Oracle® Forms Recognition
Verifier User Guide
10g Release 3 (10.1.3.5.0)

October 2009

Oracle Forms Recognition Verifier Guide

10g Release 3 (10.1.3.5.0)

Content

Chapter 1 About Verifier and Advanced Verifier

1.1 Overview

Oracle Forms Recognition is a product suite designed for automatically processing incoming documents. Oracle Forms Recognition can process documents from arbitrary physical sources and paper-based documents, and from electronic files, e-mail, or faxes.

What happens when Oracle Forms Recognition Verifier processes documents?

Structured or unstructured document input is obtained by scanning paper-based documents from fax servers or e-mail servers, or as files. All documents are stored on a computer's hard drive. Oracle Forms Recognition monitors specified directories on this hard drive for new documents. If new documents are detected, Oracle Forms Recognition imports them.

Imported documents are first analyzed to determine the document layout and to recognize structures such as words, lines, logos, or tables.

The documents are then classified according to predefined categories. Examples of typical categories used in classification are invoices, orders, offers, or resumes. Categories can be defined individually, depending on the needs of your organization. Using a set of sample documents, Oracle Forms Recognition actually learns to tell which category a previously unknown document belongs to.

For each category, the data relevant for further processing is different. For example, if you are processing invoices, you probably want to know the total sum to be paid. This information is irrelevant if you are processing resumes, where the applicant's name, the desired position, and the contact options are more important. Oracle Forms Recognition identifies and extracts data that is relevant for the respective document category. Again, the data that is to be extracted can be defined individually to suit the needs of your organization.

Finally, the documents, their category assignments, and the extracted information are released from Oracle Forms Recognition and written to designated export directories. The documents are forwarded to connected systems. For example, invoices can automatically be forwarded to the software system used in your company's accounting department, while resumes are sent to Human Resources.

Once the Oracle Forms Recognition application is set up, all this is done without human intervention. But what happens if Oracle Forms Recognition cannot properly process a document? There are several reasons this could happen:

- Paper-based documents might be "dirty" so that Oracle Forms Recognition is no longer able to read them.
- There might be stamps or notes on the documents that make important sections illegible for Oracle Forms Recognition.
- Oracle Forms Recognition may encounter a document from an unknown category. Since the software was not previously trained to recognize documents from this category, it will not be able to process the document.
- Oracle Forms Recognition may have been told to extract information that is missing, such as a form that was not filled in correctly.

That's where Oracle Forms Recognition Verifier comes in: Oracle Forms Recognition Verifier is the quality assurance utility of the Oracle Forms Recognition suite. The application detects all documents with processing problems and presents them to the operator for verification.

Since the verification step is done before the export step, only qualified output will leave the Oracle Forms Recognition process. Therefore, subsequent systems will only receive appropriate input.

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.
- Keyboard shortcuts are available for most operations, enabling you to get your job done quickly.
- Through automatic locking, document batches can safely be processed by teams of operators.

1.2 Verifier vs. Advanced Verifier

Advanced Verifier is an extension of Verifier that provides access to supervised learning – the interactive verification and training of Learn Sets.

If you've worked with Oracle Forms Recognition before, you already know that you can work with both data and text fields. Oracle Forms Recognition with Supervised Learning and its Advanced Verifier component dramatically improve the way you work with table data. In earlier versions, the analysis of table data was "rules-based," meaning that users could not improve upon the rules established for the analysis. Oracle Forms Recognition with Supervised Learning changes that by enabling you to interact with (or supervise) how Oracle Forms Recognition works with your data.

Chapter 2 About This Manual

2.1 Intended Audience

This manual is primarily for users of Oracle Forms Recognition Verifier. To use this application and its documentation, you need no special skills, but you should have basic knowledge of the Windows operating system and of Windows applications. Application designers and system administrators will also find this manual useful.

2.2 Reading Suggestions

This manual is organized as follows:

- ***Chapter 1*** explains the purpose and concepts behind Oracle Forms Recognition Verifier and Advanced Verifier.
- ***Chapter 4*** describes how to start and exit Oracle Forms Recognition Verifier.
- ***Chapter 7*** explains the user interface of Oracle Forms Recognition Verifier.
- ***Chapter 6*** describes how to configure Oracle Forms Recognition Verifier and Advanced Verifier. The configuration determines where the applications get their input from and which state the input has before and after verification.
- ***Chapter 8*** explains how to use Oracle Forms Recognition Verifier and Advanced Verifier for quality assurance in automatically processed documents.
- ***Chapter 9*** explains how to train table data using traditional methods and with Table Extraction.
- ***Chapter 10*** discusses the use of the Learn Set Manager.

2.3 Related Documentation

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

- ***Oracle Forms Recognition Installation Guide***: Explains how to install Oracle Forms Recognition and how to set up the licensing file.
- ***Oracle Forms Recognition Designer User's Guide***: Explains how to use the Oracle Forms Recognition Designer component to create custom application.
- ***Oracle Forms Recognition Runtime Server User's Guide***: Explains how to use the Oracle Forms Recognition batch processing application.
- ***Oracle Forms Recognition Scripting Guide***: Explains how to write scripts to enhance Oracle Forms Recognition projects.
- ***Release Notes***

Chapter 3 Understanding How 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 Oracle Forms Recognition, batches consist of electronic documents. The documents inside such a batch may be paper-based documents that have been scanned to transform them into a digital format, or files created using applications such as a word processor. Various documents are normally assigned to the same batch only because they have been received within the same time period. For example, all letters 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. Oracle Forms Recognition does the same thing with folders. However, in the context of Oracle Forms Recognition, a folder is always a structure inside a batch. This means that batches can either consist of document stacks, or they 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, of a situation, or of a 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 Oracle 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, a second stack of resumes, a third stack of orders, and so on. 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 information that is 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 Oracle Forms Recognition, indexing information is written 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 again, and the name of the supplier, the total sum to be paid, and the invoice number are written out again, 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 Oracle 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 Learn Set?

In classification, a Learn Set 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 Learn Set is a set of documents whose field contents are selected by the user from a set of candidates.

3.2 Oracle Forms Recognition and Quality Assurance

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

Batches are the basic entity Oracle Forms Recognition is working on

Oracle 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 get 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.

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, an Oracle 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, Oracle 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, or
- 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 Oracle Forms Recognition application may enforce that characters which could not be identified for certain 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 Oracle Forms Recognition and Teamwork

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

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

However, if Oracle Forms Recognition Runtime 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 Oracle Forms Recognition Verifier. Invalid batches are presented to you in a task list, the so-called Batch View. (see Section ***7.1 The Batch View***) You will have to resolve each problem and validate each correction before you can release the batch. Only after release can subsequent automatic steps be carried out. Finally, when Oracle Forms Recognition has finished processing a batch, the documents will be sent to their actual recipients.

One of the objectives of an Oracle Forms Recognition application is to get documents to their recipients as fast as possible. On the machine side, automatic steps can be distributed to several computers to ensure that no delays occur. They can simultaneously perform the same or different tasks. Similarly, on the human side, Oracle Forms Recognition supports a variety of task distribution 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 is verifying the extraction results. This can be realized by configuring Oracle Forms Recognition Verifier accordingly. (See Section ***6.2 Settings – Workflow***) In addition, several operators can carry out the same task at the same time, yet on different batches. This is possible via a locking mechanism that avoids conflicting results by making sure that a batch cannot be changed by several persons at the same time.

As a practical example consider two servers with Oracle Forms Recognition Runtime 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

---

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

steps. If Runtime_Gamma generates an invalid classification result, the corresponding batch is routed to a member of the QA team named Miller. Miller is correcting these results using the Oracle Forms Recognition 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. Oracle Forms Recognition Verifier sets a corresponding marker in the worklist.

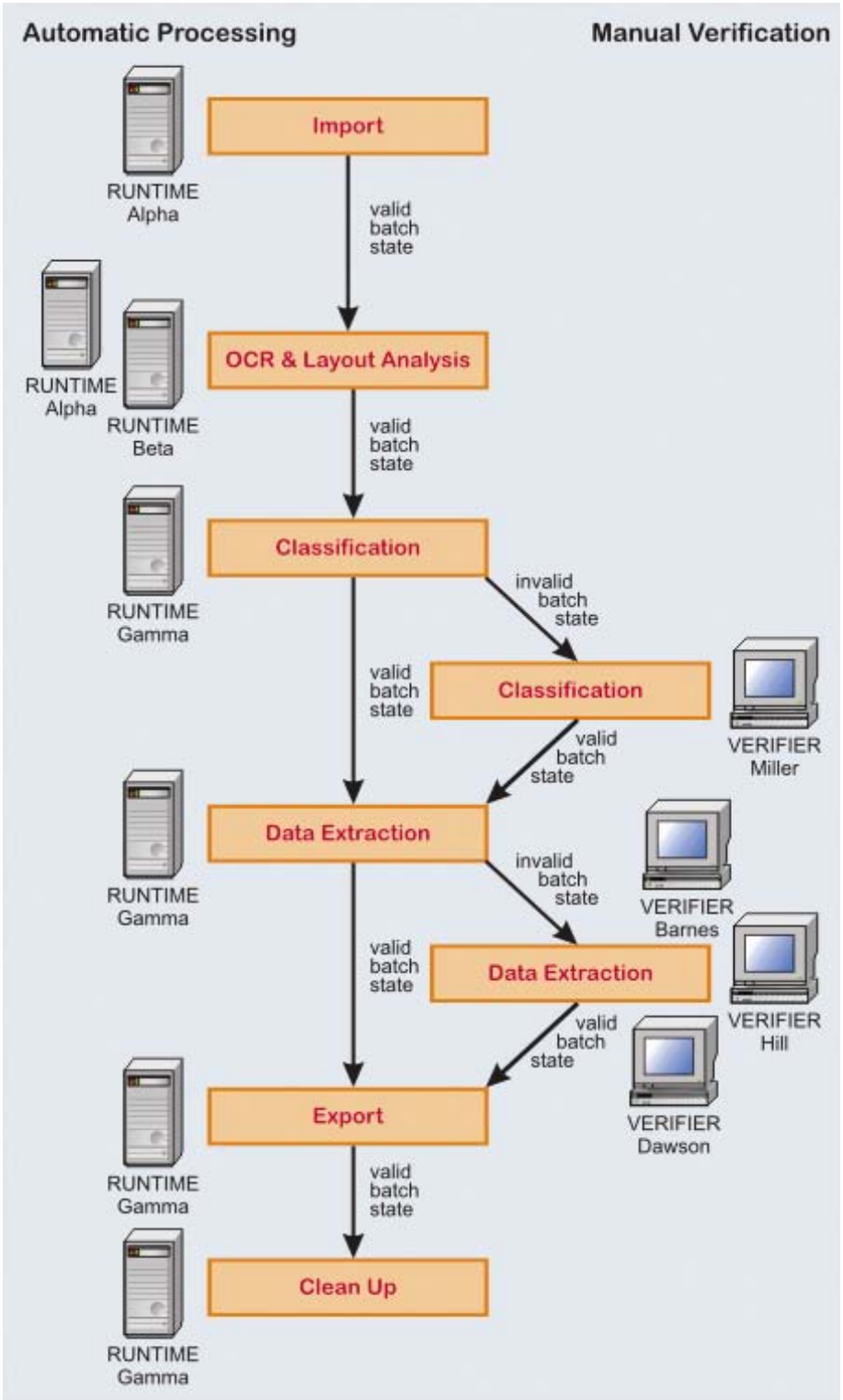


Figure 3-1: Sample workflow showing how processing steps can be distributed. This configuration involves several servers running Oracle Forms Recognition Runtime, and a number of workstations with Oracle Forms Recognition Verifier that are used by a quality assurance team.

Chapter 4 Starting and Exiting Oracle Forms Recognition Verifier

4.1 Starting

If Verifier was installed as recommended by the setup program, you can launch it from the Windows Start menu using the command sequence Start - Programs - Oracle Forms Recognition- Forms Recognition Verifier. After startup and login, the application displays the Batch View (See ***Chapter 7***).

4.2 Logging In

If the application is being launched for the first time, it is not yet configured. An error message is displayed, and the Batch View is empty. Oracle Forms Recognition Verifier still needs to be configured. This should be done by an experienced user. For instructions, please refer to **Chapter 6**.

When you log in to an existing project in Verifier, you must supply your user name and password. This password is not the same as the one you use to log in to your workstation. Instead, it is specific to Verifier, and possibly to the project. However, you probably have the same user name and password for all Verifier projects you work on. Your user name and password were assigned to you in Designer when your project administrator configured the project.

Your user name and password enable you to get into Verifier and govern what you can do once there.

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 password checkbox appears on the logon form. To remember your username and password between logons, fill in your username and password and select Remember password before clicking OK. Next time you logon to that computer, the system will fill in the username and password information automatically so that simply clicking OK will log you in.

Figure 4-1: Verifier logon with password remember feature

For more about user accounts and roles, please see ***Chapter 5***.

4.3 Ability to Specify Logon Information via Command Line Arguments

To suppress project authentication when starting Oracle Forms Recognition Verifier, logon information specified as command line arguments. The command line argument for user name is “/USR” and for password “/PWD”.

For example, the following line in a Windows batch file placed in the Oracle Forms Recognition program folder will launch Oracle Forms Recognition Verifier under John Smith’s account:

```
start /B DstVer.exe /USR "John Smith" /PWD john1234567
```

The same mechanism can be used from the Windows “Run…” menu:

```
"C:\Program Files\Oracle\Forms Recognition\DstVer.exe" /USR "John Smith" /PWD john1234567
```

If the password is empty there’s no need to specify the “/PWD” option. For example:

```
start /B DstVer.exe /USR “Guest User”
```

4.4 Exiting the Application

To quit Oracle Forms Recognition Verifier:

- On the File menu, select Exit.

Chapter 5 About Users, Groups, and Roles

To load a Verifier project, you must log in with your user name and password. If you work on more than one Verifier project, your user name and password are probably the same for all of them.

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

User name/password combinations were set up by your project administrator while configuring your project in Designer.

Your administrator also set up user groups and assigned you to at least one of these groups – perhaps more than one. In turn, the administrator assigned one or more roles to each user group.

There are five roles: Administrator, Verifier, Verifier Settings, Learn Set Manager, and Supervised Learning Verifier.

The roles of these groups follow:

- 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 five roles, because it encompasses the permissions for all other roles.
- Learn Set Manager: The Learn Set Manager role is to define, modify, and maintain the Learn Set.
- 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 Learn Set 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 Oracle Forms Recognition Verifier configuration to be changed.

5.1 Changing your Password.

1) Select Change Password from the Options menu.

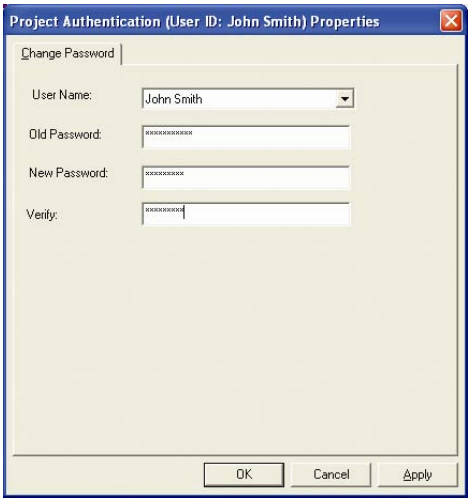


Figure 5-1: Password change dialog

2) Type in your existing password.
3) Enter the new password.
4) Re-enter the new password to verify that you typed it correctly.
5) Click OK.

Chapter 6 Configuring Verifier and Advanced Verifier

You can only change the Oracle Forms Recognition Verifier settings if you have been assigned the Verifier Settings role.

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

- The location of the batches in the file system.
- The Oracle Forms Recognition Designer project file that contains the settings used to process the documents.
- The processing steps that you want to verify: classification, extraction, or both.
- The status of batches before and after processing.

It also entails configuring 508 Compliance, but this is done at the workstation level, not the project level.

After you configure a project's settings, you can load and save them using commands on the File menu. When loading or saving a project, you can load or save a file with or without network data. When loading, click on the file type dropdown box and select either <project name> (*.sdp), or <project name> skip learn data (*.sdp).

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

To configure Oracle Forms Recognition Verifier, either:

- On the Options menu, select Properties.
- Click the Settings button in the toolbar.

This displays the Oracle Forms Recognition Verifier Properties.

6.1 Settings – General

For general settings such as 508 Compliance and referenced directories and files, select the General tab.

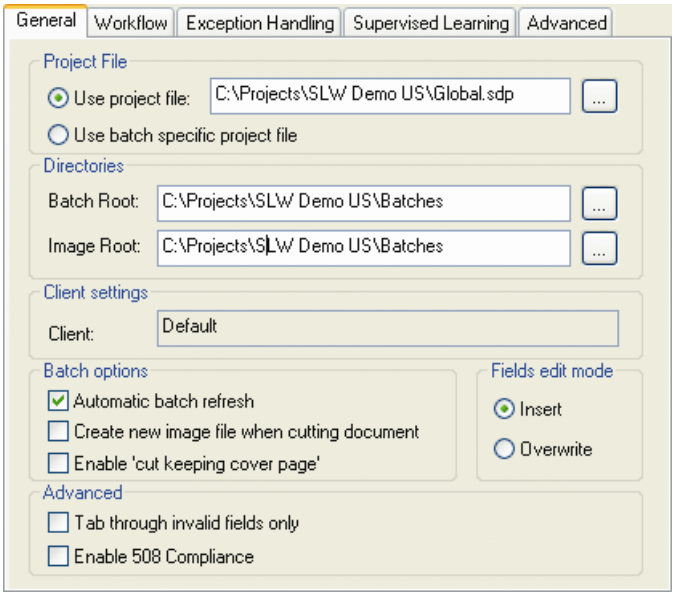


Figure 6-1: General Tab

Project File:

- Use Project File: Used to select the path and file name of the Oracle Forms Recognition application that is used to process the documents and which contains the design of the indexing windows that you will use to verify the extraction. Select this option if you are working with a copy of the project file that resides on the local machine's hard drive, or if you use a project file from the network drive but with a mapping that is different from Oracle Forms Recognition Runtime's drive allocation.

Example:
Let's assume that Oracle Forms Recognition Runtime runs on a computer called ocrsrv and uses the project file SampleCorp.sdp located on ocrsrv in C:\OCR\Projects\SampleCorporation.
Oracle Forms Recognition Verifier runs on a second computer called ocrqa. To access the SampleCorp.sdp project file on ocrsrv, ocrqa needs to be able to access the directory with the project file over the network. This is done by mapping the remote directory on ocrqa. The mapped directory will then obtain a new drive letter.

Let's assume that C:\OCR\ on ocrerv is mapped to drive E:\ on ocrqa. This means that

- C:\OCR\Projects\SampleCorporation\SampleCorp.sdp
- on ocrsrv and
- E:\Projects\SampleCorporation\SampleCorp.sdp
- on ocrqa

are one and the same file.

