Perceptive Intelligent Capture

with Supervised Learning

Verifier User's Guide

Version 5.5 SP3

perceptivesoftware

Written by: Product Documentation, QA
Date: March 2014

Contents

Chapter 1 About Verifier and Advanced Verifier

1.1 Overview

Perceptive Intelligent Capture is a product suite designed for automatically processing incoming documents. Perceptive Intelligent Capture can process documents from arbitrary physical sources and paper-based documents as well as from electronic files, e-mail, or faxes.

What happens when Perceptive Intelligent Capture Verifier processes documents?

Structured or unstructured document input is obtained by scanning paper-based documents from fax servers, e-mail servers or as files. All documents are stored on a computer's hard drive. Perceptive Intelligent Capture monitors specified directories on this hard drive for new documents. If new documents are detected, Perceptive Intelligent Capture 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, Perceptive Intelligent Capture actually learns to tell which category a previously unknown document belongs.

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. Perceptive Intelligent Capture identifies and extracts data that is relevant for the respective document category. 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 Perceptive Intelligent Capture and written to designated export directories. The documents are then 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.

All this is done without human intervention once the Perceptive Intelligent Capture application has been set up. But what happens if Perceptive Intelligent Capture cannot properly process a document? There are several reasons this could happen:

- Paper-based documents might be "dirty" so that Perceptive Intelligent Capture is not able to read them.
- There might be stamps or notes on the documents that make important sections illegible for Perceptive Intelligent Capture.
- Perceptive Intelligent Capture 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.
- Perceptive Intelligent Capture may have been told to extract information that is missing, such as a form that was not filled in correctly.

That's where Perceptive Intelligent Capture Verifier comes in: Perceptive Intelligent Capture Verifier is the quality assurance utility of the Perceptive Intelligent Capture 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 Perceptive Intelligent Capture process. Therefore, subsequent systems will only receive appropriate input.

The Perceptive Intelligent Capture database platform for Perceptive Intelligent Capture applications enables you to keep a central store of your project and authentication information. This solution also allows for central management of storage and backup and thus provides for easier security, better connectivity of your applications, and higher flexibility for your personnel.

This version of Perceptive Intelligent Capture also brings an advanced GUI.

Key Features

- Allows central Perceptive Intelligent Capture database storage of projects and of user authentication information for more flexible access.
- 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 Learnsets. Advanced Verifier requires separate licensing and may not be available at your organization.

If you've worked with Perceptive Intelligent Capture before you already know that you can work with both data and text fields. Perceptive Intelligent Capture 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. Perceptive Intelligent Capture with Supervised Learning changes that by enabling you to interact with (or supervise) how Perceptive Intelligent Capture works with your data.

Chapter 2 About This Manual

2.1 Intended Audience

This manual is primarily for users of Perceptive Intelligent Capture 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 Perceptive Intelligent Capture Verifier and Advanced Verifier.
- CHAPTER 4 describes how to start and exit Perceptive Intelligent Capture Verifier.
- CHAPTER 7 explains the user interface of Perceptive Intelligent Capture Verifier.
- CHAPTER 6 describes how to configure Perceptive Intelligent Capture 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 Perceptive Intelligent Capture Verifier and Advanced Verifier for quality assurance in automatically processed documents.
- CHAPTER 9 explains how to train table data using traditional methods and with Brainware Table Extraction.
- CHAPTER 10 discusses the use of the Learnset Manager.

