiple word data verification elements for fields such as address information or cell descriptions. Before you can insert blocks of text, first select the settings in the Workflow dialog box to immediately copy information. (See section 7.3 SETTINGS – WORKFLOW)

To insert large blocks of text:

- Click and drag over the desired text in the image viewer.

- Release the mouse button. A rectangle appears around the text. Adjust the rectangle by selecting the nodes at any corner, if necessary.

- Drag and drop the rectangle to the desired field or table cell. A copy of the rectangle appears over the field or table cell.

Or
Double click on the rectangle. The text in the rectangle replaces the text in the field or table cell.

Note: You can move or resize this rectangle by clicking in the area in the image viewer. When the rectangle appears, select the nodes to resize it, or drag it using the drag and drop method described above.

9.4.2.   Finishing the Validation

- Once a field is corrected, press ENTER to validate it. Once the validation is finished, the cursor automatically moves to the next invalid field, regardless of whether this field is in the same or next invalid document. If you leave the document this way, it is validated automatically. In the next field, proceed as described above.

    When all documents in the batch are validated, the application prompts you to select what you want to do next.

    Select the next step by clicking *Yes* or *No*, or click *Details.* The following choices display:

- *Verify next invalid batch on the list*: Releases the current batch and opens the next batch that needs verification.

- *Close batch and return to the batch list*: Releases the current batch and displays the Batch View where you can select the next batch.

- *Verify this batch with the next verification form*: Change verification forms using the next verification form.

9.5   Manual Correction of Classification and Extraction Results

Simultaneous correction of classification and extraction results is done if your workstation is configured as follows:

- Classification verification is enabled.

- Extraction verification is enabled.

- Automatic extraction after classification is disabled.

    Note: If this option is enabled, extraction will be carried out automatically by WebCenter Forms Recognition Runtime. In this case, classification verification and extraction verification are two separate steps.

    To determine your settings, check the workflow tab of the WebCenter Forms Recognition Web Verifier Settings dialog box. (See section 7.3 SETTINGS – WORKFLOW)

Note: If you do this task regularly, you may want to apply the appropriate filter in the Batch View using the menu command View - Batch Filter - Batches to verify.

9.6   Manual Indexing

Manual indexing is done if WebCenter Forms Recognition Runtime was not configured to do the extraction step. Your input consists of batches with valid classification results, but no fields have been filled so far.

For manual indexing, your WebCenter Forms Recognition Web Verifier workstation is configured as follows:

- Classification verification is disabled.

- Extraction verification is enabled.

To determine your settings, check the workflow tab of the WebCenter Forms Recognition Web Verifier Settings dialog box. (See section 7.3 SETTINGS – WORKFLOW)

Note: *If you do this task regularly, you may want to apply the appropriate filter in the Batch View using the menu command View - Batch Filter - Batches to verify, extraction only.*

To index a document manually:

1) In the Batch View, check the State column to find a batch you can verify. Then click on a batch to select it.

2) The Verification View opens in Verify Mode, with the first document being displayed. The cursor is already placed in the first field.

3) Enter the value for the first field.

4) Press *Enter* to validate your entry. The cursor will automatically move to the next field, regardless of whether this field is in the same or next invalid document. If you leave the document this way, it is validated automatically. In the next field, proceed as described above.

5) When all documents in the batch are indexed, the application prompts you to select what you want to do next.

6) The following choices display:

– Verify next invalid batch on the list: Releases the current batch and opens the next batch that needs verification.

– Close batch and return to the batch list: Releases the current batch and displays the Batch View where you can select the next batch.

– Verify this batch with the next verification form: Changes verification forms, using the next verification form.

7) Move to the next step by clicking Yes or No.

9.7   Checking Entire Batches

To browse through all documents in a batch:

1) In the Batch View, use the status value to determine a batch you can browse through. Select a batch by clicking on it.

2) The Verification View with the first document requiring correction is displayed.

3) To display the first document in the batch, press the following button:

4) You may encounter a document that has been classified incorrectly. To correct this result, select *Reclassify Document Manually* from the Options menu to open the classification window. To correct the class, select the corresponding entry from the list box at the bottom, then confirm by pressing ENTER. This displays the indexing window again.