If you use this option and your documents are processed using more than one project file, you must change the settings every time you change the project file.

- Use batch specific project file: Uses the Oracle Forms Recognition project file that is referenced in the batch itself. Select this option if Verifier and Runtime Server run on the same machine.
  You can also select this option if they run on different machines, but in this case you need to make sure that Verifier and Runtime Server use identical paths and file names for all project files.

Directories:

- Batch Root: This is the directory where the batch control files are.
- Image Root: This is the directory where subdirectories with the scanned images can be found. As a rule, batch root and image root should be the same. In special cases –for security reasons, for example – the image root can be different from the batch root.

Client Settings:

- Client: Currently, only the default setting is available. This option refers to the intention to use client-specific variables. In Oracle Forms Recognition Designer, project administrators can define global variables for different clients. With the default entry, global variables do not vary by client.

Batch Options:

- 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 (using the refresh option on the View menu) when you need up-to-date information.

- Create New Image File When Cutting Document: This option enables Verifier users to create new TIFFs when a Workdoc is split into multiples. The TIFFs correspond to the new Workdocs.
- Enable “cut keeping cover page:” This option enables Verifier users to cut a long document, such as a multi-page fax, into several shorter documents while still retaining the cover page of the original Workdoc as the cover page for each of the newly created shorter documents. If this is checked, the shortcut menu in document browsing view has a new menu entry, Cut page keeping cover page. The new documents must then be re-OCR’d.

Fields Edit Mode:

- Insert: When a document is opened that requires correction or confirmation of extraction results, the cursor is automatically placed in the first invalid field. If you select Insert mode, the cursor is inserted to the left of the field contents.
- Overwrite: When a document is opened that requires correction or confirmation of extraction results, the cursor automatically appears in the first invalid field. If you select Overwrite mode, the entire field content is selected.

Advanced:

- Tab through Invalid Field Only: When the user presses TAB, SHIFT+TAB, CTRL+TAB, or CTRL+SHIFT+TAB to tab through the fields in Document Verification mode, the system tabs through invalid fields only. Likewise, when the user presses TAB inside of a table control, the system tabs through invalid table cells only.
- Enable 508 Compliance: This option activates 508 Compliance settings for your workstation. This option will enable 508 Compliance for all projects you work with from this station. Users at other workstations who do not want to use these features do not have to use them, even if they work on the same projects you do. The following features are available if 508 Compliance is enabled:
  - A blue arrow shows which field has focus.
  - Additional visual indicators besides color highlighting help distinguish between invalid fields, valid fields, and questionable fields. These indicators are present in table fields and form fields. Green check marks show valid fields, red Xs show invalid fields, and orange question marks show questionable fields. Field candidates are highlighted in yellow, but do not have additional validity icons.
  - All menu items have underscored letters available by ALT menu shortcuts.
  - Pop-up menus for workflow state lists and exception handling can be activated by the right-click key on the keyboard. This key is on the right of the standard keyboard, in between the Windows key and the CTRL key.
  - In Show Selected Batch, the right-click keyboard key activates the shortcut menu for Append this document to previous one and Cut pages into a new document.
  - During document verification, pressing CTRL+M or selecting Show Selection Context Menu activates the shortcut menu for the currently selected item.

- In the highlight columns for interactive learning mode, unmapped columns' items are indicated by a blue rectangle without icons and valid / invalid columns' items, via rectangles with an “valid” / “invalid” icon at the left side of every item.
- If input focus is lost for any reason, the user can manually restore it from the Main Menu (by selecting Restore Focus) or by pressing CTRL+N.

6.2 Settings – Workflow

If 508 Compliance is enabled, a blue arrow indicates the current focus for each field when the TAB key is used to move between the fields. The input batch state fields are not highlighted when focus is applied via the TAB key.

Oracle 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 Oracle 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 Oracle 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.

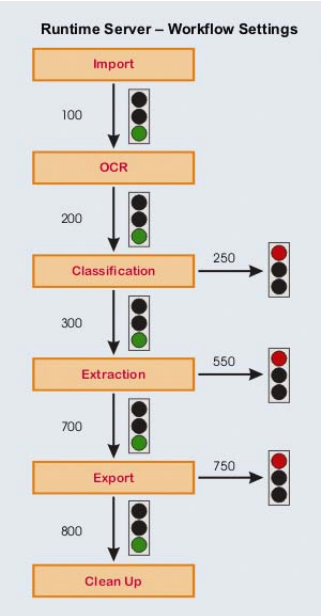


Figure 6-2: Sample Runtime server workflow configuration

In this case, an Oracle Forms Recognition Verifier station conducting both classification and extraction verification should use the settings shown to verify whether this is correct.

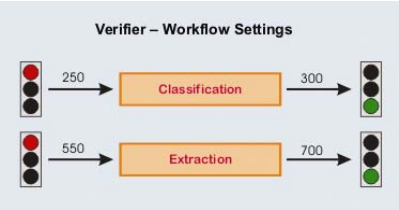


Figure 6-3: An Oracle Forms Recognition Verifier workflow configuration matching the Runtime server settings above

This produces the following combined workflow:

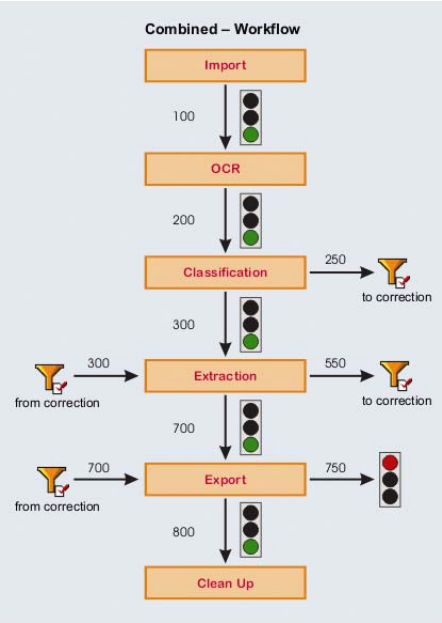


Figure 6-4: Combined workflow resulting from matching Runtime server and Verifier workflow configurations above.

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.

6.2.1. Configuring Tasks to Perform at the Workstation

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

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

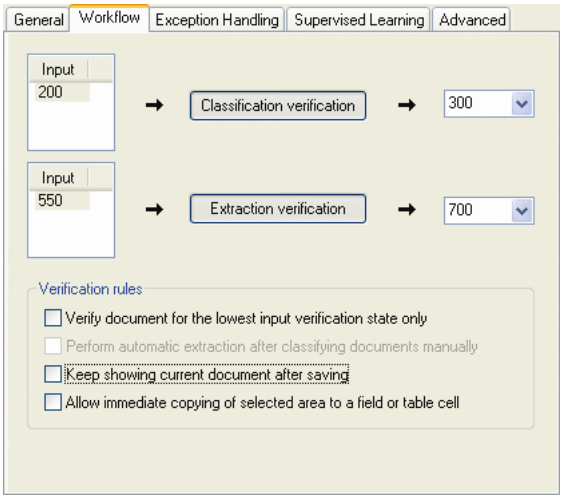


Figure 6-5: Workflow tab

- To configure classification verification at this workstation, click the Classification Verification button. Likewise, you can figure extraction verification by clicking its button. One or both steps can be performed at the workstation.
- After you've 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 delete states this way. Deleting states are available only for those that you have added, but not for preset ones.) To set an output value, select it from the dropdown box to the right of the classification or extraction button.
- Next, configure verification rules:
  - Verify document for the lowest input verification state only: Verifies a document using the lowest input verification state. When this option is selected, the correction of the documents is grouped. After the verification of each input state, the user is asked to release the batch even if there are still documents with a higher input state left to be corrected. This option is valuable when you use several forms to verify extraction fields. If you have several forms defined for default processing (meaning that this option is not selected) all forms will be shown for the document that is corrected. In the example shown below, first form1 and afterwards Form 2 for Document 1 are shown, afterwards Form 1 for Document 2 that has only Form 1. Then all forms for Document 3 will be shown. See ***Figure 6-6***)

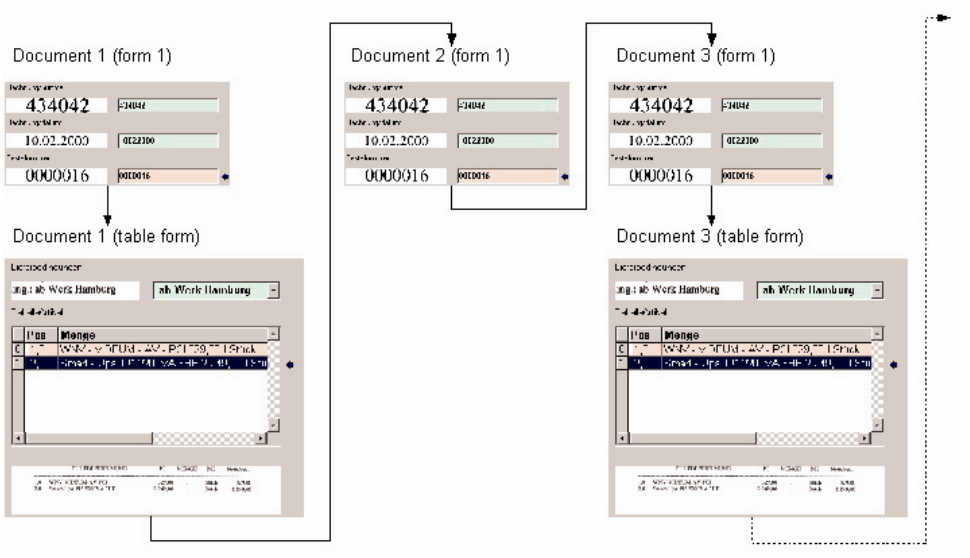


Figure 6-6: Default form sequence for extraction (check box not selected)

If this option is selected, the processing can be changed so that Form 1 for Document 1 is shown first, followed by Form 1 for Document 2 and Form 1 for Document 3 until all documents of the lowest input level are processed. Then the documents for the next input state and other forms are displayed for correction. (***Figure 6-7***) Use the Designer Verifier Design Mode to design verification forms and to define for which input state is shown for which form. (***Figure 6-8***)

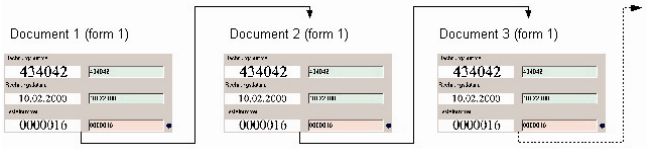


Figure 6-7: Sequence of verification forms depending on input states

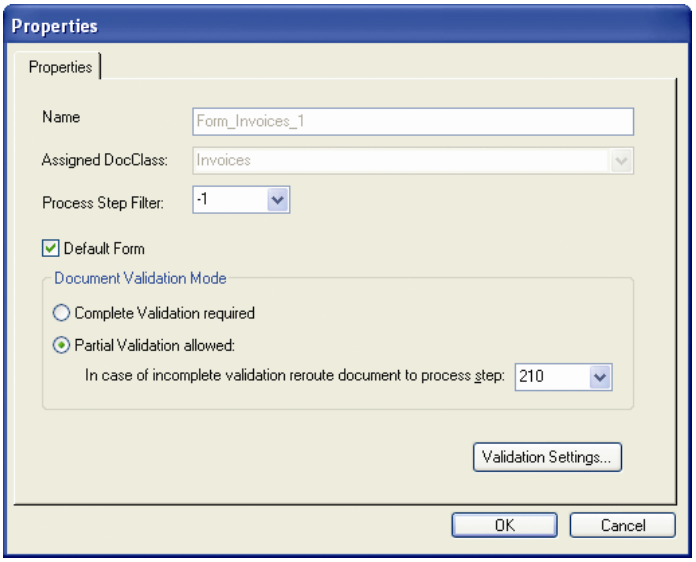


Figure 6-8: Oracle Forms Recognition Designer - Verifier Design Mode Verification Form Settings

In the settings of the first Verifier form of this document class, Partial Validation is selected and the value is set (in this example, the value is 210) that will be the output status for the document when verification with this form is completed. When a document has the input state 210, the

second verification form is used to validate the fields. Therefore, the Process Step Filter was set to 210.

- Perform automatic extraction after classifying documents manually: Forces Oracle 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 Verification workflow step must be entered as an input state for the Extraction Verification input step.
- Keep showing current document after saving: Displays the current document after performing a Save, instead of automatically displaying the next document.
- Allow immediate copying of selected area to a field or table cell: Allows copying of a selected area to a field or table cell when verifying. Speeds up the process by copying single words and candidates to verification elements.

6.3 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 cannot be verified. Without a mechanism to handle unexpected failures, operators would not be able to get the batch with this document out of their task list. This is why Oracle Forms Recognition Verifier incorporates 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 are specialized in collecting exceptions.

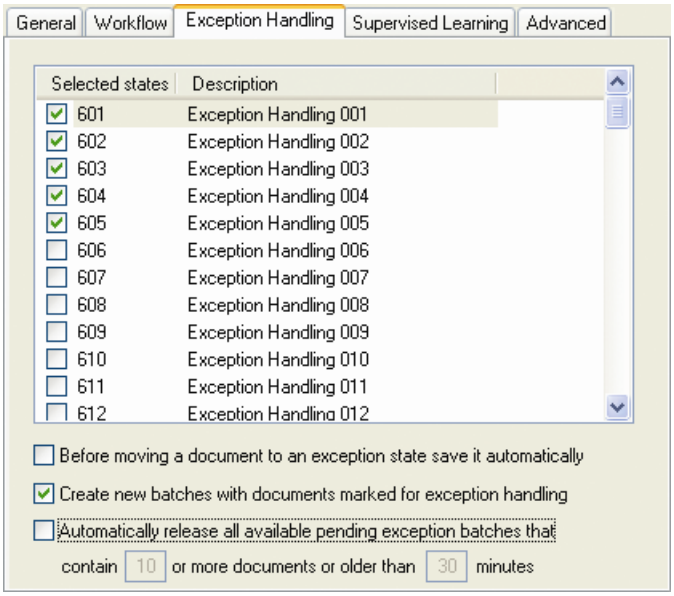


Figure 6-9: Exception Handling tab

6.3.1. 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-value 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.

6.3.2. Editing the Description

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.

- 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. 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 also Section ***7.1.1***)
- 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

6.4 Settings – Supervised Learning

This tab is not available unless Supervised Learning was enabled for the project in Designer.

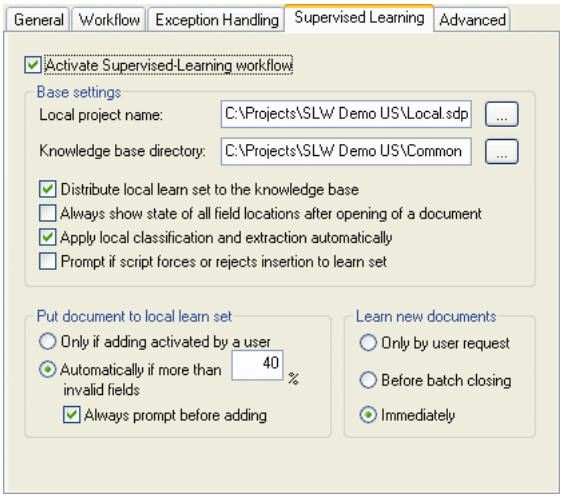


Figure 6-10: Supervised Learning tab

- Activate Supervised Learning Workflow. If this checkbox is not enabled, Supervised Learning will not be available.
- Base Settings (This information should be inherited from the settings your project administrator established in Designer.)
  - Local Project Name: The file and pathname for the local project.
  - Local Learn Set Directory: The folder the new Local Learn Set will be in.
  - Knowledge Base Directory: The file and pathname of the Global Learn Set.
    The Global Learn Set will be updated whenever the Local Learn Set is migrated to it.
  - Distribute Local Learn Set to the knowledge base. Automatically updates the Global Learn Set with valid new local data.
  - Always show state of all field locations after opening of a document. The state of all fields will be displayed when a document is open.
  - Apply local classification and extraction automatically. New classes will be created using the supplier's name. A Learn Set should also be created if you select this setting.
  - Prompt if script forces or rejects insertion to Learn Set
- Under Put document to Local Learn Set:
  - Only if adding activated by a user with Always prompt before Adding selected. A document will be added to a Learn Set when a user requests it. The system will display a dialog box to confirm that the document should be added.
  - Only if adding activated by a user with Always prompt before adding cleared. A document will be added to the Learn Set only if the user requests it. This will be done automatically with no confirmation.
  - Automatically if more than N% invalid fields with Always prompt before adding selected. Documents will automatically be added to the Learn Set if the threshold you set is exceeded. The system will display a dialog box to confirm that the document should be added.
  - Automatically if more than N% invalid fields with Always prompt before adding cleared. Documents will automatically be added to the Learn Set if the threshold you set is exceeded. No confirmation will occur.
- Under Learn new documents:
  - Only by user request: Learning is initiated only when a user asks for it.
  - Before batch closing: Learning is initiated for every batch in the project each time any batch is closed.
  - Immediately: Learning is initiated anytime a document is added to the Learn Set.

6.5 Settings – Advanced

The Advanced tab is used for additional features.

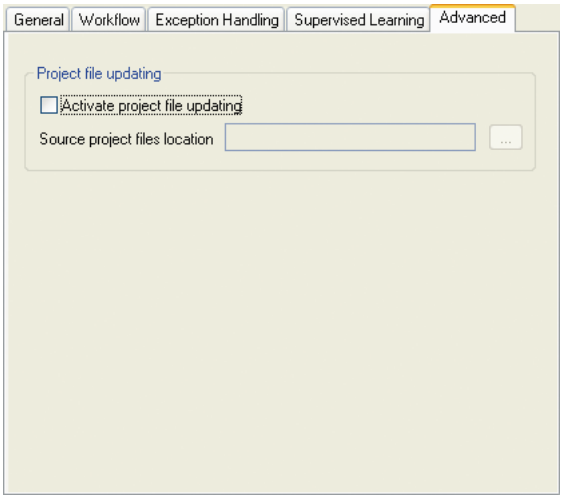


Figure 7-11: Advanced Settings tab

- Under Project file updating:
  - Activate project file updating. Checking this box activates the Project File updating feature.
  - Source project file location. The file and pathname of the source project file.