2.3 Related Documentation

In addition to this manual, Perceptive Intelligent Capture comes with the following documentation:

- ***Perceptive Intelligent Capture Web Verifier Users Guide***: Explains how to use the new web based extension of the Verifier.
- ***Perceptive Intelligent Capture Installation Guide***: Explains how to install Perceptive Intelligent Capture and how to set up the licensing file.
- ***Perceptive Intelligent Capture Designer User's Guide***: Explains how to use the Perceptive Intelligent Capture Designer component to create custom application.
- ***Perceptive Intelligent Capture Runtime Server User's Guide***: Explains how to use the Perceptive Intelligent Capture batch processing application.
- ***Perceptive Intelligent Capture Scripting Guide***: Explains how to write scripts to enhance Perceptive Intelligent Capture projects.
- ***Release Notes***
- ***Product Licensing Guide***
- ***Migration Guide***

In addition, a variety of white papers, solution descriptions, product bulletins, and case studies about Perceptive Intelligent Capture and other Perceptive products are available on Perceptive's Web site, www.perceptivesoftware.com

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 Perceptive Intelligent Capture, 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. Perceptive Intelligent Capture does the same thing with folders however in the context of Perceptive Intelligent Capture, 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, situation, or business transaction. For example, a packing slip may provide evidence that an order has actually been shipped. People are used to working with paper therefore electronic documents strongly resemble paper-based documents. You will notice that Perceptive Intelligent Capture 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 is 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 Perceptive Intelligent Capture, 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?

Extraction is a means for automatic document indexing. Take the stack of invoices again and electronically extract the name of the supplier, the total sum to be paid, and the invoice number. This procedure is called extraction. 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 Perceptive Intelligent Capture, states are determined hierarchically from the bottom up: If anything is wrong with a document, then there is also something wrong with the batch it belongs to.

3.1.8. What is verification?

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

3.1.9. What is validation?

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

3.1.10. What is a Learnset?

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

3.2 Perceptive Intelligent Capture and Quality Assurance

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

Note: Batches are the basic entity Perceptive Intelligent Capture is working on processing.

Perceptive Intelligent Capture works on processing 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.

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.

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

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

Similarly, a Perceptive Intelligent Capture 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, Perceptive Intelligent Capture 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 Perceptive Intelligent Capture 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 Perceptive Intelligent Capture and Teamwork

In Perceptive Intelligent Capture the flow of incoming documents follows a sequence of standard processing steps. Some steps can be skipped but the order of the steps is fixed.

Automatic steps are executed by Perceptive Intelligent Capture Runtime and include document import with batch creation, OCR and layout analysis, classification, extraction, export, and clean-up (This step deletes files that are no longer required from the hard drive.). 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 Perceptive Intelligent Capture 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 Perceptive Intelligent Capture 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 Perceptive Intelligent Capture 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 are carried out. Finally, when Perceptive Intelligent Capture has finished processing a batch, the documents will be sent to their actual recipients.

One of the objectives of a Perceptive Intelligent Capture 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, Perceptive Intelligent Capture 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 Perceptive Intelligent Capture Verifier accordingly. (See section 6.2 SETTINGS – WORKFLOW) In addition, several operators can carry out the same task at the same time, but on different batches. This is through 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 Perceptive Intelligent Capture 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 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 Perceptive Intelligent Capture 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. Perceptive Intelligent Capture Verifier sets a corresponding marker in the worklist.