5) To correct extraction results, type your corrections into the corresponding field. If a field has been changed, its state is set to invalid. Press ENTER to validate the field you modified.

6) To get to the next document, press the following button:

Proceed as described above until the last document is reached.

10   Working with Tables

Note: *You can correct invalid cells the same way you would correct an invalid text field.*

In WebCenter Forms Recognition, it's possible that any given table was either trained by an extraction engine – Brainware Table Extraction – or the traditional Table Analysis Engine. A Verifier user does not necessarily know which one was used, and the user will not be able to see a difference in Web Verifier. The process of table fields is similar, regardless of whether you're using Verifier or Web Verifier.

10.1   Automatic Training and Extraction of Verified Table Data

Brainware Table Extraction supports automatic learning of verified table data. Brainware Table Extraction trains documents using only the information in verified table data, and the content and position of every data cell.

10.2   Traditional Training and Correction Methods

10.2.1. Inserting Words in Table Cells

You can insert single words in table cells.

10.2.1.1.   Words That Are Candidates for Cells

If the word belongs in a cell area, you can copy it to the cell. The Copy feature fills the cell with the desired text. Or, you can replace text.

To replace text with the new text, double click on the desired word or right-click in the image viewer and select *Copy to current field*. If you have candidates, double click on the desired candidate to replace it, or right-click in the document, and then select *Copy to current field* from the shortcut menu.

10.2.1.2.   Words That Are Not Candidates for a Cell

If the word does not belong to cell areas, it will display in orange when selected. Even if it is not a candidate, you can append or replace the word. Appending places the text in the best location, either right or left of the word, by text or location of the word. For example, a cell named "C2658" might be appended by "number." Or, you can replace the cell text and location by the text and location of a word. To append text with the new text, double click on the desired word or right click in the image viewer and select Append Cell Text by Word. To replace text, select the word, select *Replace Cell Text by Word* in the shortcut menu.
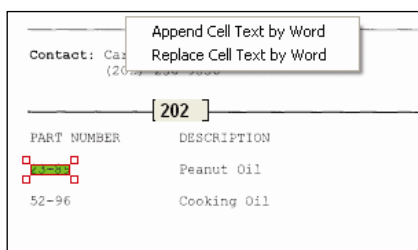
Figure 10-1: Table cell, word not in cell area

10.2.2. Correcting Table Structure

You may also need to correct the table structure. Under the table within the verification mask, you will find buttons for modifying the table structure.

Figure 10-2: Correcting table structure

The available commands are summarized below.

Button	Command	Description
	Delete all rows	Deletes all rows from the table. You will need to relearn the lines on the invoice.
	Insert new row above	Inserts an empty row above the current one.
	Delete selected row	Deletes the currently selected row. To select a row, click on its number. You will notice that the appropriate line on the invoice will be unlearned and loses its color highlighting.
	Append	Appends an empty row at the bottom of the table.

Table 10-1: Table editing

10.3 Brainware Table Extraction and Correction

Note: This functionality is available for the Supervised Learning Verifiers.

10.3.1. About Brainware Table Extraction

The learning process for the Brainware Table Extraction engine consists of two phases:

- Learning lines
- Learning mappings of columns

These are discussed in detail in the following sections.

10.3.1.1. Learning Lines

The Brainware Table Extraction engine considers the following main types of the lines:

- **Primary line**: A line that defines table structure. The BTE engine applies advanced and precise similarity analysis for all primary lines. It is important that all primary lines are well-structured and that they look similar in many of the rows to extract. The engine easily supports an unlimited number of types of primary lines for one table definition. The

primary line must contain at least four words. Otherwise, the BTE engine will not learn it. Also, the primary line must be the first line in the table row.

- **Secondary line**: A line between primary lines. The engine applies smooth similarity analysis for these types of lines, which is possible because BTE only searches the area between two neighboring primary lines. This allows the engine to extract data that varies widely, which often happens with multi-line descriptions. There is also no limit to the number of words in secondary lines, and no limit to the number of secondary lines. However, a document's page must have at least one primary line; otherwise secondary lines on this page will not be extracted.