Chapter 7 Getting Familiar with the User Interface

7.1 The Batch View

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

To select this view, press CTRL+1, or click the following button:

If Oracle Forms Recognition Verifier is not yet configured, the list of batches is empty. For information about configuring Oracle Forms Recognition Verifier, please see **Chapter 6 Configuring Verifier and Advanced Verifier**

The Batch View looks like this:

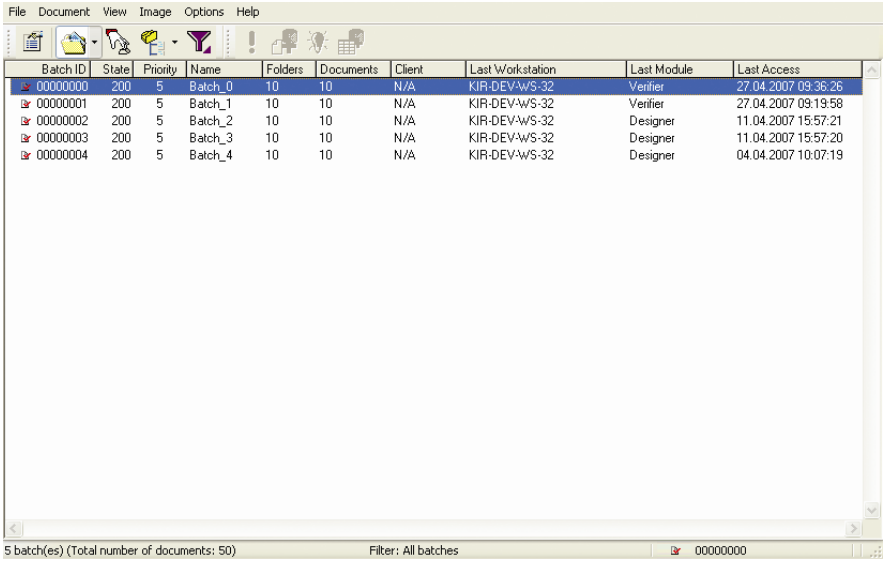


Figure 8-1: Batch View

Menu Bar (See Section ***7.1.1***)

Toolbar (See Section ***7.1.2***)

Table of Batches (See Section ***7.1.3***)

7.1.1. Menu Commands and Keyboard Shortcuts -- Batch View

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
File				
	Open	N/a	N/a	Opens Verifier settings file.
	Save	N/a	N/a	Saves Verifier settings file.
	Save As	N/a	N/a	Saves a Verifier settings file under another name.
	Exit	N/a	N/a	Quits Oracle Forms Recognition Verifier.

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
Document				
	No commands are available in the Batch View context.			
View				
	Toolbar	n/a	n/a	Switch that can be used to show or hide the toolbar.
	Status Bar	n/a	n/a	Switch that can be used to show or hide the status bar.
	Show Batches	n/a	Ctrl 1	Displays the Batch View. (see Section 8.1)
	Verify Selected Batch	n/a	Ctrl 2	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)
	Show Selected Batch	n/a	Ctrl 3	Displays the Document View of the currently selected batch. The Document View provides an overview of the documents within the batch.
	Highlight Mode	n/a	n/a	n/a
	Batch Filter	n/a	n/a	Filters the table of batches as specified below.
		All batches	n/a	All batches from the batch root directory.
		Batches to verify, classification only	n/a	Only batches waiting for classification verification.
		Batches to verify, extraction only	n/a	Only batches waiting for extraction verification.
		Batches to verify	n/a	Only batches waiting for verification.
	Document Filter	n/a	n/a	n/a
	Restore Focus		Ctrl X	Manually restores input focus without using the mouse
	Refresh	n/a	F5	Checks the directory that contains the batches for changes and updates the table of batches accordingly.

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
Image				
	No commands are available in the Batch View context.			
Options				
	Properties	Property Settings		Displays a dialog box where you can configure Oracle Forms Recognition Verifier. (see Chapter 7)
	License	License Settings		
	Change Client	n/a	n/a	Displays a dialog box where you can select the active client.
	Release Exception Batches	n/a	Ctrl E	Manually releases all batches that were created when documents are marked as exceptions. See the settings for the Exception Handling tab, Section 7.3.

	Learn Set Manager	n/a	n/a	Starts the Learn Set Manager (if licensing and permissions are configured.)

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

7.1.2. Toolbar Buttons

The toolbar provides quick access to some frequently used commands:

Button	Description
	Displays a property sheet where you can configure Oracle Forms Recognition Verifier. (See ***Chapter 6***)
	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 ***7.4 The Verification View - Indexing Window***)
	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 ***7.2 The Document View***)
	Starts Learn Set Manager. (***Chapter 10 Working with the Learn Set Manager***)

Table 8-4: Batch view controls

7.1.3. 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 that has 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 of another application. Therefore it can not 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. (Section ***6.2 Settings – Workflow***)

in the batch list can be sorted for 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.
- Client:
  The owner of the Oracle Forms Recognition license. Contains N/A; not in use yet.
- Last user:
- Computer name of the operator who has previously processed the batch.
- Last module:
  Name of the application that most recently processed the batch.
- Locked:
  This column is empty or contains the letter L. The latter indicates that the batch is locked. You cannot work with this batch because it is being used by someone else or by an instance of Oracle Forms Recognition Runtime. If a batch is locked, the batch information is displayed in red.

7.1.3.1. 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, then click the Batch column label; it will revert to the last batch on the list.

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

In the table of batches, select one batch and then move through the list using the following keyboard commands:

To move to the first document, press HOME.

- To move to the next document, press the ARROW DOWN key.
- To move to the previous document, press the ARROW UP key.
- To move to the last document, press END.
- To move one page down, press PAGE DOWN.
- To move one page up, press PAGE UP.

You can leave the Batch View and switch to another views using the following keyboard commands:

- To verify the selected batch, press CTRL+2.
- To view the selected batch, press CTRL+3.

7.2 The Document View

The Document View can be used to investigate the documents in a selected batch. Only the first page of a document is displayed.

To select this view, press CTRL+3, or click the “Show selected batch” button:

The Document View looks like this:

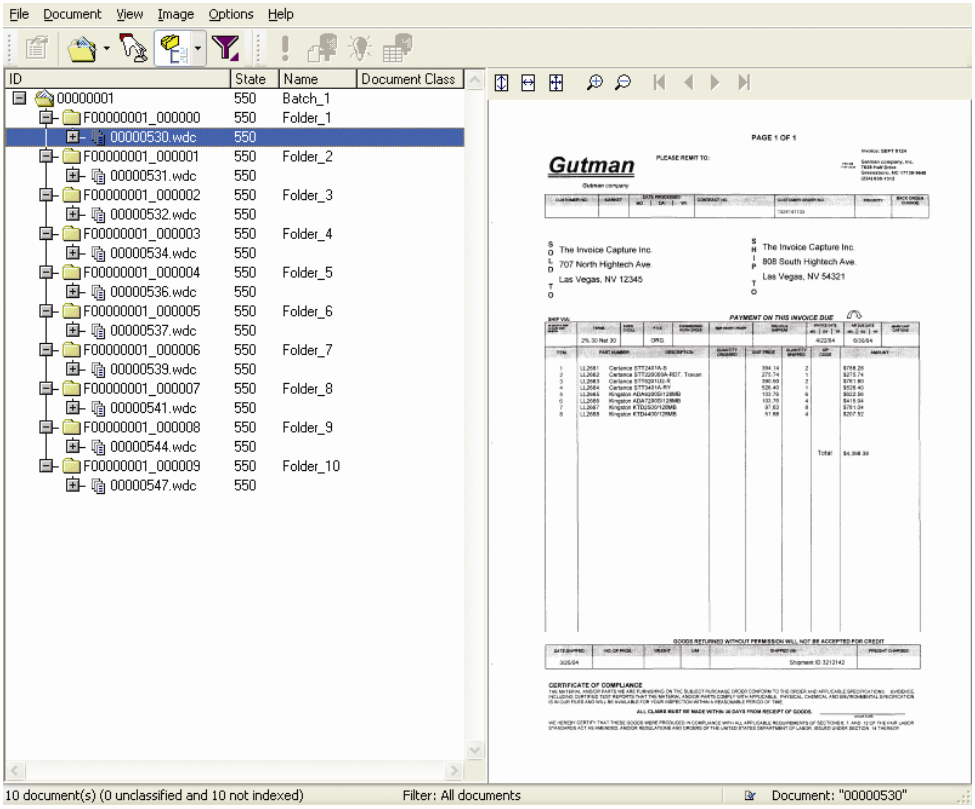


Table 8-6: The Document View

Menu Bar (See Section ***7.2.1***)
Main Toolbar (See Section ***7.2.2***)
Batch Structure (See Section ***7.2.3***)
Viewer Toolbar (See Section ***7.2.4***)
Document (See Section ***7.2.5***)

7.2.1. Menu Commands and Keyboard Shortcuts

The following commands can be accessed via the menu bar:

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
File				
	Exit	n/a	n/a	Quits Oracle Forms Recognition Verifier.
Document				
	First Document	n/a	Ctrl Alt Home	Moves to the first document.
	Preceding Document	n/a	Ctrl Alt Page up	Moves to the previous document.
	Next Document	n/a	Ctrl Alt Page down	Moves to the next document.
	Last Document	n/a	Ctrl Alt End	Moves to the last document.

Table 8-7: Document view menu commands (File & Document)

Menu	Submenu/ Command	Command	Key board	Description
View				
	Toolbar	n/a	n/a	Switch that can be used to show or hide the toolbar.
	Status Bar	n/a	n/a	Switch that can be used to show or hide the status bar.
	Show Batches	n/a	Ctrl 1	Displays the Batch View. (see Section ***7.1***) The check mark shows that this view is selected.
	Verify Selected Batch	n/a	Ctrl 2	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 ***7.3***) or in Indexing Mode. (see Section ***7.4***)
	Show Selected Batch	n/a	Ctrl 3	Displays the selected batch in Batch View.
	Highlight Mode	n/a	n/a	n/a
	Batch Filter	n/a	n/a	n/a
	Document Filter			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	n/a	Filters the structure of all of the documents.
		Documents to classify or index	n/a	Filters the structure of all documents that will be classified or indexed.
		Documents to classify	n/a	Filters the structure of all documents that will be classified.
		Documents to index	n/a	Filters the structure of all documents that will be indexed.
	Restore Focus			Manually restores input focus without using the mouse.
	Refresh	n/a	n/a	Checks the directory that contains the documents for changes and updates the table of documents accordingly.

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

Menu	Submenu/ Command	Cmd	Keyboard Shortcut	Description
Image				
				Navigates the document area as specified below
	Zoom In	n/a	Ctrl +	Zooms in.
	Zoom Out	n/a	Ctrl -	Zooms out.
	Move Image to Left	n/a	Ctrl ←	Scrolls the document area to the left.
	Move Image to Right	n/a	Ctrl →	Scrolls the document area to the right.
	Move Image Upwards	n/a	Ctrl ↑	Scrolls the document area upward.
	Move Image Downward	n/a	Ctrl ↓	Scrolls the document area downward.
	Rotate	n/a	Ctrl R	Rotates image 90 degrees to the right.
	First page in document	n/a	Ctrl Home	Moves to the first page of the document.
	Previous page in document	n/a	Ctrl Page down	Moves to the preceding page of the document.
	Next page in document	n/a	Ctrl Page up	Moves to the following page of the document.
	Last page in document	n/a	Ctrl End	Moves to the final page of the document.
	Brightness and Contrast	n/a	n/a	Opens a dialog box with sliders to control the image's brightness and contrast.
	Show Selection Context Menu		Ctrl M	Invokes the shortcut menu applicable for the selection
	Fit to Height		n/a	Fits the document to the height of the viewer.
	Fit to Width		n/a	Fits to document to the width of the viewer
	Best fit		n/a	Forces the document to fit into the viewer so all of it is displayed in the viewer.
	Keep focus on field	n/a	n/a	Keeps the focus on the same field for each document you view in the batch.
	Keep zoom	n/a	n/a	Keeps the established zoom settings on each document you view in the batch.
	Print	n/a	Ctrl P	Prints the image see section ***7.5***

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

Menu	Submenu/ Command	Cmd	Keyboard Shortcut	Description
Options				
	Undo			Can be used in Document View
	Settings	n/a	n/a	Displays a property sheet where you can configure Oracle Forms Recognition Verifier. (See ***Chapter 6***)
	License	n/a	n/a	Modify license connections: Use specified path or use License directory.
	Reclassify manually	n/a	n/a	Option available in Advanced Verifier.
	Show last verified document	n/a	n/a	Option available in Advanced Verifier.
	Show last value for selected field	n/a	n/a	Option available in Advanced Verifier.
	Show location states for all fields	n/a	n/a	Option available in Advanced Verifier.
	Apply local extraction	n/a	Ctrl L	Option available in Advanced Verifier.
	Add document to Learn Set	n/a	Ctrl A	Option available in Advanced Verifier.
	Learn Documents	n/a	n/a	Option available in Advanced Verifier.
	Correct Tables	n/a	n/a	Option available in both Verifier and Advanced Verifier, but enhanced functionality only available in Advanced Verifier.
	Switch table highlighting	n/a	n/a	Option available in Advanced Verifier.
	Learn Set Manager	n/a	n/a	Starts the Learn Set Manager.
	Change Password	n/a	n/a	Allows you to change your password.

Table 8-10: Document view menu commands (Options)

7.2.2. Main Toolbar Buttons

The toolbar provides quick access to some frequently used commands:

Button	Description
	Displays a property sheet where you can configure Oracle Forms Recognition Verifier. (See ***Chapter 6***)
	Displays the Batch View is displayed
	Starts the verification of the currently selected batch. Depending on the batch state, the batch is either displayed in the classification window (Section ***7.3 The Verification View – Classification Window***) or in the indexing window. (Section ***7.4 The Verification View - Indexing Window***)
	Displays the available filters for the batch structure. You can select from among the following options: • All documents • Batches to verify / docs to verify • Batches to classify / all docs • Batches to classify / docs to verify • Batches to extract / all docs • Batches to extract / docs to verify • Batches to classify or extract / all docs • Batches to classify or extract / docs to verify
	Starts Learn Set Manager (Available only if configured in Designer and properly licensed.) ***Chapter 10 Working with the Learn Set Manager***

Table 8-11: Document view controls

7.2.3. Batch Structure Area

In the batch structure, a hierarchical representation of the batch contents is displayed. 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.

7.2.3.1. 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, then press the batch column label. It will revert to the last batch on 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 use the following keyboard commands to navigate:

- To move to the first document, press CTRL+ALT+HOME.
- To move to the next document, press CTRL+ALT+PAGE DOWN.
- To move to the previous document, press CTRL+ALT+PAGE UP.
- To move to the last document, press CTRL+ALT+END.

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

7.2.3.2. Splitting and Merging Documents

In the document list, you can split multipage documents into separate documents, with the exception that the first page of a document cannot be split. You can also merge consecutive documents into one with multiple pages. (***Figure 7-2*** and ***Figure 7-3***)

Splitting Multipage documents

To split a multipage document, do the following:

1) Select View, Show Selected Batch, All Documents from the Main Menu, or press the Show Selected Batch, All Documents button.
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 underline and one 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.
5) Or select Cut pages into new document keeping cover page. Like the option above, this option splits a single document into several smaller documents and corresponding TIFFs. However, it also includes the cover page of the original document as the cover page for the newly created documents (See also ***cut keeping cover page***.)

Figure 7-2: Splitting pages in a new document

Appending Two Documents

To append two documents:

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

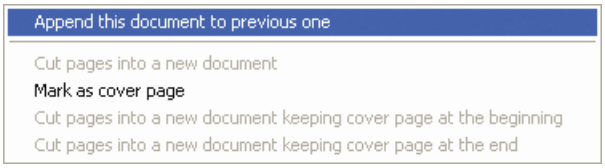


Figure 7-3: Appending pages to a document

7.2.4. 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. Alternatively, press Ctrl +
	Zooms out. Alternatively, press Ctrl -

Table 8-12: Viewer Toolbar Controls

7.2.5. Document Area

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

7.3 The Verification View – Classification Window

To display the Verification View, select a batch from the list that requires verification. Then press CTRL+2, or click the “verify selected batch” button.

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:

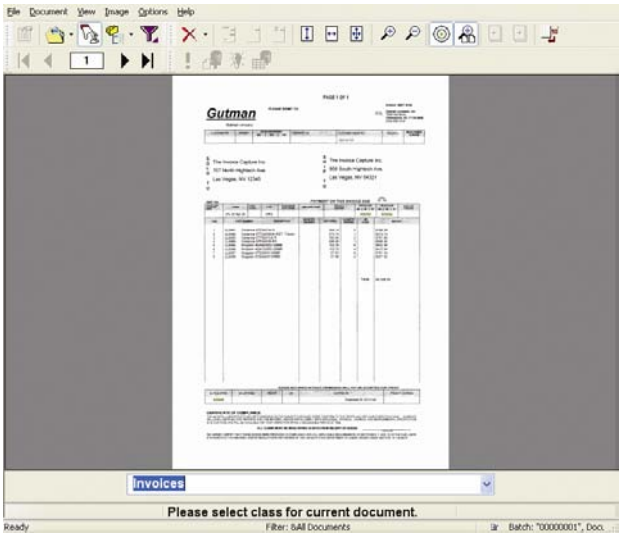


Figure 8-4: The Classification window

Menu Bar (See Section ***7.3.1***)

Toolbar (See Section ***7.3.2***)

Document (See Section ***7.3.3***)

Class Selection (See Section ***7.3.4***)

7.3.1. Menu Commands and Keyboard Shortcuts

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
File				
	Exit	n/a	n/a	Quits Oracle Forms Recognition Verifier.
Document				
	First Document	n/a	Ctrl Alt Home	Moves to the first document.
	Preceding Document	n/a	Ctrl Alt Page up	Moves to the previous document.
	Next Document	n/a	Ctrl Alt Page down	Moves to the next document.
	Last Document	n/a	Ctrl Alt End	Moves to the last document.

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

Menu	Submenu/ Command	Cmd	Keyboard SHortcut	Description
View				
	Toolbar	n/a	n/a	Switch that can be used to show or hide the toolbar.
	Status Bar	n/a	n/a	Switch that can be used to show or hide the status bar.
	Show Batches	n/a	Ctrl 1	Displays the Batch View (see Section ***8.1***)
	Verify Selected Batch	n/a	Ctrl 2	Switches from Batch Mode to Verify Mode.
	Show Selected Batch	n/a	Ctrl 3	Displays the Document View of the currently selected batch (see Section ***8.2***) The Document View provides an overview of the documents within the batch.
	Highlight Mode	n/a	n/a	n/a
	Batch Filter	n/a	n/a	n/a
	Document Filter			n/a
	Restore Focus			Restores input focus
	Refresh	n/a	n/a	Switch that can be used to show or hide the toolbar.