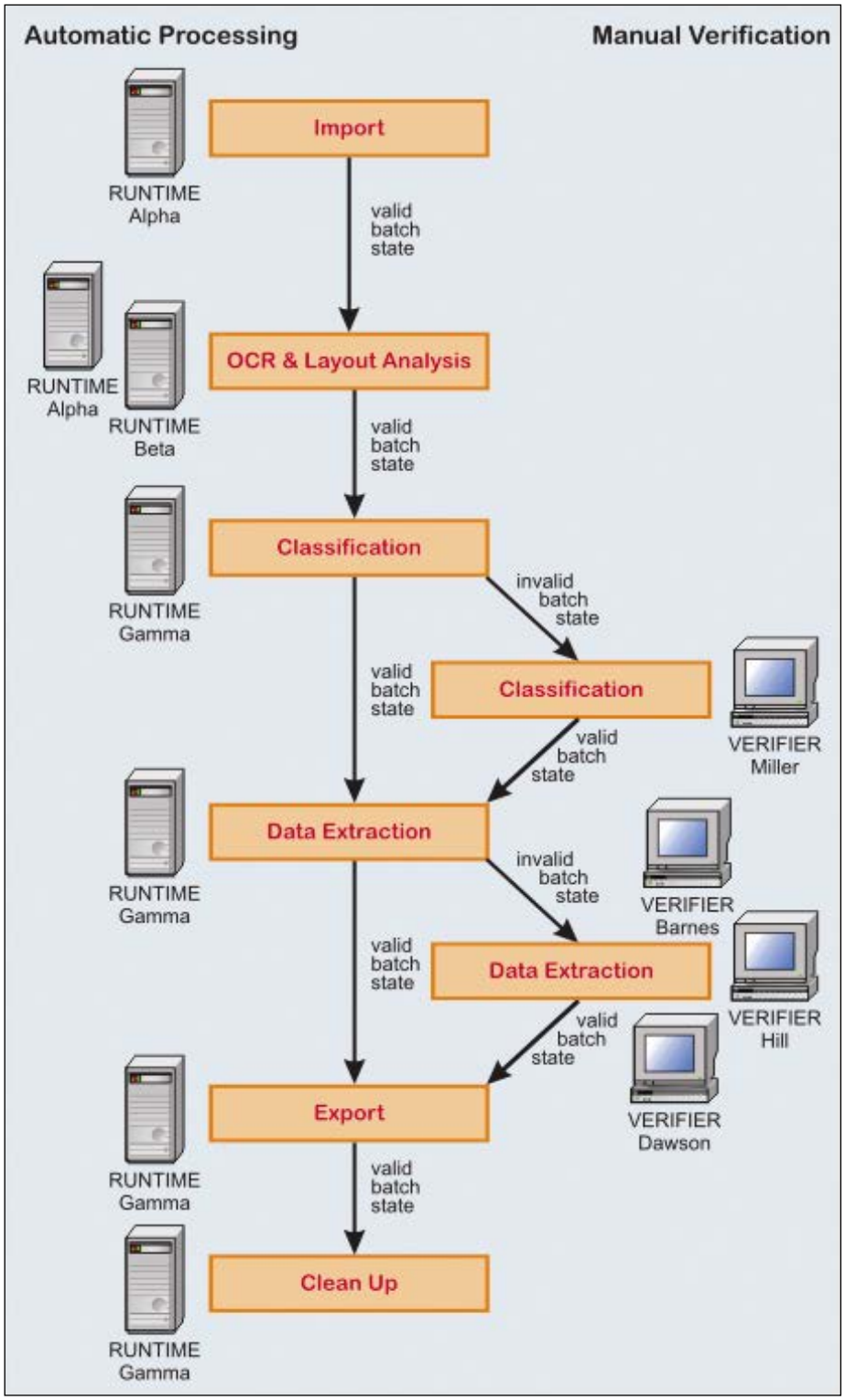


Figure 3-1: Sample workflow showing how processing steps can be distributed. This configuration involves several servers running Perceptive Intelligent Capture Runtime, and a number of workstations with Perceptive Intelligent Capture Verifier that are used by a quality assurance team.

Chapter 4 Starting and Exiting 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* - Perceptive Intelligent Capture - Perceptive Intelligent Capture *Verifier*. After startup and login, the application displays the Batch View (See CHAPTER 7 GETTING FAMILIAR WITH THE USER INTERFACE).

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

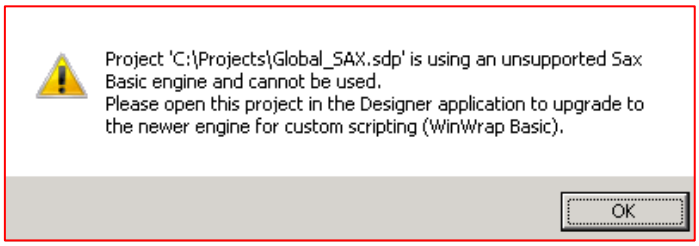


Figure 4-1: Error Message for a Sax Basic project

Click OK. This will close the Verifier application.