- **Wrong line**: A primary line that is learned as a negative line sample. In other words, all lines classified by the engine as members of one particular "wrong" line class will not be extracted. In principle, it is possible to learn an unlimited number of wrong lines, though the current restriction is that this will only take effect during in-document learning. Cross-document learning (that is, learning the whole document after all the fields are completely valid) may not automatically train the wrong lines.

  After it learns any type of line, the BTE engine automatically creates and manages a new line class (cluster.) Afterward, all lines in the document considered by the engine to be members of the line class (similar to the learned line sample) will be extracted, or not extracted in the case of "wrong" lines.

Note: It is possible to learn an unlimited number of different line classes. However, the overall quality may suffer if too many lines are learned.

Learning lines can be applied in lines learning (or lines highlighting) mode. Mapping of the column data in the lines can be done in column mapping learning (or columns highlighting) mode. The user can switch between learning (highlighting) modes via the Switch Table Highlighting menu option in Web Verifier Options menu or via the context menu options *Show Lines* and *Show Columns*.

If you have documents with many line items, you will benefit from the table navigation feature. In the web.config, you are able to define how many lines should be displayed on one table page (please refer to the Installation Guide). If this number is exceeded, you'll be able to browse between the table pages by using the table controls below the table.

10.3.1.2. Learning Mappings of Columns

When learning columns' mapping, the user trains the engine on how the data from the extracted lines must be mapped to the user's table data. For primary lines, this mapping can be defined differently for different line classes. For example, if a user learned two different line samples that went to two different lines classes internally in one document, the user can then map "Unit Price" in the document to the "Unit Price" data column and the "Total Price" to the "Total Price" for the first line sample. For all lines of the second line type, the user can map "Unit Price" to "Total Price" and "Total Price" to "Unit Price." For the next document, the BTE engine will always use mapping rules #1 for the lines classified to the first line type and mapping rules #2 for the lines classified as the second line type.

If you have several BTE tables in one WebCenter Forms Recognition class, the Learnset is shared between these tables. In other words, if you used interactive learning for one BTE table, cross-document learning (which happens if the system added the document to the Learnset after document validation) will be applied for all BTE tables in the document.

10.3.1.3. Correcting Fields in Tables Created with Brainware Table Extraction

Note: Because of the way interactive table verification works, you cannot manually delete data from a cell. Rather, if you want to discard cell data, un-map the column and re-extract the table to re-map the column. Although it will seem as if you deleted the data, the data will actually still be there until you un-map the column.

Anytime you train a table interactively, do the entire required training first and then verify manually.

Brainware Table Extraction can train line types and column mapping for each type of line.

When working with interactive table extraction learn lines before you map columns.

10.3.2. How to learn tables with Brainware Table Extraction (Standard Method)

This section describes the simplest way to use interactive BTE learning. If this method does not work, proceed to the advanced method described in the following sections. This recommended method consists of six steps:

4. Show the first row sample.
5. Learn mapping in the learned row.
6. Learn missing lines.
7. Learn and adjust the mapping of missing or wrong columns.
8. Manually correct the table date and validate the table.
9. Learn the document.

These steps are discussed in detail below.

10.3.2.1.   Step 1: Show the First Row Sample

1) Select your BTE table by clicking any table field inside the table grid.

2) Click the *Correct Tables* button.

3) In the lines highlighting mode, use the *Learn As Row* function to show the row sample. This function will automatically learn the first line as a primary line and the rest of the lines as secondary lines. This function is also available by double clicking on the selected row area. Select the whole first row and learn it.

Note: The visual indicators for valid, invalid and questionable table lines are the same as for header fields: Valid lines have a green check mark; invalid lines have a red X, and questionable fields have an orange question mark.

Turning off the Correct Tables button will cause the loss of learned rows and columns. Turning the Correct table button back on will lead to the loss of mapped table data.

Learning Lines as Primary Lines

1) Right click on any line marked in gray in the TIFF.

2) On the shortcut menu, select Learn Line.

Learning a Block of Lines as Primary Lines

1) In the document viewer, draw a rectangular selection over the primary lines in a single row.

2) Right click on the selection.

3) On the shortcut menu, select *Learn as Primary Line(s)*.

All correctly selected primary lines will be learned, and all other lines are similarly extracted and displayed.