Table 8-14: Verification View - Classification Mode menu commands (View)

Menu	Submenu/ Command	Cmd	Keyboard Shortcut	Description
Image (Navigates the document area specified below)				
	Zoom In	n/a	Ctrl +	Zooms in.
	Zoom Out	n/a	Ctrl -	Zooms out.
	Move Image to Left	n/a	Ctrl ←	Scrolls the document area to the left.
	Move Image to Right	n/a	Ctrl →	Scrolls the document area to the right.
	Move Image Upwards	n/a	Ctrl ↑	Scrolls the document area upward.
	Move Image Downward	n/a	Ctrl ↓	Scrolls the document area downward.
	Rotate	n/a	Ctrl R	Rotates the document clockwise.
	First page in document	n/a	Ctrl Home	Moves to the first document page of multipage document.
	Previous page in document	n/a	Ctrl Page down	Moves to the previous document page of multipage document.
	Next page in document	n/a	Ctrl Page up	Moves to the next document page of multipage document.
	Last page in document	n/a	Ctrl End	Moves to the last document page of multipage document.
	Brightness and Contrast	n/a	n/a	Adjusts image brightness and contrast. Use slider to view. Brightness adjusts line darkness; Contrast adjusts the difference between text and page.
	Print	n/a	Ctrl P	Prints the image see section ***8.5***

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
Options				
	Undo	n/a	n/a	Retracts the last action.
	License	n/a	n/a	A dialog box is shown to select the path for the license file.
	Show Last Verified Document	n/a	n/a	Moves from the current document to the last document that has been verified, and switches from Verify Mode to Batch Mode.
	Change Password	n/a	n/a	Allows you to change your password.

Table 8-16: Verification View - Classification Mode menu commands (Object)

7.3.2. Toolbar Buttons

The toolbar provides quick access to some frequently used commands:

Button	Description
	If you click on this button, the Batch View is displayed (See Section ***7.1 The Batch View***)The Batch View
	Click this button to verify the selected batch. (See Section ***7.4 The Verification View - Indexing Window***)
	If you click this button, the Document View is displayed (See Section ***7.2 The Document View***)
	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.
	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.
	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.
	Rotates the current document 90 degrees clockwise.
	This button is only enabled if the current document has more than one page. If you click on it, the previous page is displayed.
	This button is only enabled if the current document has more than one page. If you click on it, the next page is displayed.

Table 8-17: Verification View - Classification Mode Controls

7.3.3. Document Area

This area shows the current document.

7.3.4. Class Selection List

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

To set or change a classification result, be 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 and 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.

7.4 The Verification View - Indexing Window

To display the Verification View, select a batch from the list that requires verification. Then press CTRL+2, or click the button shown at left.

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:

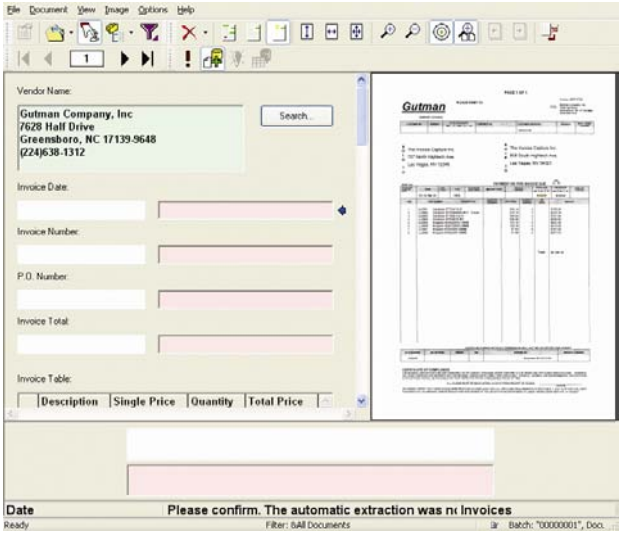


Figure 7-5: Indexing Mode

Menu Bar (see Section ***7.4.1***)	Document (see Section ***7.4.5***)
Toolbar (see Section ***7.4.2***)	Current Area (see Section ***y***)
Field Area (see Section ***7.4.4***)	User Info Area (see Section ***7.4.7***)

Your indexing window may look quite different than the one shown in ***Figure 8-5: 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 changed 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.

7.4.1. Menu Commands and Keyboard Shortcuts

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
File				
	Exit	n/a	n/a	Quits Oracle Forms Recognition Verifier.
Document (All commands switch to Browsing Mode)				
	First Document	n/a	Ctrl Alt Home	Moves to the first document.
	Preceding Document	n/a	Ctrl Alt Page up	Moves to the previous document.
	Next Document	n/a	Ctrl Alt Page down	Moves to the next document.
	Last Document	n/a	Ctrl Alt End	Moves to the last document.

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
View				
	Toolbar	n/a	n/a	Switch that can be used to show or hide the toolbar.
	Status Bar	n/a	n/a	Switch that can be used to show or hide the status bar.
	Show Batches	n/a	Ctrl 1	Displays the Batch View. (See Section 8.1)
	Verify Selected Batch	n/a	Ctrl 2	Switches from Batch Mode to Verify Mode.
	Show Selected Batch	n/a	Ctrl 3	Displays the Document View of the currently selected batch. (See Section 8.2)
	Highlight Mode	n/a	n/a	Marks certain areas on the document in color. Green - result is valid Red - result is invalid. Yellow - considered candidates (during the extraction as a candidate, but were not used to fill a field).
		Highlight Fields	n/a	Marks all areas that have been used to fill fields.
		Selected Fields	n/a	Marks only the area used to fill the current field.
		Highlight Candidates	n/a	Marks the area used to fill the current field and all candidates for the current field.
	Batch Filter	n/a	n/a	Displays documents from all batches.
		All batches	n/a	
		Classifica-tion only	n/a	
		Indexing only	n/a	
		Batches to verify	n/a	
	Document Filter			
		All documents	n/a	
		Documents to classify or index	n/a	
		Documents to classify	n/a	
		Documents to index	n/a	
	Restore Focus		Ctrl N	
	Refresh	n/a	F5	

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

Menu	Submenu/ Command	Cmd	Keyboard Shortcut	Description
Image				
	Zoom In	n/a	Ctrl +	Zooms in.
	Zoom Out	n/a	Ctrl -	Zooms out.
	Move Image to Left	n/a	Ctrl ←	Scrolls the document area to the left.
	Move Image to Right	n/a	Ctrl →	Scrolls the document area to the right.
	Move Image Upwards	n/a	Ctrl ↑	Scrolls the document area upward.
	Move Image Downward	n/a	Ctrl ↓	Scrolls the document area downward.
	Rotate	n/a	Ctrl R	Rotates the document clockwise.
	First page in document	n/a	Ctrl Home	Moves to the first document page of multipage document.
	Previous page in document	n/a	Ctrl Page up	Moves to the previous document page of multipage document.
	Next page in document	n/a	Ctrl Page down	Moves to the next document page of multipage document.
	Last page in document	n/a	Ctrl End	Moves to the last document page of multipage document.
	Brightness and Contrast	n/a	n/a	
	Show Selection Context Menu	n/a	Ctrl M	
	Fit to Height	n/a	Ctrl H	
	Fit to Width	n/a	Ctrl W	
	Best Fit	n/a	Ctrl F	
	Keep Focus on Field	n/a	n/a	
	Keep Zoom	n/a	n/a	
	Increase Image Area	n/a	Ctrl J	
	Decrease Image Area	n/a	Ctrl K	
	Print	n/a	Ctrl P	Prints the image see section ***8.5***

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

Menu	Submenu/ Command	Cmd	Keyboard Shortcut	Description
Options				
	Undo			Retracts the last action.
	Settings	n/a	n/a	Open Settings dialog.
	License	n/a	n/a	A dialog box is shown, to select the path for the license file.
	Change Client...	n/a	n/a	
	Reclassify manually	n/a	n/a	Opens the classification window (see Section ***7.3***) for the current document.
	Show last verified document	n/a	n/a	Moves to the last document that has been verified, and switches from Verify Mode to Batch Mode.
	Get last value for selected field	n/a	n/a	Takes the field value that has been validated previously and copies it into the currently selected field.
	Move Document to Exception State	Not Applicable	F9	
	Release Exception Batches	Not Applicable	Ctrl E	
	Apply local extraction	n/a	Ctrl L	Option available in Advanced Verifier.
	Add document to Learn Set	n/a	Ctrl A	Option available in Advanced Verifier.
	Learn Documents	n/a	n/a	Option available in Advanced Verifier.
	Correct Tables	n/a	n/a	Option available in both Verifier and Advanced Verifier, but enhanced functionality only available in Advanced Verifier.
	Switch table highlighting	n/a	n/a	Option available in Advanced Verifier.
	Learn Set Manager	n/a	n/a	Starts the Learn Set Manager.
	Change Password	n/a	n/a	Allows you to change your password.

Table 8-21: Verification View - Indexing Mode menu commands (Options)

7.4.2. Toolbar Buttons

The toolbar provides quick access to some frequently used commands:

Button	Description
	Displays the Batch View is displayed. (See Section ***7.1 The Batch View***)
	Displays the Verification window (See Section ***7.3 The Verification View – Classification Window***) for the current document. You can verify the selected document manually.
	Displays the Document View. (See Section ***7.2 The Document View***)
	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.
	If you click the arrow next to this button, a list of exceptions is displayed. You can use these exceptions if you cannot correct a document at all, for example because the required data is illegible. Please ask your supervisor which exceptions should be used in which case.
	Marks all areas on the current document that have been used to fill the fields. If the result is valid, the area is highlighted in green. If the result is invalid, the area is highlighted in red.
	Marks only the area on the current document that was used to fill the field that is currently selected in the field area. If the extraction result is valid, the area is highlighted in green. If the extraction result is invalid, the area is highlighted in red.
	Marks the area that was used to fill the field that is currently selected in the field area. This area either appears in green or in red. In addition, all other areas that were taken into account to fill this field, are highlighted in yellow.
	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, the application always displays the document area that is associated with the currently selected field.
	If this button appears pressed, all documents in a batch are displayed with the same magnification.
	Rotates the current document 90 degrees clockwise.
	This button is only enabled if the current document has more than one page. If you click it, the previous page is displayed.
	This button is only enabled if the current documents has more than one page. If you click it, the next page is displayed.

Table 8-22: Verification View - Indexing Mode Controls

7.4.3. Support of Mouse Wheel

7.4.3.1. Description

Oracle Forms Recognition Verifier application supports the mouse wheel usage when validating documents in document verification mode. In this connection, for the user's convenience, the mouse wheel rolling has the following effect depending on where the mouse cursor is or where the keyboard focus is:

Case	Wheel rolling effect
Input focus is in a multi-line header field	Scrolls between lines of the header field
Input focus is in a single line header field or at the first line / row of any field (scrolling up only) or at the last line / row (scrolling down only)	Scrolls the entire verification form
Input focus is in a table field	Scrolls between table rows or between multiple lines of the currently selected table cell (when multi-line)
Mouse pointer is in the document viewer area	Scrolls the currently viewed page image up and down

7.4.3.2. Usage

The present feature is useful in terms of convenience and performance of Verifier application's navigation using mouse. As a result it speeds up performance of documents' verification in Verifier and saves certain amount of time per Verifier user.

7.4.4. Field Area

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

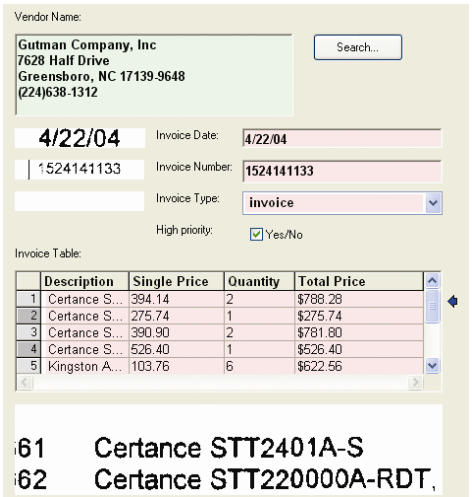


Figure 8-6: Field area example

Form fields

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

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

Labels

Captions that help users to identify form fields and – if desired – viewers and tables.

Viewer

Snippets of document areas, normally those that were extracted to fill fields or tables.

Buttons

Buttons that fire actions for a new script event.

Tables

Relevant when table extraction is configured. The 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.

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

Icon	Description
	Indicates the currently selected field.
	Indicates a smart index field that can be filled by a database lookup. This icon is visible only when a smart index field is selected.
	Indicates a smart index field that can be used to start a lookup. This icon is visible regardless of whether a smart index field is selected.
	Indicates a valid extracted field.
	Indicates a field that needs to be validated because it was extracted with low confidence.

Table 8-23: Field area icons

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. Similarly, pressing SHIFT+TAB gets you to the preceding 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 then focus moves to the next field that requires correction. This field may also be within another document.

7.4.5. Document Area

The document area shows the currently selected document or page along with highlights (See ***Figure 7-5: Indexing Mode***).

- Red areas indicate an invalid result.
- Green areas indicate a valid result.
- 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.

In practice, red, green, and yellow areas never appear in the same document at the same time.

Table fields provide additional highlighting options, as shown in ***Figure 7-7***. Many of these options will seem familiar if you've worked with a spreadsheet program such as Microsoft Excel or Lotus 1-2-3.

	Description	Single Price	Quantity	Total F
1	Certance STT2401A - S	394.14	2	$788.2
2	Certance STT220000A - R...	275.74	1	$275.7
3	**Certance STT6201U2 - R**	390.90	2	$781.8
4	Certance STT3401A - RY	526.40	1	$526.4
5	Kingston ADA6200S / 128...	103.76	6	$622.5

Figure 7-7: Highlighting for table fields

- Clicking the square in the upper-left corner of a table field highlights the entire document table. In ***Figure 7-7***, this square is above the 1 and to the left of Description.
- Clicking a column label of a table field highlights the document column. In ***Figure 7-7***, the column labels are Description, Single Price, Quantity, and Total Price.
- Clicking a row label of a table field highlights the document row. In ***Figure 7-7***, the row tables are 1, 2, and …
- Clicking a cell of a table field highlights the document cell.
- Valid areas are green; invalid are red. These areas may also contain validity icons: green check marks for valid fields' red Xs for invalid fields.

Only one table will display per verification form, even if you are able to define multiple tables. However, you can display different tables on different forms.

If you only need to verify certain columns in a table, you can make the other columns invisible. All invisible columns must be valid for the entire table can be valid.

7.4.6. Current Input Area

The current input area shows enlarged information for the currently select field and provides a spacious edit box.

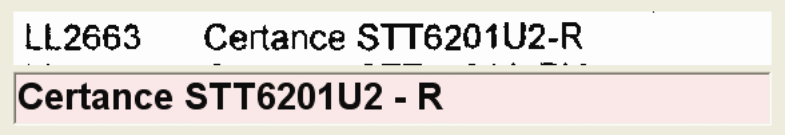


Figure 7-8: The input area

- A snippet that shows an enlargement of the document area that was used to fill the field.
- The extracted data. Color coding is used in the same way as in the field area. (See Section **7.4.4**) You can edit the data here.

7.4.7. User Info Area

The user info area (See ***Figure 7-5: Indexing Mode***) consists of three fields that display the following information (from left to right):

- The name of the currently selected field.
- If the current field is invalid, the reason is displayed.
  If the current field is valid, the field is normally empty.
- The classification result of the current document.

7.5 Printing

Oracle Forms Recognition Verifier allows the currently opened image document to be printed. The user can invoke this method by selecting “Print…” menu item of the “Image” main menu option or by using “Ctrl + P” shortcut.

The print dialog looks like:

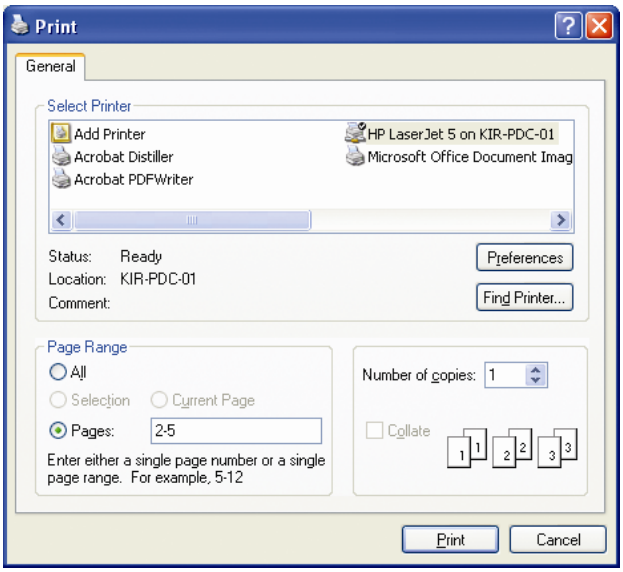


Figure 7-9: Verifier Print dialog

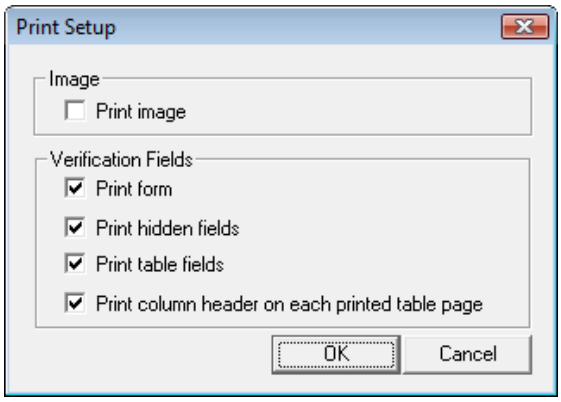
Verifier allows all the pages of the currently selected document to be printed by selecting the “All” radio button, or a desired page or page range to be specified in the “Pages” field of the Print dialog.

Note that the function is available in all modes of Verifier (classification verification, extraction verification and document browsing mode) with the exception of batch browsing mode.