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

4.2 Login procedure

Note: When launching Verifier for the first time the application is not yet configured. The Batch View will be empty with an error message displayed. Perceptive Intelligent Capture Verifier needs to be configured. This should be done by an experienced user. For instructions, please refer to *CHAPTER 6 CONFIGURING VERIFIER AND ADVANCED VERIFIER*.

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.

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

Your project administrator can give you the option to remember your username and password between logons. This has been enabled if the *Remember 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 when you logon to the same computer, the system will fill in the username and password automatically so that simply clicking *OK* will log you in.

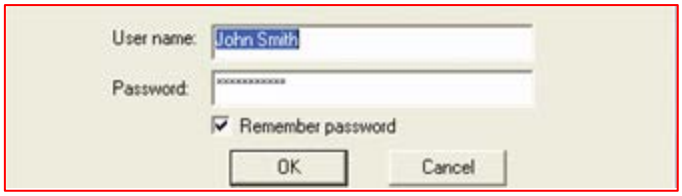


Figure 4-2: Verifier logon with password remember feature

For details about user accounts and roles, see CHAPTER 5 ABOUT USERS, GROUPS, AND ROLES.

4.3 Ability to Specify Logon Information via Command Line Arguments

To suppress project authentication when starting Perceptive Intelligent Capture Verifier, logon information can be 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 Perceptive Intelligent Capture program folder will launch Perceptive Intelligent Capture 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\Perceptive\Perceptive Intelligent Capture\DstVer.exe" /USR "John Smith" /PWD john1234567
```

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

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

Administrator can also review who has logged into the application by entering certain script. See **Scripting Guide** for detail.

4.4 Exiting the Application

To quit Perceptive Intelligent Capture 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, Learnset Manager, and Supervised Learning Verifier.

The roles of these groups follow:

- **Administrator:** The Administrator's 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.
- **Learnset Manager:** The Learnset Manager's role is to define, modify, and maintain the Learnset.
- **Supervised Learning Verifier:** The Supervised Learning Verifier's role is to collect and manage local training data. Supervised Learning Verifiers are subject-matter experts who can propose Learnset candidates to improve system performance.
- **Verifier:** The role of the Verifier group is to verify documents that could not be automatically processed.
- **Verifier Settings:** The role of the Verifier Settings group is to allow the Perceptive Intelligent Capture Verifier configuration to be changed.
- **Verifier Filtering:** The FLT role is to allow Verifier user to configure custom filtering of batches. By application design, FLT users would be able to use the filtering feature even if they do not have the SET role. This solution provides more flexibility and security.

5.1 Changing your Password

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

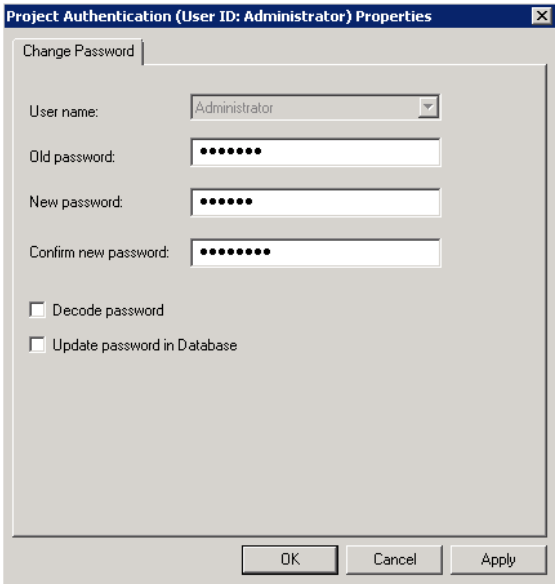


Figure 5-1: Password change dialog