If some lines were not extracted, try relearning the lines alone or in a block.

Learning a Lines Block as a Table Row

1) In the document viewer, draw a rectangular selection over the required multi-line (or single-line) table row.

2) Double click or right click on the selection.

3) From the shortcut menu, select *Learn as Row*.

All correctly selected primary lines will be learned, and all other lines are similarly extracted and displayed.

If some lines were not extracted, repeat the procedure described above.

Do not try to learn the rest of the missing secondary or primary lines now. This is because mapping is defined on the basis of line type. If you were to train all different line samples now, you would need to learn the columns mapping separately for every line class. If you first learn the column mapping for the row you just learned, next time you learn another line sample, the engine will try to apply existing mapping rules for the newly learned row automatically, significantly reducing the time to train the table.

10.3.2.2.　　Step 2: Learn Columns in the Rows You Learned

1) Switch to the columns highlighting mode now (select *Switch Table Highlighting* from Option menu or right-click on a line and select *Show column*) and mark the location of the first cell item in the row you learned..

2) Right-click on the first cell item in a column. Choose *Map Column*, select the required data column and click on the selection.

3) Repeat this step for the rest of the cell items in the first row.

10.3.2.3.　　Step 3: Learn Missing Lines

1) Switch back to the lines highlighting mode.

2) Mark the next missing row and learn it as before.

3) Repeat this step for all rows on all pages where something is missing. Go to the next step only after you are sure nothing is missing.

10.3.2.4.　　Step 4: Learn and Adjust the Mapping of Missing or Wrong Columns

1) Return to columns mapping learning mode and look for wrong or missing mapping. Correct any missing mapping.
2) If you can't map the missing columns, switch back to the lines highlighting mode and try to learn the row where the mapping is missing.
3) Switch to columns highlighting. If the mapping is still missing, mark the missing part and map it.

Note: The BTE engine may determine the mapping automatically.

Repeat these steps until the data is completely extracted or cannot be learned correctly. (There is always a chance that you will not get 100 percent extraction results.)

10.3.2.5.　　Step 5: Manually Correct the Table fields and Validate the Table

Now switch to cells highlighting mode and manually correct missing data, OCR errors, etc. Do not use interactive learning anymore because every BTE learning action will reactivate extraction and will replace all your manual inputs instantly.

To validate the table press the *Enter* key.

Note: The only requirement for cross-document learning is correctness and completeness of the table data to train. This means that location and content of every cell item should be correct. Also, ideally, the content of cell items should not be formatted.

10.3.3. Advanced Learning with Brainware Table Extraction

10.3.3.1.   When to Learn Secondary Lines

This section discusses the special cases in which it is necessary to use secondary lines explicitly. There are two such cases:

- Case 1: Table row begins on one page and ends on the next.

- Case 2. Learning of unmapped secondary lines leads to unwanted extraction.

Case 1: Table row begins on one page and ends on the next

If a table row begins on one page and ends on the next page, you must use the *Learn as Secondary Lines* function (in lines learning mode) to train missing secondary lines (on the next page.) In this case, these secondary lines will be placed right before the first primary line on the page. Mark all the secondary lines as before: Right click and select *Learn as Secondary Lines*.

Note: Never use the Learn as Row function in this case, as this will tell the engine that the first secondary line is actually a new sample of primary line. As a result, the engine may split extracted table data into new rows.

Case 2: Learning of unmapped secondary lines leads to unwanted extraction

Your project may require that data from secondary lines not be extracted. Usually, this will not be a problem, but sometimes the engine extracts the data from these lines anyway. In this case, not learning these secondary lines will prevent unwanted extractions. Use the *Learn as Secondary Lines* function instead of *Learn as Row* if you would like to learn just selected lines and not all lines that belong to the row. You can also Unlearn Line to correct or adjust the extraction.

How to Learn a Block of Secondary Lines

1) In the document viewer, draw a rectangular selection over the required secondary lines of a desired multi-line row.

2) Right click on the selection.

3) On the shortcut menu, select *Learn as Secondary Line(s)*.

All correctly selected secondary lines will be learned, and all other lines are similarly extracted and displayed.

If some lines were not extracted, repeat the procedure described above.