7.6 Printing of verified Data Content

7.6.1. Description

The amount of a printed form’s data can be configured via “Page Setup” dialog available via “File \ Page Setup…“ menu item of Oracle Forms Recognition Verifier application:



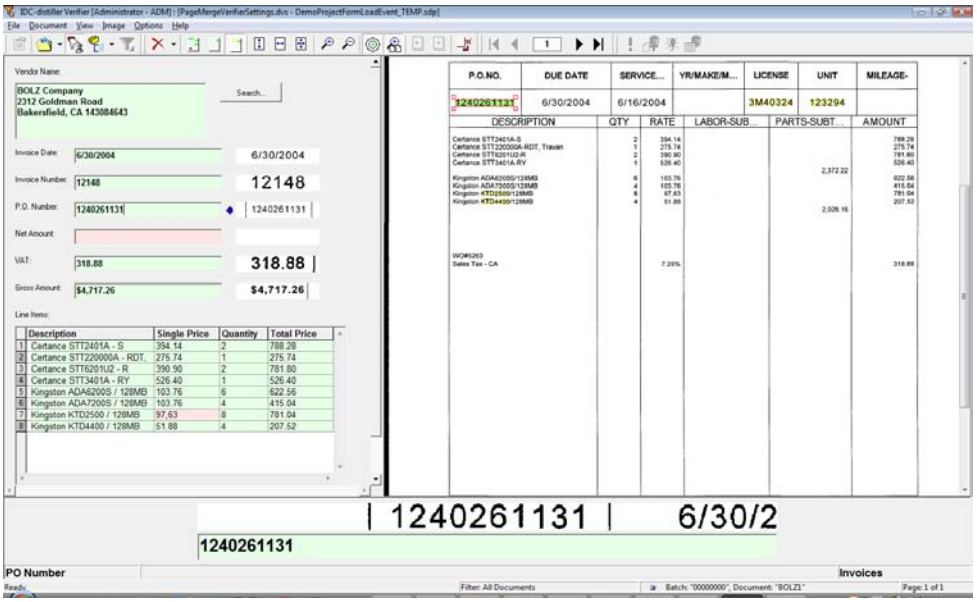
Here is the description of the available options:

Option Name	Description
Print image	When selected, also prints pages of the document file (image).
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, 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, 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	Has effect (enabled) only if “Print table fields” option is turned on. When turned on (default), Verifier prints table header on each page (this option is useful for printing of long tables).

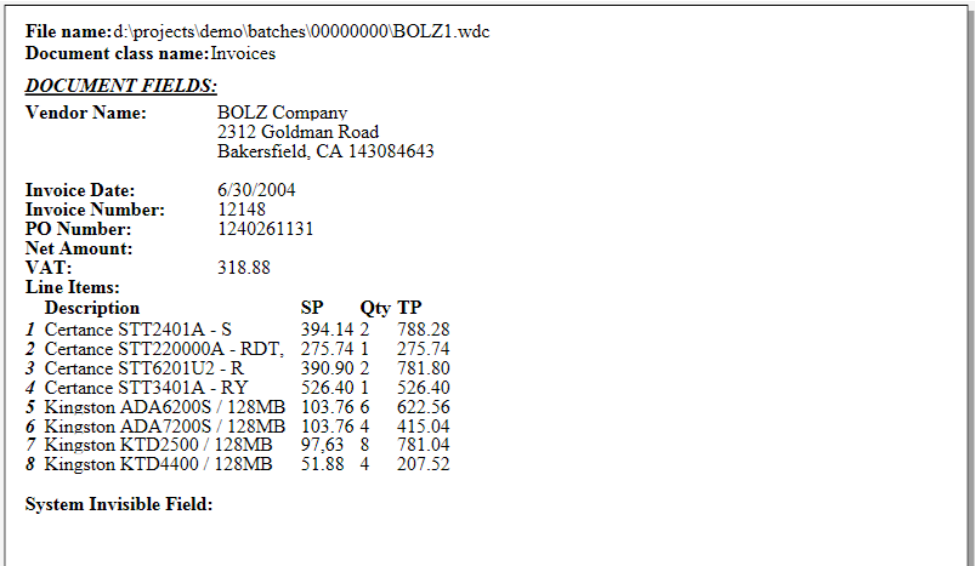
Save your Verifier Settings, if you would like your current “Page Setup” preferences to be restored next time you start Oracle Forms Recognition Verifier application.

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

For example, for the following document and the form:



The Verifier application is going to print out the following information (“Page Setup” settings from the screenshot above were used):



File name: d:\projects\demo\batches\00000000\BOLZ1.wdc
Document class name: Invoices

DOCUMENT FIELDS:

Vendor Name: BOLZ Company
2312 Goldman Road
Bakersfield, CA 143084643

Invoice Date: 6/30/2004
Invoice Number: 12148
PO Number: 1240261131
Net Amount:
VAT: 318.88
Line Items:

	Description	SP	Qty	TP
1	Certance STT2401A - S	394.14	2	788.28
2	Certance STT220000A - RDT,	275.74	1	275.74
3	Certance STT6201U2 - R	390.90	2	781.80
4	Certance STT3401A - RY	526.40	1	526.40
5	Kingston ADA6200S / 128MB	103.76	6	622.56
6	Kingston ADA7200S / 128MB	103.76	4	415.04
7	Kingston KTD2500 / 128MB	97.63	8	781.04
8	Kingston KTD4400 / 128MB	51.88	4	207.52

System Invisible Field:

In this connection, the order the fields are going to be printed is defined by the custom fields’ order configured via Form Design mode of Oracle Forms Recognition Designer application.

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

7.6.2. Usage

The form’s printing function can be used in case the user would like to review a difficult validated document using its paper copy.

Chapter 8 Working with Verifier and Advanced Verifier

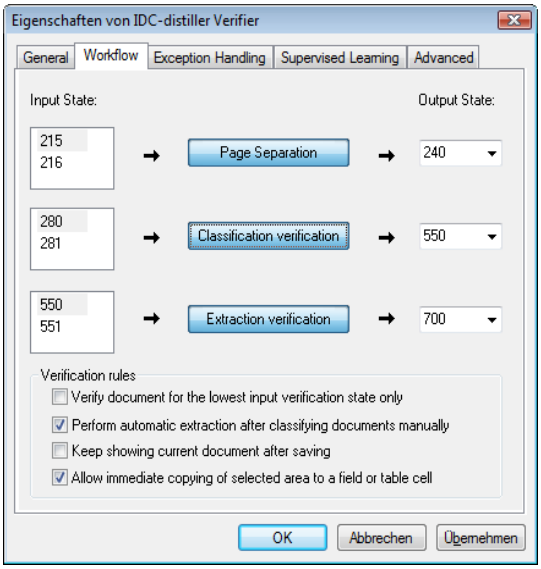
This chapter provides step-by-step instructions for the main tasks that can be carried out with Oracle Forms Recognition Verifier. We recommend that you read ***Chapter 7*** before you start working with Chapter 9.

8.1 Page Separation Workflow in Verifier

8.1.1. Verifier Settings for “Page Separation”

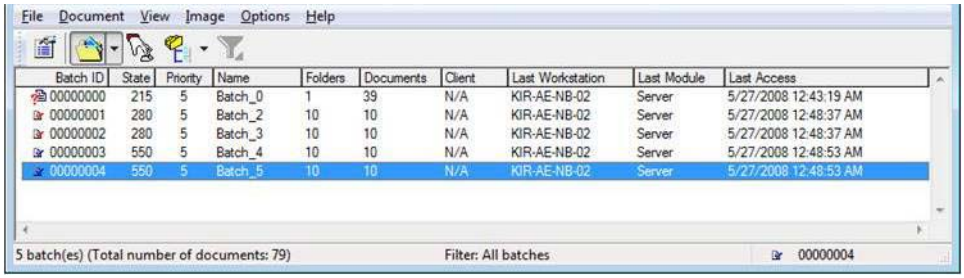
The verifier settings for “Page Separation” are similar to a simple verifier project in that the project and batch root paths must be specified in the general tab of the verifier settings.

The only entry pertaining to “Page Separation” is in the “Workflow” tab. Define herein the input state for “Page Separation”. The final step is to activate this workflow step by clicking on the “Page Separation” button.

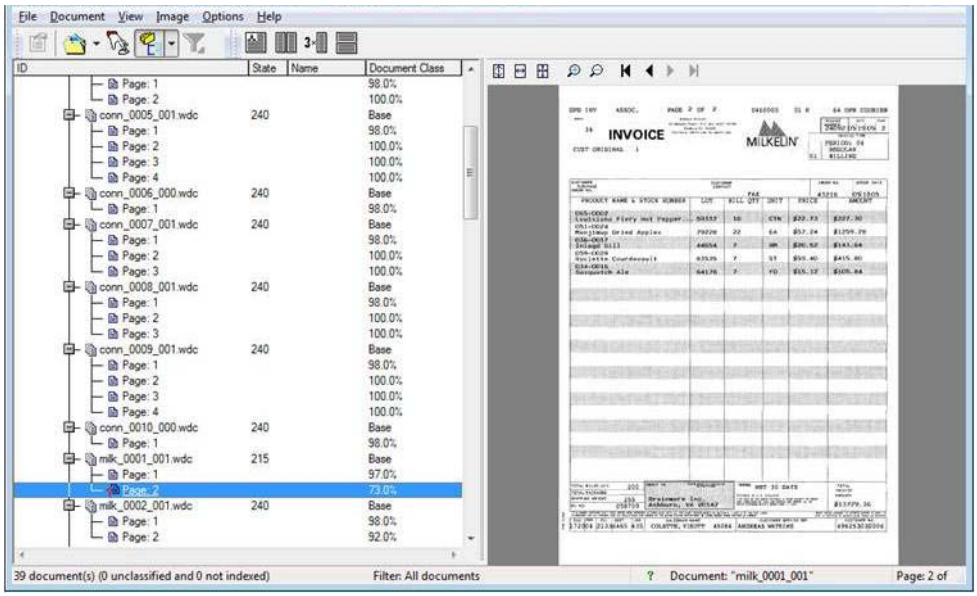


8.1.1.1. Manual Correction of Automatic Page Separation in Oracle Forms Recognition Verifier

In case during automatic page separation in Runtime Server there was at least one unsure page-level decision for a batch of documents, the whole batch gets state “failed page separation”. Such a batch is supposed to be manually reviewed and, if required, corrected in Oracle Forms Recognition Verifier application (see the first element on the screenshot below shown with a special icon indicating that this batch requires manual validation of page separation’s correctness):

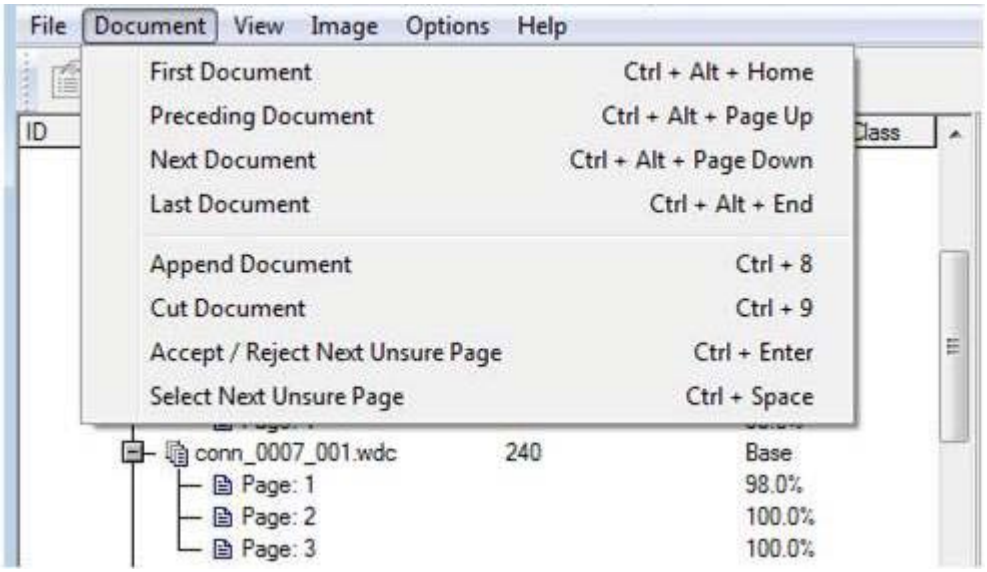


The automatic page separation results can be corrected in Document Browsing mode of Verifier application. When the next batch is opened the system automatically displays the first unsurely split / merged page:

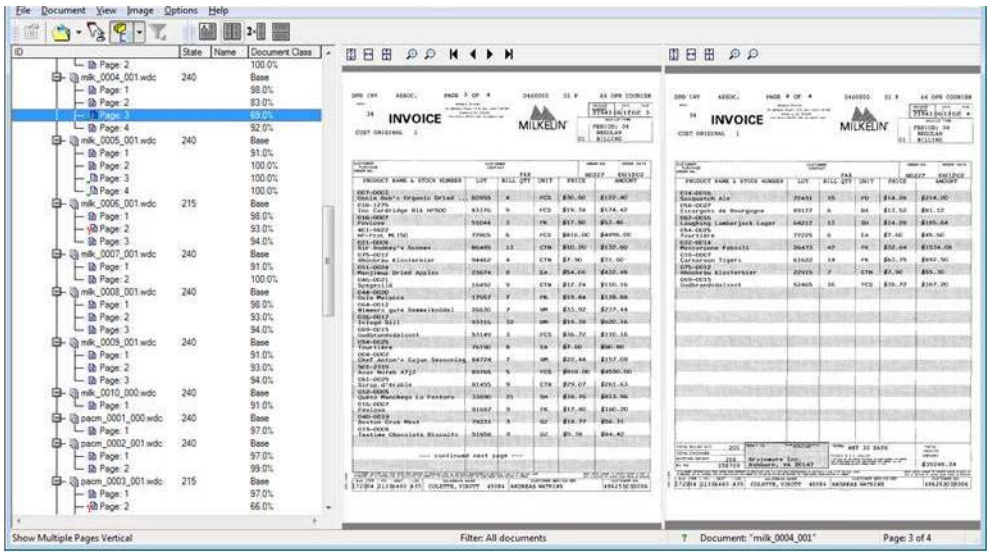


Here the user has the following options:

- Toggle the unsure status (“Accept / Reject Next Unsure Page” menu command; shortcut “Ctrl + Enter”). 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 sign 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 retain after the user closes the batch in Verifier and can be reviewed by the other users. If all pages of a documents become accepted (the pages extracted with high confidence are accepted by default), the document is redirected to successful page separation state (in the example above – “240”). 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 Verifier settings (in the example above – “215”).
- Split the document into two separate documents (“Cut Document” menu command; short cut “Ctrl + 9”). 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, as well as the preceding page automatically get “manually accepted” page correction status.
- Merge selected document with the previous one (“Append Document” menu command; shortcut “Ctrl + 8”). In this case the first page of the currently selected document, as well as 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; shortcut “Ctrl + Space”). This action selects the next unsure page to verify (the one with red question sign) without changing any page states.



When verifying correctness of automatic page separation in Oracle Forms Recognition Verifier, the user can switch to more convenient page view modes, for example having to consequent pages displayed simultaneously:



There are 4 different page view modes available in Oracle Forms Recognition Verifier 10.1.3.5.0:

- Single page view (default).
- Two pages displayed horizontally view.
- Three pages displayed horizontally view.
- Two pages displayed vertically view.

8.2 Manual Correction of Classification Results

Manual correction of classification results is done if your Oracle Forms Recognition Verifier workstation is configured as follows:

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

To determine your settings, check the workflow tab of the Oracle Forms Recognition Verifier Properties dialog box. (See Section ***6.2 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 in ***step 3) page 54***.
5) When all documents in the batch are validated, the application prompts you to select what you want to do next

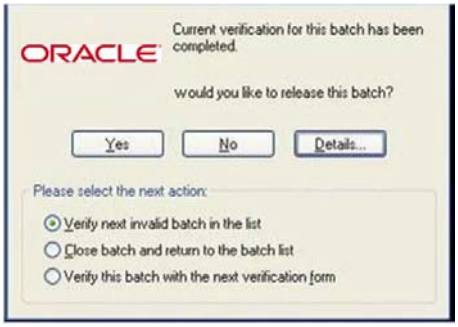


Figure 8-1: Finishing a batch

6) 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.

8.3 Processing of Documents Classified to No Longer Existing Document Classes

8.3.1.1. Description

Verifier application is now 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 be quite often the case in content of supervised learning workflow) can only be processed if their former parent class still exists in the project the document is being processed with.

8.3.1.2. 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 getting deleted (or not inserted) from the global project’s configuration.

8.4 Manual Correction of Extraction Results

Manual correction of extraction results is done if your Oracle Forms Recognition Verifier workstation is configured as follows:

- Classification verification is disabled.
- Extraction verification is enabled.
- Depending on the typical problems, Overwrite Mode should be enabled.

To determine your settings, check the workflow tab of the Oracle Forms Recognition Verifier Properties dialog box. (See Section ***7.2 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, extraction only.

8.4.1. 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. Then use the arrow keys and to select a batch.
2) Once the batch is selected, 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 first invalid field. In Overwrite Mode, the field content is also selected. The user info area contains a message indicating why the field is invalid.

8.4.1.1. 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, Auto-completion, 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 description:

- **Read Only**: When selected, information on a field is dimmed and cannot be selected or edited.
- **Auto-completion**: Enables you to edit text in a field by typing the first two letters of a word. Auto-completion finishes the word with the best matching candidates.
- **Multi-line fields** are 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.
- The **Check Box**: A toggle selection for one of two choices of the data input for a field. Example: Yes/No.

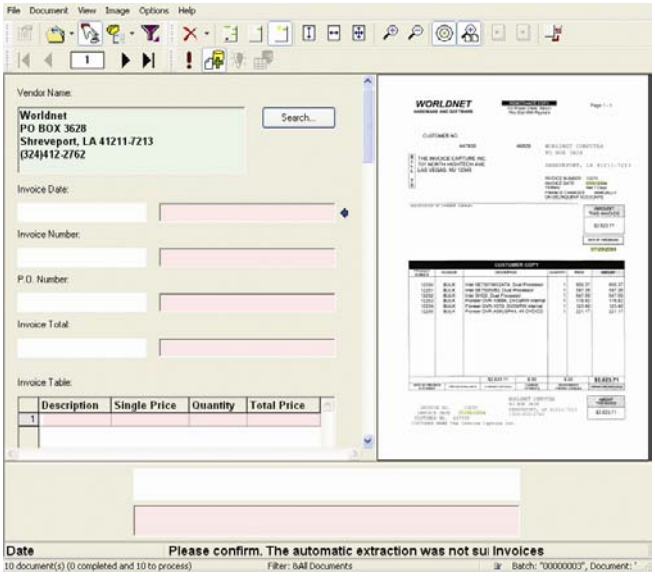


Figure 9-2: A document open in Verifier Plus

You'll notice a Search button in many of the illustrations in this document. This button was created and programmed in Oracle Forms Recognition's scripting language. In these illustrations, the button is used to quickly classify or reclassify as shown in ***Figure 9-3: Vendor Candidates List****. To learn more about the scripting language, please see the Scripting Documentation.*

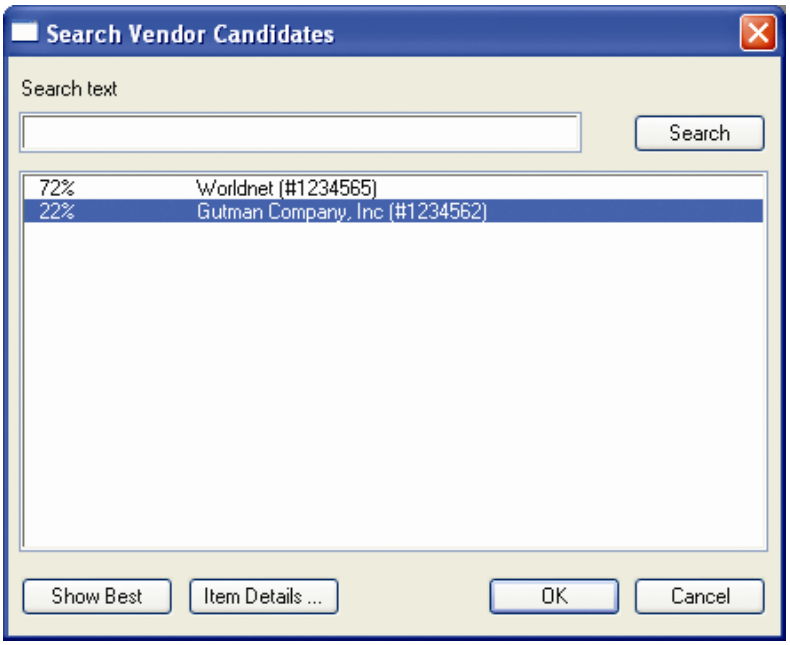


Figure 9-3: Vendor Candidates List

8.4.1.2. Editing Text Fields

Oracle Forms Recognition Verifier includes automated features for editing text fields that can speed up text entry and correction. You can use automatic character entry when the auto-completion is selected in the form field Properties dialog box to edit text fields and cells. Other 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 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. To add a new line to a multi-line field, press CTRL+ENTER.

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.

8.4.1.3. Auto-Completion

Do not enter formatted text for auto-completion. Auto-completion does not work on formatted text and characters incorrectly read by OCR.

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.

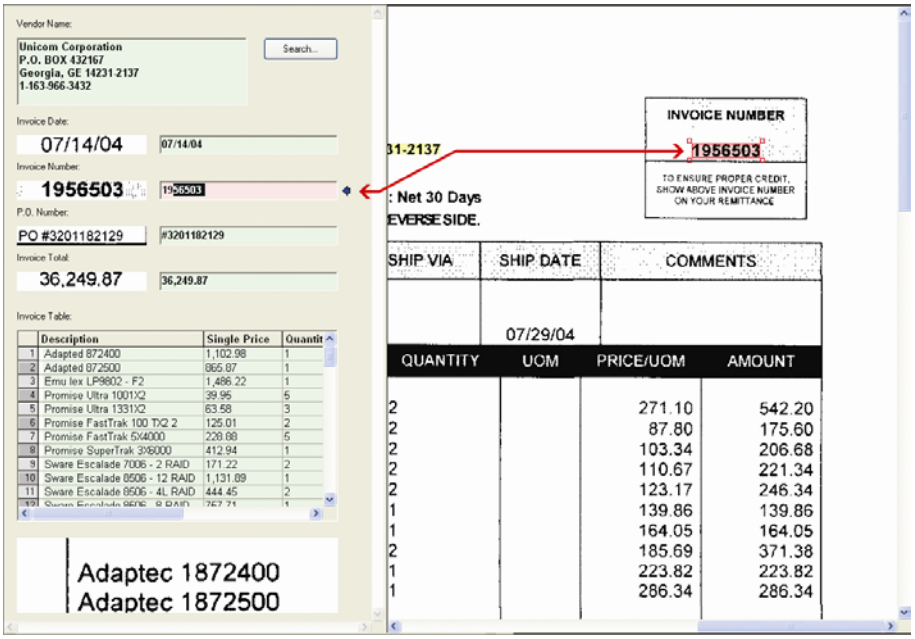


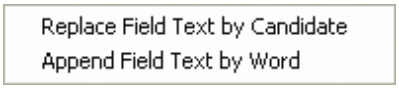
Figure 9-4: Example of auto-completion

8.4.1.4. 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. It will appear in green (with a border of green check marks if that visual indicator is enabled in Batch Options) when you select it after selecting the field. Non-candidates will display in orange when selected. 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 append to or replace a word in a field box. A candidate is one that matches the learned selections for the field. It will show up in green when you select it after selecting the field, or with the additional green check marks around it if that visual indicator is enabled in Batch Options. The append feature takes the current word left of the candidate and appends the field text. It places the text in the best location, either right or left of the word, and places the field by text or location of the word. Or, you can replace text. For example, a blank candidate might be replaced by “285.98.”



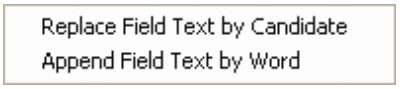
- To append text with the new text, click on the desired word next to the text that you want to append (this text will appear in green if it is a candidate.) A box appears around the word. Double-click on the box, or right-click in the document and select Append Field Text by Word.
- To replace a word, click on the desired word. A box will appear around the word. Double-click on the desired candidate, drag and drop the word to the field, or right-click in the document and select Replace Field Text by Candidate from the shortcut menu to replace it.

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

Make certain that this word fits the format analysis rules defined for that field. If not, the word is highlighted in orange (with orange exclamation marks around it if validity icons are enabled.) In this case, it would not be a good candidate for the field.

Words Without Candidates

Even if the word does not belong to any candidates for the field, you can append or replace a word with a new one. Appending places the text in the best location, either right or left of the word, by text or location of the word. Or you can replace the field text and location by the text and location of a word. A word that does not belong to any candidates for that field will display in orange when selected. For example, a field named "sales total" might be replaced by "invoice total."



- To append text with the new text, drag a box around the desired word. Double-click on the desired word in the box, or right click in the document and select Append Field Text by Word.
- To replace text, select the desired word with the mouse. A box will appear around the word. Double-click it, or select Replace Field Text by Candidate in the shortcut menu.

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. If not, the word is highlighted in orange (and with a border of orange exclamation marks if validity icons are enabled) to help distinguish it. If so, it would not be a good candidate for the field.

8.4.1.5. 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 for cell descriptions. Before you can insert blocks of text, first select the settings in the Workflow dialog box to immediately copy information. (See Section ***7.2 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.

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.

8.4.2. Finishing the Validation

- Once a field is corrected, press ENTER to validate it. During validation, the field's background color appears in yellow, and the cursor becomes an hourglass. Once the validation is finished, the cursor moves automatically to the next invalid field, regardless of whether this field is still in the same or already in the next invalid document. If you leave a document this way, it is validated automatically. In the next document, proceed as described in step 1 of ***8.4.1 Correcting Invalid Results*** or in step 2 of ***8.2 Manual Correction of Classification Results*** 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.

8.5 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 disabled3.

To determine your settings, check the workflow tab of the Oracle Forms Recognition Verifier Properties dialog box. (See Section ***7.2 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.

8.6 Manual Indexing

Manual indexing is done if Oracle 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 Oracle Forms Recognition Verifier workstation is configured as follows:

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

To determine your settings, check the workflow tab of the Oracle Forms Recognition Verifier Properties dialog box. (See Section ***7.2 Settings – Workflow***)

---

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

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 document manually:

1) In the Batch View, check the State column to find a batch you can verify. Then use the arrow keys , and · to select a batch.
2) Press ENTER to open the Verification View. 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. To fill the current field with the value from the previously validated document, press F8. To undo this, press F8 again.
4) Press ENTER to validate your entry. The cursor moves automatically to the next field, no matter whether this field is still in the same document or in the next invalid document. If you leave the current document this way, it is validated automatically. For the next field, proceed as described in step 1 of ***8.4.1 Correcting Invalid Results***.
5) When all documents in the batch are indexed, the application prompts you to select what you want to do next
6) Do 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: Changes verification forms, using the next verification form.

8.7 Smart Indexing

Organizations, in particular commercial ones, usually collect legions of information about themselves and everybody they do business with. Much of this information is stored in databases. For example, there will hardly be a company without a customer database that contains addresses, contacts, and so on. Financial transactions are also recorded in databases. Databases can be excellent support for indexing because they store related information, and this information can easily be retrieved. During indexing, if you have one piece of information from a document, you can obtain related pieces from the database and fill the associated fields automatically. This method is called smart indexing.

Normally, smart indexing is combined with manual indexing. Some fields of a form have to be filled in manually; some fields can be filled automatically.

Example:
Let’s assume that your organization saves information related to orders in the database of its ERP system. Every order is characterized with a unique identifier and some attributes about the supplier and the items that have been ordered. Soon after an order is placed, the ordered items are delivered, and a delivery note is attached. The corresponding invoice follows soon. Delivery note and invoice refer to the original order: They have the order’s unique identifier printed on them. With this identifier, you can look up supplier information from the database when you verify the delivery note and invoice. However, new information such as the invoice

date has not yet been entered into the database. This information can be supplied manually.

To use smart indexing:

1) Smart index fields can be recognized by the Key icon that is displayed next to them. Select a smart index field. The field itself and all the fields that can be filled via the database lookup are marked with a yellow cylinder.
2) If the field is still empty, enter the field value. Alternatively, enter a wildcard expression, using the * to represent a sequence character, or the ? to represent a single character.
3) Do one of the following to start the lookup:
   - If your application is configured accordingly and the field content is correct, validate the smart index field by pressing ENTER.
   - Press ALT+F12.
4) The system may respond as follows:
   - If the lookup yields no results, a corresponding message is displayed. Fill the lookup fields manually. If you cannot complete the fields, send the document to exception handling.
   - If the lookup yields one result, the lookup fields are filled.
   - If the lookup yields multiple results and this is allowed in your application, the lookup fields are filled.
   - If the lookup yields multiple results and this is not allowed in your application, a dialog box is displayed where you can select the correct record. The lookup fields are then filled accordingly.

8.8 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. Use the arrow keys to select a batch.
2) Once the batch is selected, press ENTER to open the Verification View. The Verification View with the first document requiring correction is displayed.
3) To display the first document in the batch, press CTRL+ALT+HOME.
4) You may encounter a document that has been classified incorrectly. To correct this result, press F7 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, and then press F3 to return to the document.

To get to the next document, press CTRL+ALT+PAGE DOWN. Proceed like this until the last document is reached.

Chapter 9 Working with Tables

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

In Oracle Forms Recognition, it's possible that any given table was either trained by the new extraction engine – Table Extraction – or the traditional Table Analysis Engine. As a Verifier user, you won't necessarily know which was used, and you won't be able to see a difference in Verifier. The process of table fields is similar.

The difference between the two workflows – Verifier or Advanced Verifier – lies in the fact that tables extracted through Table Extraction can be learned and corrected differently by Supervised Learning Verifier users. Even if you're a Supervised Learning Verifier, you won't be able to train and correct traditionally created tables the same way you can a Table Extraction table.

9.1 Automatic Training and Extraction of Verified Table Data

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

9.2 Traditional Training and Correction Methods

9.2.1. Using Auto-Completion

Auto-completion works in table cells and with text fields. When you type two or more characters, auto-complete suggests a word or phrase for that cell. The candidate appears in green if the field is valid, and red if the field is invalid. (If the visual validity icons are enabled in Batch Options, valid fields also have a border of green check marks and invalid fields have a border of red question marks.)

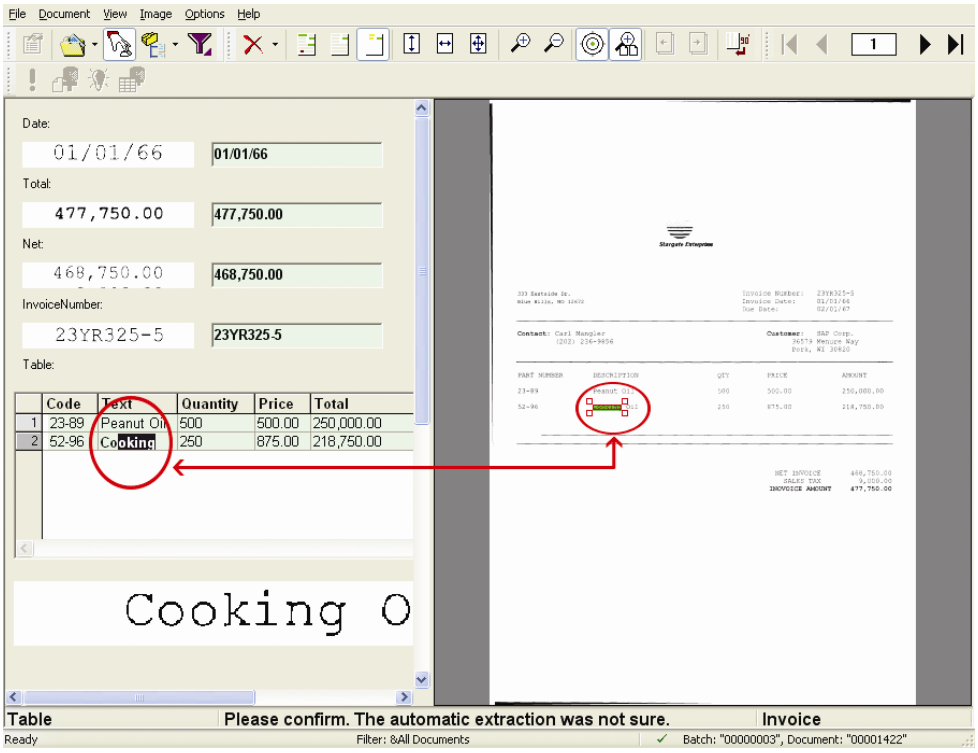


Figure 10-1: Table with text auto-completion

9.2.2. Inserting Words in Table Cells

You can insert single words or append existing text in table cells.

9.2.2.1. Words That Are Candidates for Cells

If the word belongs in a cell area, you can append or replace a word in a cell. The Append feature takes the current word behind the candidate and appends the cell text. It places the text in the best location, either right or left of the word, and cell location by text or location of the word. The word belonging to a cell area will highlight in green when selected. Or, you can replace text.

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. If you have candidates, double-click on the desired candidate to replace it, or right-click in the document, then select Select Cell from the shortcut menu.

In the search region, word candidates are all words that are not covered (by location) by other table cells and that have the same beginnings as the whole text of the cell.

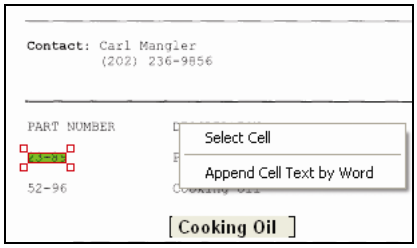


Figure 9-2: Table cell, word in cell area

9.2.2.2. 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, then press CTRL + left click, or select Replace Cell Text by Word in the shortcut menu.

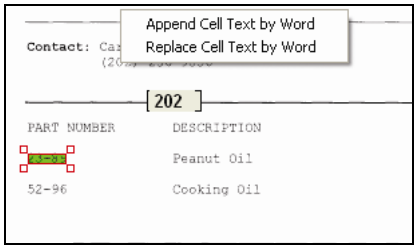


Figure 9-3: Table cell, word not in cell area

9.2.2.3. Correcting Table Structure

You may also need to correct the table structure. Table rows and columns have shortcut menus with options for modifying the table structure. To invoke them, right-click on the row or column label. The available commands are summarized below.

Shortcut menu	Command	Description
Column		
	Unmap	Clears all data for the selected verification column and turns the state of the corresponding column of the recognized table back to “unmapped.” To view an unmapped column, double-click on the table header in the verification form. All unmapped columns are highlighted in red.
	Map	Adds the column selected from the shortcut menu. Or you can right-click on an unmapped column to map it to a column in the verification form.
	Swap	Exchanges the position of the current column and the one selected from the drop-down menu.
Row		
	Insert	Inserts an empty row above the current one.
	Delete	Deletes the current row.
	Duplicate	Duplicates the current row.
	Append	Appends an empty row at the bottom of the table.
	Merge	Merges cells in a row.

Table 9-1: Table shortcut menus

To do this task, proceed as described in Section ***9.3.3*** (***How Table Extraction Learns (Standard Method)***), Section ***9.3.4*** (***Advanced Learning with Table Extraction***), and Section ***9.3.5*** (***Advanced Learning: Additional Functions***). In addition, 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. Press F7 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.

9.3 Table Extraction and Correction

This functionality is available on the Supervised Learning Verifiers.

9.3.1. About Table Extraction

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

- Learning lines
- Learning mappings of columns

These are discussed in detail in the following sections.

9.3.1.1. Learning Lines

The 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 is searching only in 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 on the number of words in secondary lines, and no limit on 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 this will take effect only 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.

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 (CTRL+Q) menu option in Verifier or via pop-up menu options Show Lines and Show Columns of the document viewer in Verifier or Designer.

9.3.1.2. 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 Oracle Forms Recognition class, the Learn Set 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 Learn Set after document validation) will be applied for all BTE tables in the document.

9.3.2. Correcting Fields in Tables Created with Table Extraction

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

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

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

When you work with interactive table extraction, you must learn lines before you map columns.

9.3.3. How Table Extraction Learns (Standard Method)

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

1) Show the first row sample.
2) Learn mapping in the learned row.
3) Learn missing lines.
4) Learn and adjust the mapping of missing or wrong columns.
5) Manually correct the table date and validate the table.
6) Learn the document.

These steps are discussed in detail below.

9.3.3.1. Step 1. Show The First Row Sample

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

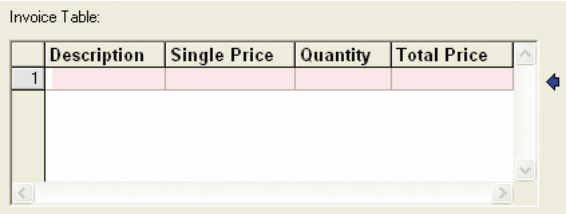


Table 10-2: BTE Table

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-click on the selected row area. Select the whole first row and learn it.

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.

Learning a New Line as Primary Line

1) Right-click on any line marked in gray in the TIFF.
2) On the shortcut menu, select Learn Line.

The learned lines change from gray to green, or to blue if the line is extracted with low confidence.

Learning a Block of Lines as Primary Lines

1) In the document viewer draw a rectangular selection over the primary lines 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 highlighted in green (or blue if the line is selected with low confidence,) and all other lines are similarly extracted and displayed.

If some lines were not extracted (these lines won't be color-coded,) try relearning the lines singly 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 highlighted in green (or blue if the line is extracted with low confidence,) and all other lines are similarly extracted and displayed.

If some lines were not extracted (these lines won't be color-coded,) repeat the procedure described immediately above.

Do not try to learn the rest of 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 will learn another line sample the engine will try to apply existing mapping rules for the newly learned row automatically, which significantly reduces time to train the table.

Green highlighting means a line is extracted with high confidence, blue highlighting - with low confidence. If the confidence for a blue line is less than 0.3 (moving the mouse cursor over the highlighted lines highlights the confidence value as a tool-tip) then the lines will not be extracted. Blue highlighting has also the following important meaning: this line can be trained by the engine as a new line class.

9.3.3.2. Step 2: Learn Mapping in the Row You Learned

1) Switch to the columns highlighting mode now (using CTRL+Q) and mark location of your first cell item in the row you learned. The system

will pop up a special mapping control asking for the desired data column to extract the data to.