10.3.4. Advanced Learning: Additional Functions

This section discusses two additional functions: Unmap Column and Unlearn Line.

10.3.4.1.   Un-map Column

This feature is not available in this version of the Web Verifier.

10.3.4.2.   Unlearn Line

The Unlearn Line function can be used to discard previously applied learning for a particular line. To do this, Brainware Table Extraction uses a line sample, searches for the line type and removes the line type from the Learnset.

1) Switch to lines learning mode and right click on the line you want to unlearn.

2) On the shortcut menu, click *Unlearn Line*. Unlearned lines change from green to gray.

10.3.4.3. Learn Line as Wrong Line

Learning a Wrong Line means to train the table such that a particular line will not be extracted. This applies to other lines of the same type in the table.

1) Right click on any learned line or draw a rectangular selection over the required lines.

   On the shortcut menu, select Learn as Wrong Line. The selected lines and lines similar to it are now highlighted in gray. Information from these lines will not be extracted.

11   Tips for Tricky Situations

My document contains an invalid extraction result. However, this result is precisely what I need, and I want to validate the field. What can I do?

This depends a bit on the design of your application. In most cases, you will have to press Enter three times.

In one of my batches, there is a document that must be classified manually, but it does not belong to one of the available classes. I cannot release the batch as it is. What can I do to finish my job?

Normally, your organization will have specialized workstations where people are in charge of handling special cases that only occur as exceptions. For more information about exception handling, refer to section *7.4 Settings – Exception Handling*.

In one of my batches, there is a document I have already validated. However, I've overlooked a mistake in this document. I don't want to release the batch without correcting it.

You can use the Document Mode (See section *8.2.5 Document Area*) to get to the document. Select the document and switch to Verify Mode. Make corrections and press Enter.

Sometimes the indexing window looks weird: It has no field area, only the current input area. How do I get to the next field?

Not a problem. You can use all keyboard shortcuts for field navigation from within the current input area.

When I switch from one field to the next, the document is not moving as well. I find this annoying. Is there a way to stop that?

Yes, there is. With your current settings, the application always searches the document area associated with the current field's content. This area is then displayed. To turn this off, click on *keep focus*: Alternatively, you could just use a different magnification ratio.

I want to change the Default Colors, Background Colors & Fonts for Elements of Verification Forms. Can I do it?

WebCenter Forms Recognition supports a set of script methods to dynamically or statically adjust fonts, colors, and background colors for verification forms and their verification elements. Please refer to the Scripting Guide documentation for more details.

Appendix A   Quick Reference

Main Controls		Verification View		Document View	
	Batch View		Exception State		Fit to Height
	Start Verification		Highlight Candidates		Fit to Width
	Batch Structure		Keep Focus on Field		Fit to Size
	Display Properties		Keep Zoom		Zoom In
	Display Batch Filter Dialog		Previous Page in Document		Zoom Out
	Start Learnset Manager		Next Page in Document		Single Page View
	First Batch Page		Rotate Image (90 degrees)		Two Pages View horizontally
	Previous Batch Page		First Document in Batch		Three Pages View horizontally
	Next Batch Page		Previous Document in Batch		Two Pages View vertically
	Last batch Page		Next Document in Batch		Move Image Downwards
	Refresh		Last Document in Batch		Move Image to Left
			Correct Tables		Move Image to Right
			Add Current Document to Local Learnset		Move Image Upwards

			Classify and Analyze Current Document		

Glossary

Accumulative Learnset	The Common Learnset.
Administrator	In WebCenter Forms Recognition, an administrator is a power user who creates user accounts, passwords, and groups, and assigns users to groups.
Analysis	In this processing step, the document content is analyzed and a set of possible values for a field is generated. These values are called candidates.
Associative Search Engine	Uses a reference field to extract results.
Automatic Supervised Learning	Uses the Associative Search Engine to process, classify, and extract information.
Base class	The highest level of a classification.
Batch	A logical organizational structure to control a set of documents during a process. A batch is normally created during the scan process from a batch of paper. The status of a batch is used to manage the input flow.
Brainware Table Extraction	An extraction method that facilitates interactive table training.
Candidate	Set of possible values for a field.
Child class	A class spawned by a parent class. See also base class and parent class. Also called a sub-class.
Class	A set of documents that are grouped by common content. Each class usually has a mnemonic name that describes its contents from the user's point of view.
Classification	The process of assigning one or more classes and corresponding confidence values to one or more unknown documents.
Common Learnset	An accumulation of Local Learnsets.
DocClass	A parent document class.
Document	Any electronic file mainly consisting of ASCII text. If this is initially the case, OCR or filtering must be applied to create the text representation. A document can be classified, have fields used for extraction, and have one or more images attached.
DPI	Dots per inch. Affects the size and clarity of an image file.
Evaluation	The process of determining a class or the contents of a field from confidence levels, weights, or distances for classes or candidates.
Export	In WebCenter Forms Recognition, document export releases the documents so that they are no longer managed by the software.
Extraction	The process of automatically finding specified information within a document and writing the information to data fields associated with the document. Extraction is used for automatic indexing.
Folder	A logical structure inside a batch for coherent documents. For example, a folder may consist of all pages of a correspondence with many folders inside one batch.
Form	(1) A structured, standardized document that is used to support business processes. (2) A custom dialog box in a software application.
Global Learnset	A general Learnset that encompasses similar classes or projects. See also Local Learnset.
Importing	Bringing documents into WebCenter Forms Recognition for management and processing.
Indexing	The process of assigning attributes to a document. This can either be done manually, semi-automatically (Smart Indexing), or entirely automatically (Extraction).
Knowledge Base	A database of knowledge about a subject; used in artificial intelligence. The knowledge base comes partly from human experience and partly from the computer's experience.
Learnset	In classification, a Learnset is a set of documents whose class assignments are specified by the user. For each view and each class, the user must provide a sufficient number of representative documents. Similarly, in extraction, a Learnset is a set of documents whose field contents are selected by the user from a set of candidates.
Learnset Manager	A user who designs, modifies, and maintains Learnsets.
Learning	Given a view with a set of documents in vector representation and their class assignments, a neural network is created, so that the defined classes can be reproduced without error. This neural network is then used in all subsequent classification tasks.
Literal character	Normal alphanumeric characters that are not used as operators.
Local Learnset	Learnset specific to a document class.
Neural network	An artificial neural network is an application that in some ways works like a human brain. This

	includes the ability to learn. It consists of artificial neurons that are linked into a network of layers. The neural network can receive signals through an input layer, process it within the internal layers, and send signals through the output layer. During learning, a specified input (called a teacher signal, such as documents from a Learnset) and the desired output (such as the corresponding classes) are presented to the network together. Processing is then adjusted until the desired output can be produced from the teacher signal.
OCR	Optical Character Recognition. The reading and recognition of symbols of text from a piece of paper or a scanned image. OCR detects the symbols and converts them into characters and words that can be read electronically.
Parent class	A class with derived classes, called children.
Persistent	Permanent; something that is saved persistently is saved permanently, unless a user or process deletes it.
Project	Project files are used to persistently save custom settings for WebCenter Forms Recognition applications. They are created in WebCenter Forms Recognition Verifier and handed over to WebCenter Forms Recognition Runtime for productive operation.
Smart Indexing	Smart indexing uses a database lookup to determine document attributes. It can be used for automatic indexing and to support manual indexing.
Sub-class	A derivative class. Also called a child class.
Supervised Learning Verifier	A user who collects and maintains local training data.
Validation	A quality assurance task that involves confirming whether a processing result is correct. This can be done at several levels: for the class or a field associated with a document, for the document as a whole or for an entire batch.
Verification	A quality assurance task that involves checking and correcting processing results.
Verifier	WebCenter Forms Recognition's QA application.
View	A set of documents that represent at least two classes. A view is usually defined using a small set of documents that represent the domain of interest. In a view, classes compete for documents; that is, a document may only be assigned to one class within the view.
Web Verifier	WebCenter Forms Recognition's web based extension of the Verifier Thick client.
Workdoc	An internal structure representing the logical structure of a document. The Workdoc represents the data created during processing of a single document and is stored in a file with the extension *.wdc. Since the Workdoc includes all OCR and analysis results, it may exceed the document file by size.