2) Select the required data column and double-click on the selection.
3) Repeat this step for the rest of the cell items in the first row.

9.3.3.3. Step 3: Learn Missing Lines

1) Switch back to the lines highlighting mode.

Pressing CTRL+Q switches the highlight between three modes: Cells, lines, and columns. Press CTRL+Q twice to switch from columns learning to lines learning.

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 "***Step 4: Learn and Adjust the Mapping of Missing or Wrong Columns***" only after you are sure nothing is missing. Go to the next step only after you are sure nothing is missing.

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

Return to columns mapping learning mode and look for wrong or missing mapping. Correct any missing mapping.

If you can't map the missing columns, switch back to lines highlighting mode and try to learn the row where the mapping was missing.

Switch to columns highlighting. If the mapping is still missing, mark the missing part and map it.

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.)

9.3.3.5. Step 5: Manually Correct the Table Data and Validate the Table

Now and only 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 input in one shot.

9.3.3.6. Step 6: Learn the Document If Required

After table learning and validation have been completed and the rest of document's fields are validated, you may want to add this document to the Learn Set and then learn it (so-called "cross-document" learning in contrast with "in-document" interactive BTE learning.) If the system did not suggest learning the document automatically (The Add Current Document to Learn Set toolbar button is not pushed,) but you still would like to learn your table, activate learning by clicking the Add Current Document to Learn Set button.

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.

9.3.4. Advanced Learning with Table Extraction

9.3.4.1. 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 not mapped 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 the Learn as Secondary Lines option.

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 not-mapped 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 primary lines will be learned and highlighted in green (or blue if the line is selected with low confidence,) and all other lines are similarly extracted and displayed.

If some lines were not extracted (these lines won't be color-coded,) repeat the procedure described immediately above.

9.3.5. Advanced Learning: Additional Functions

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

9.3.5.1. Unmap Column

The Unmap Column method can undo mapping for the specified cell item.

This will undo mapping for all cell items that were extracted from the lines that belong to the same line type as the cell item used to invoke Unmap Column method.

Undo Column Mapping

To undo incorrect column mapping:

1) Right-click on any unassigned column – they are marked in blue – or draw a rectangular selection over the cell items to be mapped to a table column.
2) On the shortcut menu, select Undo Mapping.

The previously assigned column (highlighted in red) is now unassigned (and displayed in blue.) The values are no longer extracted or in the table grid.

9.3.5.2. Unlearn Line

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

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.

9.3.5.3. Learn Line as Wrong Line

Learning a Wrong Line means training the table so 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.
2) On the shortcut menu, select Learn as Wrong Line. The selected lines and similar lines to it are now highlighted in gray. Information from these lines will not be extracted.

Chapter 10 Working with the Learn Set Manager

10.1 What Does Supervised Learning Do?

You have access to the Learn Set Manager mode only if you have been assigned to a group that has permission to work with the mode.

There is no limit on the number of users who can simultaneously access the Learn Set Manager.

The basic purpose of the Learn Set Manager is to use Supervised Learning (interactive training) to improve the quality and usefulness of your enterprise's Learn Sets. With Supervised Learning, Supervised Learning Verifiers and Learn Set Managers can customize your project's Learn Sets by adding or subtracting documents, reclassifying them, creating altogether new classes or Learn Sets and migrating documents there, and promoting Local Learn Sets to a Global Learn Set that can be shared across the enterprise.

In general, Learn Set Management consists of:

1) Creating new classes based upon documents themselves and supplier information
2) Learn documents and adding them to the Local Learn Set.
3) Using the Local Learn Set to improve the extraction of low-quality documents.
4) Maintaining Local Learn Sets.
5) Updating and enhancing the Global Learn Set with information from the Local Learn Sets.

All this is done through a simple interface that closely resembles that of Verifier and Advanced Verifier.

Although Supervised Learning was created for use with vendors' invoices, it can also be used with other types of knowledge. For example, a library might create classes based on type of material, subject matter, or author. Most of the illustrations and examples in this chapter use invoices.

10.2 Starting and Exiting the Learn Set Manager

To start the Learn Set Manager, click the funnel-shaped button in Verifier.

To exit the Learn Set Manager, go to the File menu in the Learn Set Manager and click Exit, or, on your keyboard, press CTRL+E.

10.3 Getting Familiar with the Learn Set Manager User Interface

The Learn Set Manager has two basic modes, or views. These are the Accumulated Documents View (where you work with Local Learn Sets) and the Global Learn Set View (where you work with Common Learn Sets and Global Learn Sets.)

When you are working with Local Learn Sets, you can further refine the appearance of the Accumulated Documents Browser when you verify documents or manually reclassify them.

10.3.1. Working with the Accumulated Documents Browser

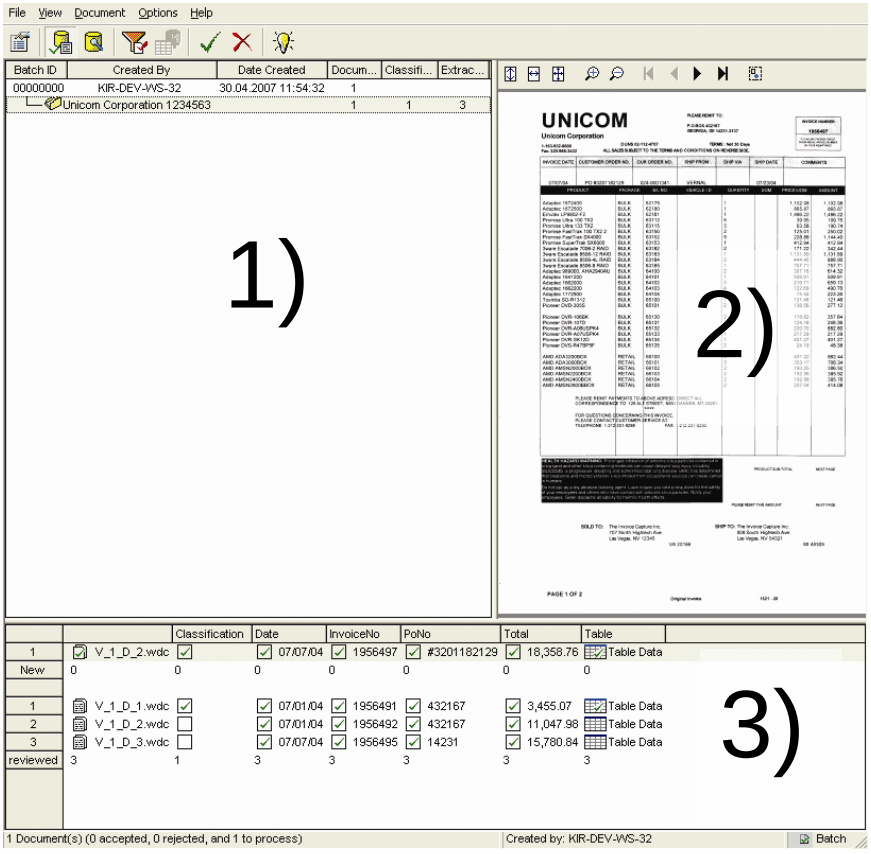


Figure 10-1: Accumulated Learn Set (Local Learn Set) View

1) **The Batch viewer**. Enables you to see each class in the batch you are working on. The Batch Viewer shows each class as part of the batch, the user who created the batch, the date it was created, the number of documents in the batch and in each class, the number of documents successfully classified, and the number of documents successfully extracted. You'll probably have to enlarge the window to see all of these categories.
2) **The Document Viewer**. As with the Document Viewer in Verifier, this window enables you to see (and therefore verify) each document in the batch you're working on.
3) **The Learning Statistics Window**. Shows the documents that have been processed by Brainware. Documents that are awaiting processing have a question mark beside them. Successfully processed documents have a check mark, while documents that failed processing have an X.

10.3.1.1. Global Learn Set Browsing

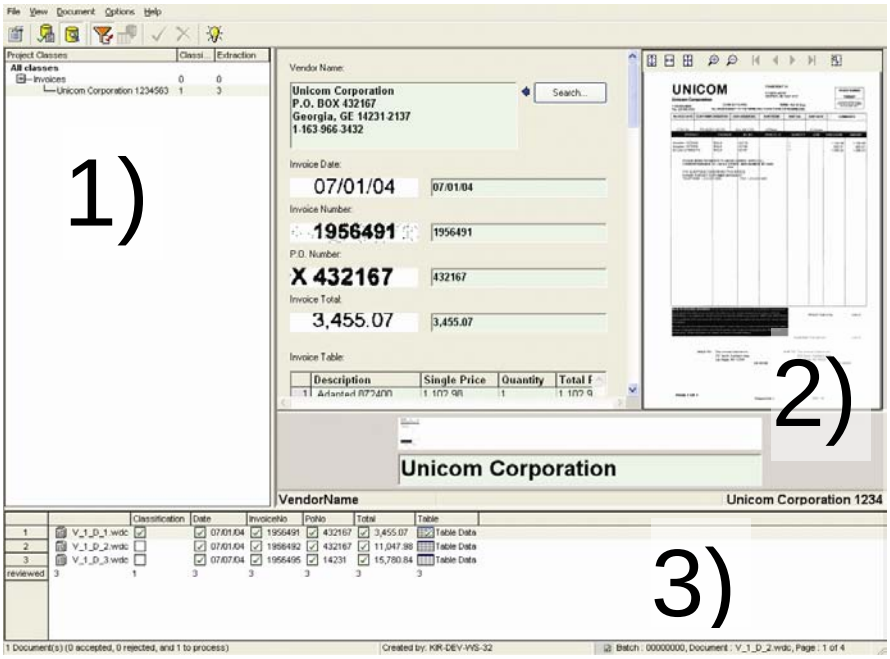


Figure 10-2: Global Learn Set Viewer

1) The Batch viewer. Enables you to see each class in the batch you are working on. Shows the classes in the Global Learn Set, and the number of documents classified or extracted in each.
2) The Document Viewer. As with the Document Viewer in Verifier, this window enables you to see (and therefore verify) each document in the batch you're working on.
3) The Learning Statistics Pane. Shows the documents that have been processed by Oracle Forms Recognition. Documents that are awaiting processing have a question mark beside them. Successfully processed documents have a check mark, while documents that failed processing have an X.

10.3.1.2. Menu Commands and Keyboard Shortcuts

Menu	Submenu/ Command	Cmd	Keyboard Shortcut	Description
File				
	Exit	n/a	Ctrl E	Closes the Supervised Learning Manager
View				
	Toolbar	n/a	n/a	Toggles the toolbar on or off
	Status bar	n/a	n/a	Toggles the status bar on or off
	Process Accumulated Documents	n/a	n/a	Switches to the Local Learn Set Viewer
	Global Learn Set Browsing	n/a	n/a	Switches to the Global Learn Set Viewer
	Verify Documents	n/a	F5	Verifies documents in Supervised Learning
	Refresh	n/a	n/a	

Table 10-1: Menu Commands and Keyboard Shortcuts for Supervised Learning (File & View)

Menu	Submenu/ Command	Cmd	Keyboard Shortcut	Description
Document				
	Accept	n/a	n/a	
	Reject	n/a	n/a	
	Learn All Processed	n/a	n/a	
	Learn Accepted and Remove Rejected	n/a	n/a	
	Learn Accepted Only	n/a	n/a	
	Remove Rejected	n/a	n/a	
	Learn All New	n/a	n/a	
	Relearn All in Global Learn Set	n/a	n/a	Available only to Supervised Learning Managers.
	Relearn Changed in Global Learn Set	n/a	n/a	Available only to Supervised Learning Managers.
	Correct Tables	n/a	n/a	Use to improve the quality of data in tables. Available to Verifiers and Managers.
	ReClassify	n/a	F7	Manually reclassifies documents. Available to Verifiers and Managers.

Table 10-2: Menu Commands and Keyboard Shortcuts for Supervised Learning (Document)

Menu	Submenu/ Command	Cmd	Keyboard Shortcut	Description
Options				
	Train Base Classes...	n/a	n/a	Available to Verifiers and Managers.
	Sort Batches by Vendor...	n/a	n/a	Available to Verifiers and Managers. Not available in Global Learn Set Browsing.
	Update Local Projects...	n/a	n/a	Available to Verifiers and Managers.
	Open with Oracle Forms Recognition Workdoc Browser	n/a	n/a	Opens the Workdoc Browser.
	Settings...	n/a	n/a	Shows settings for the Learn Set Manager
Help				
	Oracle Forms Recognition Verifier Help	n/a	n/a	Opens Verifier Help
	Learn Set Manager Help	n/a	n/a	
	Supervised Learning Help	n/a	F1	Opens Verifier Help

Table 10-2: Menu Commands and Keyboard Shortcuts for Supervised Learning (Options & Help)

10.3.1.3. Toolbar Buttons

The toolbar provides quick access to some frequently used commands.

Button	Description
	Show Settings.
	Switch to Accumulated Documents Processing (The Local Learn Set.)
	Switch to Global Learn Set Processing.
	Verify Documents. Activates the Advanced Verifier mode.
	Correct tables. Lets you correct data in the tables. You must have clicked a table field for this to be active.
	Accept documents.
	Reject documents.
	Learn documents. (Adds them to Global Learn Set.)

Table 10-3: Toolbar buttons for supervised learning

10.3.1.4. Viewer Toolbar Buttons

On the Viewer toolbar, you can use the following commands to adjust the size of a document relative to the width of the Document Viewer window:

Button	Description
	Fits the document to window height.
	Fits the document to window width.
	Best fit.
	Zooms in. Alternatively, press Ctrl + .
	Zooms out. Alternatively, press Ctrl - .

Table 10-4: Viewer toolbar buttons for supervised learning

10.4 Using the Learn Set Manager

10.4.1. Overview of the Process

Using the Learn Set Manager to work with Local Learn Sets is basically a six-step process: You verify documents, decide whether they belong in the Learn Set, add them to the Common Learn Set, and train the Learn Set.

In the Common Learn Set, you examine the documents for inclusion in the Global Learn Set, accept or reject them and add them to the Global Learn Set.

After that, you train the Global Learn Set.

10.4.2. Getting Ready to Use the Learn Set Manager

To use the Learn Set Manager:

1) In Verifier, examine the properties for Learn Set Manager. (***Figure 10-3***) Most were established in Designer. However, you need to ensure that Learn Set Manager is enabled – that the Activate Supervised Learning workflow checkbox is marked. Also, ensure that the paths for Local Project Name, Local Learn Set Directory, and Knowledge Base Directory are correct.
   As a rule, the addresses for each of these base settings should be the same except for the last part.
   - Local Project Name should end with the project name (for example, newproject.sdp).
   - Local Learn Set Directory should end with the Learn Set Directory folder (it's probably called Learn).
   - Knowledge Base Directory should end with the common Learn Set folder. (Likewise, it's probably called Common.)

   You can change these settings only if you are an Administrator, Supervised Learning Verifier, or Learn Set Manager.

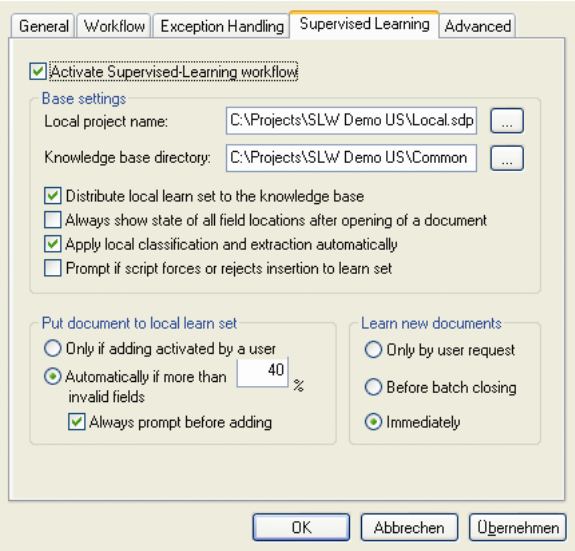


Figure 10-3: Learn Set Manager Settings Tab

2) Launch the module from Verifier.
3) On the Learn Set Manager toolbar, click the Settings button. Make sure Show learned state using engine-level information is

checked.This setting indicates whether the particular field or document was used by the system for learning. If required, a user can also disable learning for the desired field / document.

10.4.3. Working with Common Learn Sets

Your first task in Learn Set Manager is to attain the highest-quality Learn Sets possible. You'll use Learn Set Manager to build upon Local Learn Sets created earlier in the process.

1) When you launch Supervised Learning from Verifier, the Accumulated Learn Set Mode should appear by default. The Accumulated Learn Set Mode is the one you use to work with Common Learn Sets. The Accumulated Learn Set Mode is activated when you click the Switch to Accumulated Documents Processing button or from the Options Menu >Process Accumulated Documents.
2) In the Accumulated Learn Set Mode, select a batch to work on.
3) Double-click on a class to select it.

4) Select a document to work on and verify the document just as you would in the traditional Verifier. Click on the Funnel button in the toolbar. This opens the Advanced Verifier Mode, where you can correct or verify the contents of each field and table.

5) After you have verified the document click the Accept button. This marks the document for learning as the first step for promotion into the Common Learn Set. (You could have also clicked the Reject button to eliminate the document from being considered for the Common Learn Set.)
6) To select another document from the batch. This is done from the Learn Statistics Pane at the bottom of the screen, where you'll double-click on a document to open it in the Document Viewer and work on it.

7) When you've verified all the documents you need to verify, click the Learn button. This promotes all the accepted documents to the Common Learn Set. Notice that the Learn Statistics window for the Local Learn Set is now empty.

10.4.3.1. Correcting Tables

To correct tables, you must first select a table in a document and then click the Correct Tables button. This enables Supervised Learning Managers and Verifiers to interactively train all the tables on a document form (not just the table you selected so you could activate the Correct Tables button.)

From there, table correction in Learn Set Manager proceeds just like it would in Verifier itself. (Section ***8.8***)

10.4.3.2. Reclassifying a Document

To assign a document to a different class, select the Document menu from the Main Menu and select Reclassify (or press F7). This opens a dialog box in the Verifier Document Viewer where you can assign the document to a new class. When you've selected the new class, press ENTER

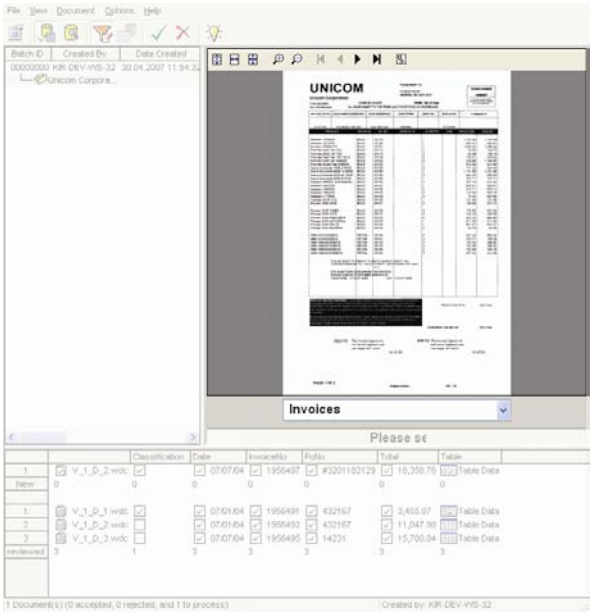


Figure 10-4: Assigning a document to a new class.

Manual reclassification will only work with the classes that Verifier is currently using, not the classes that have already been learned.

10.4.3.3. Accepting and Rejecting Documents

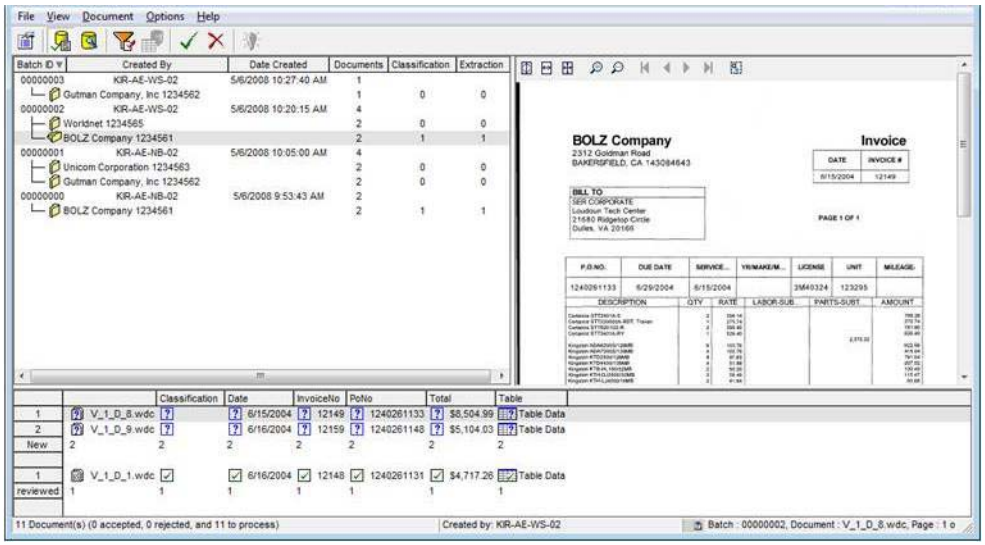
There are two ways to accept or reject a document or batch from a Learn Set.

The first is the traditional way, by using Verifier to manually screen and verify the document or batch. The other method is by comparing documents in the Common Learn Set to the corresponding batch in the Global Learn Set.

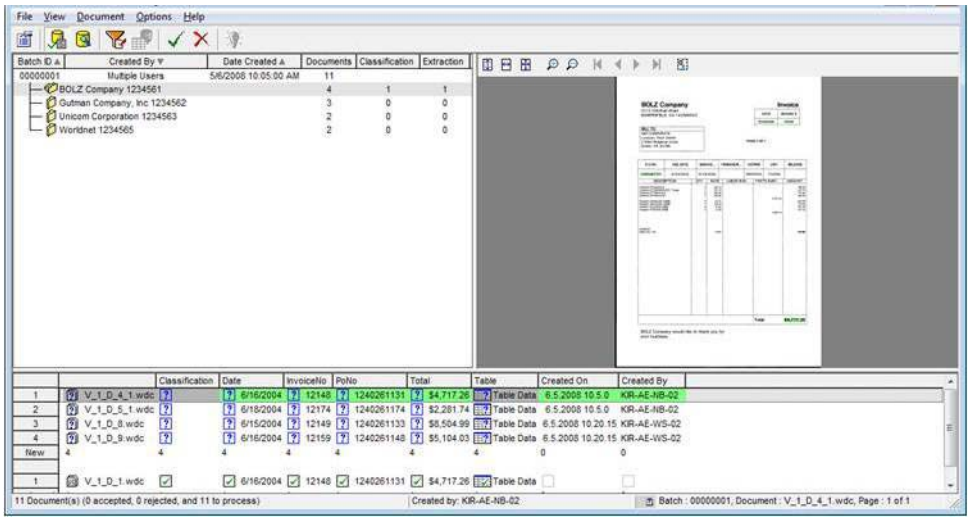
10.4.4. Ability to Sort by Vendor & Other Sorting Extensions in Learnset Manager

10.4.4.1. Description

The Learnset Manager tool of Oracle Forms Recognition Verifier can now sort by vendor name across multiple batches produced by different local supervised learning Verifiers.



By selecting the “Options \ Sort Batches by Vendor” menu item, the system is going to rebuild the batches of documents created via multiple sessions by multiple Advanced Verifier workstations and allow the Learnset Manager user to sort by vendor name. In this connection, each vendor folder is going to cumulate all available documents for this vendor, so that the user could select the best documents to train the global project with.



The “Created On” and “Created By” data fields are then showed separately for each particular document in the document view at the bottom of the Learnset Manager screen.

User sorting preferences in both “Global Learnset Browsing” and “Accumulated Documents Processing” modes are supposed to be saved and then restored next time the user starts the Learnset Manager tool.

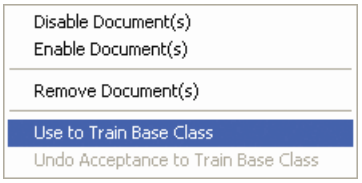
10.4.4.2. Usage

This simplifies SLW decisions as to which documents to train for a specific vendor class. With the help of the present feature, the user is now able to review all documents (for the same vendor class) created by different Advanced Verifier workstations at once.

10.4.5. Working with Global Learn Sets

In the previous section, you learned how to work with Common Learn Sets and to promote documents to the Global Learn Set. The Global Learn Set is where you further refine the quality of your data so that it can be migrated into an effective, useful Global Knowledgebase.

1) To work with Global Learn Sets, click on the Global Learn Set Viewing button.
2) Again, begin your work at the document level by examining the quality of the data in the document. As before, you select the document from the Learning Statistics window at the bottom of the screen. If you are satisfied with the document, right-click on it, select Enable, and then select Use to Train Base Classes to Train Base Classes.



3) Click the Learn button.
4) Confirm that you do want to learn the document

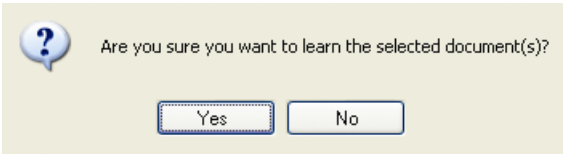


Figure 10-5: Confirming learning


5) After you've verified each field in the document, click the Check Mark button on the toolbar to accept the document for processing (or click the X button to reject it.)

6) Now retrain the Learn Set by clicking the Light Bulb button on the toolbar.

10.4.6. Training Base Classes

The final milestone in creating or enhancing your Global Learn Set is to train base classes.

On the Options menu, select Train Base Classes. Select the base class to train.

Under Train Selected Base Classes, select a value. To avoid errors while maintaining the quality of your sample, select the lowest value possible.

Click OK.

10.5 Updating Local Projects

The ability to update local projects (***Figure 10-6***) is important for keeping your Learn Sets in sync with each other by enabling you to automatically synchronize them.

To update local projects:

1) On the Options menu, click Update Local Projects. Configure the selections described below and click Update. This procedure can take a while, especially if you're updating projects on a network.

Locked projects will not be updated.

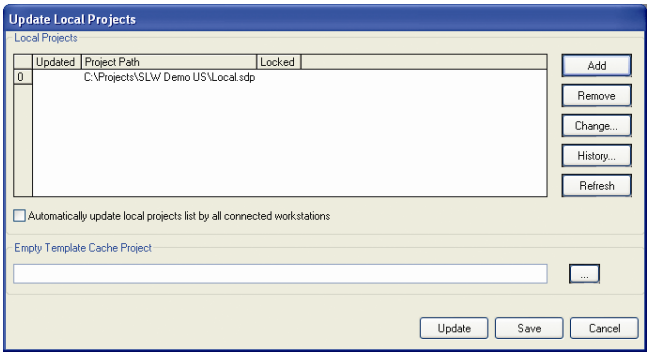


Figure 10-6: Updating Local Projects

This dialog box specifies a list of local or network project paths to be managed. Here, you can:

- Add project paths to the Update list (by clicking the Add button and browsing to the project.)
- Remove paths from the Update list (by clicking the Remove button and browsing to the project.)
- Change existing project paths (by clicking the Change button and browsing to the project.)
- View a history of all Verifier workstations that have connected to the Common Learn Set. The History shows the workstation name, the time and date of its last connection and the local project path.
- This list is updated every time a Verifier station creates a new batch of locally learned documents in the common Learn Set.
- Refresh the list of projects to see the most recent update information about them
- Empty the template cache.
- Update the list of projects or save the new criteria without actually updating the projects.

For each configured path, the dialog box shows whether a project is up to date, whether it is locked (more on this in Section ***10.6***) and whether the project is available. (See ***Using Learn Set Manager on Several Workstations***)

An up-to-date project has a green check mark beside it; a project that has not been updated has a red X. Pathnames of unavailable projects are dimmed.

The settings you establish above will be available for any workstation on which the Learn Set Manager is opened.

10.6 Using Learn Set Manager on Several Workstations

Learn Set Manager can be used simultaneously on more than one workstation, thanks to the application's ability to lock projects and files.

Learn Set Manager supports three levels of protection that facilitate this ability:

- Batch-level locking.
- Allowing all Learn Set Manager workstations to view changes made by all Supervised Learning Managers.
- Locking project files and Learn Sets while they are being trained.

10.6.1. Batch-Level Locking

Batches are locked during processing to prevent several users from updating the same batch at the same time. No one else can access the batch until processing is complete and the batch is closed.

10.6.2. Allowing all Learn Set Manager Workstations to View Changes Made by all Supervised Learning Managers

The changes applied by one Learn Set Manager user should be visible to all other Learn Set Manager users. To accomplish this, Learn Set Managers must use two predefined batch document states: “981, accepted” and “982, rejected.”

When learning is executed, the Learn Set Manager application checks to see if documents with either of these assigned states have been added to the Local Learn Set. Documents with a state of 981 are added to the Global Learn Set. Documents with a state of 982 are not added to the Global Learn Set.

10.6.3. Project file and Learn Set Locking During Training

Only one workstation can do learning at one time. This means that the learning process is locked, and therefore not available for other users, if one user has initiated learning.

10.7 Using the Workdoc Browser

The Workdoc Browser utility is available as a menu item in Learn Set Manager.

The Workdoc Browser enables you to examine the attributes and characteristics for any or all Workdocs in a batch. When you launch the utility from Learn Set Manager, the Workdoc for the document displayed in your Document Viewer is called. The chief reason for using the Workdoc Browser would be to examine the relationships between attributes and characteristics for the purpose of writing scripts to enhance Oracle Forms Recognition.

To open a document in the Workdoc Browser:

1) Select the Options Menu.
2) Click on open with Oracle Forms Recognition Workdoc Browser.

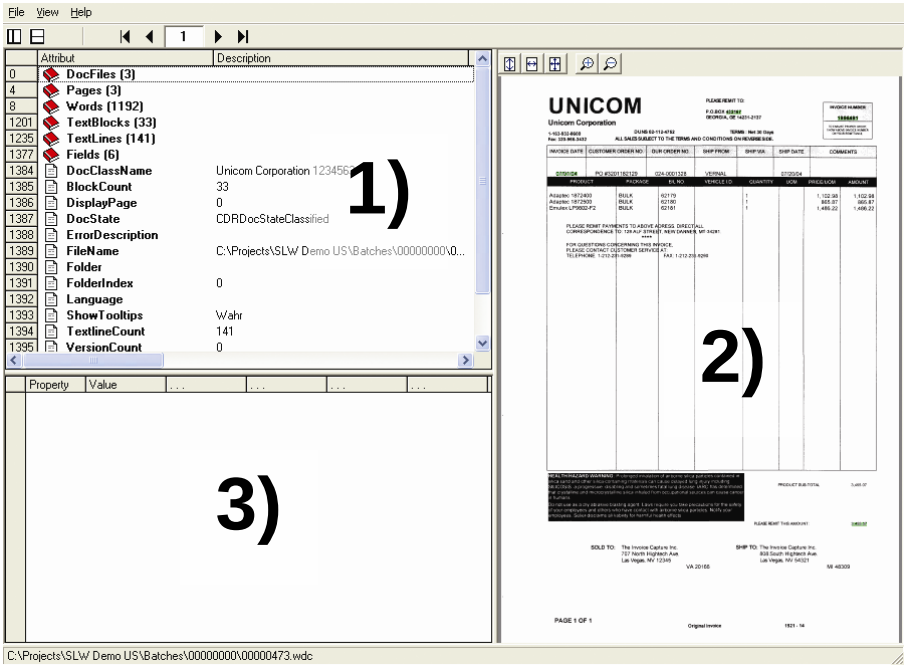


Figure 10-7: Workdoc Browser User Interface

1) Attribute Viewer Enables you to choose an attribute to view.
2) The Document Viewer. Displays the currently selected Workdoc.
3) The Property Viewer. Shows the properties of the selected attribute.

Chapter 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 ***7.3***.

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***) 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?

No 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 the "keep focus" toolbar button: Alternatively, you could just use a different magnification ratio.

I want to start the Learn Set Manager, but I don't want to have to go through Verifier first. Can I do it?

In a word, no. The Learn Set Manager is an add-in that can only be started in Verifier.

OK, I tried to start the Learn Set Manager in Verifier, and I still can't do it. Why?

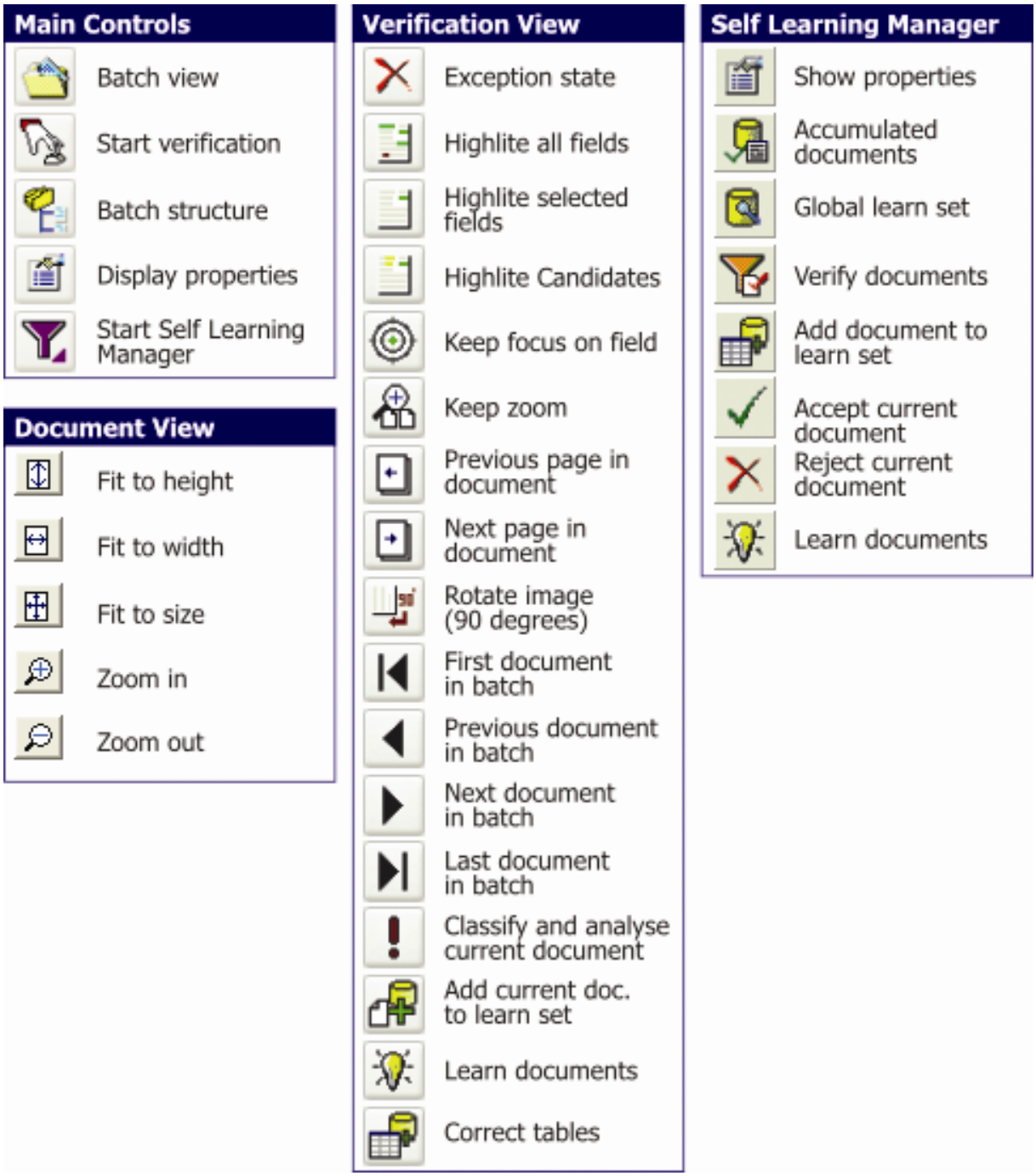
There are three main reasons:

- You may not have permission to use this add-in. Check with your project administrator to see if you are assigned to a group that can work with the Learn Set Manager mode.
- Learn Set Manager might not be enabled for the project. Again, contact your project administrator.

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

Oracle Forms Recognition 10.1.3.5.0 or higher supports a set of new 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. Please refer to the Scripting Guide documentation for more details.

Appendix A Quick Reference



Appendix B Index of Figures and Tables

Figures

Tables

Glossary

Accumulative Learn Set	The Common Learn Set.
Administrator	In Oracle 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.
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 Learn Set	An accumulation of Local Learn Sets.
DocClass	A parent document class.
Document	Any electronic file mainly consisting of ASCII text. If this 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 Oracle 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 Learn Set	A general Learn Set that encompasses similar classes or projects. See also Local Learn Set.
Importing	Bringing documents into Oracle 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.
Learn Set	In classification, a Learn Set 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 Learn Set is a set of documents whose field contents are selected by the user from a set of candidates.
Learn Set Manager	A user who designs, modifies, and maintains Learn Sets.
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 Learn Set	Learn Set 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 Learn Set) 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 Oracle Forms Recognition applications. They are created in Oracle Forms Recognition Verifier and handed over to Oracle 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	Oracle 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.
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.

Index