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

Decodes password: Allows you to see the password you typed in when changing password.

- If *Decode Password* is unchecked, when you type the password in the Designer *Change Password* dialog box, it is masked with ****
- If *Decode Password* is checked, when you type the password in the Designer *Change Password* dialog box, is shows letters typed

Update password in Database: This checkbox is only available if you have enabled the *Allow Database Authentication* option on the *Users* tab in Designer's Authentication Settings.

Chapter 6 Configuring Verifier and Advanced Verifier

You can only change the Perceptive Intelligent Capture Verifier settings if you have been assigned the Verifier Settings role.

Configuring Perceptive Intelligent Capture Verifier entails specifying which batches of documents will be processed at a given station. This includes:

- Sourcing of the batches either from the file system or from the Perceptive Intelligent Capture database.
- The location of the batches in the file system.
- The Perceptive Intelligent Capture 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).

Note: You can only work with Perceptive Intelligent Capture Verifier after these settings are established. Only experienced users should change the settings.

To configure Perceptive Intelligent Capture Verifier either:

Select *Settings* on the *Options* menu, or click *Settings* in the toolbar. This will display the Perceptive Intelligent Capture *Verifier Properties*.

6.1 Settings – General

The General tab is the place for general settings. It allows you to configure your referenced directories and files. Also, you can choose the Perceptive Intelligent Capture database as your document and statistics source here.

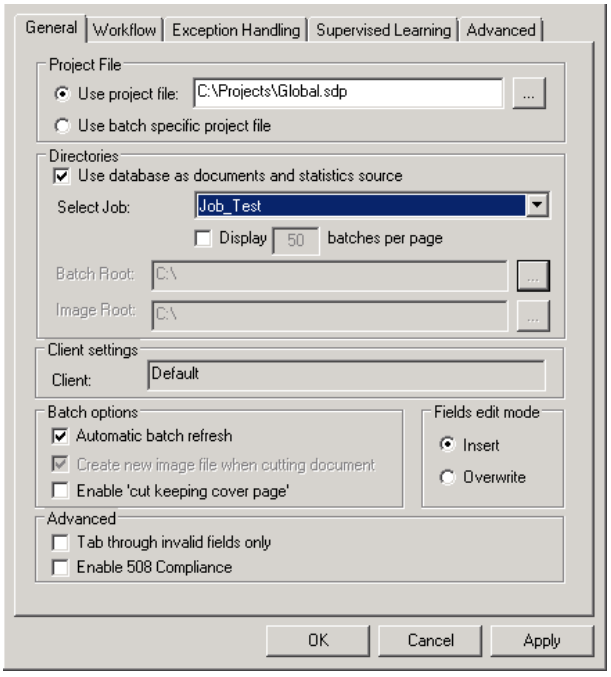


Figure 6-1: General Tab

Project File:

- *Use Project File*: Used to select the path and file name of the Perceptive Intelligent Capture application that processes 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 Perceptive Intelligent Capture Runtime's drive allocation.

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

  Let's assume that C:\Perceptive\ on system_srv is mapped to drive E:\ on system_qa. This means that…

  C:\Perceptive\Projects\SampleCorporation\SampleCorp.sdp on system**srv** and E:\Projects\SampleCorporation\SampleCorp.sdp on system_qa …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 Perceptive Intelligent Capture 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:

- *Use Database as Document and Statistics Source*: Perceptive Intelligent Capture core information can be stored in the Perceptive Intelligent Capture database. Furthermore, you are able to select the desired job from the *Select Job* dropdown list if you have selected the Database as your source.
  Note: The file system functionality is still supported.
  Display … batches per page: This option enables you to set the number of batches displayed per page. It allows values from 1 to 200 batches to be displayed per page. This is only available if *Use Database as documents and statistics source* is selected. The value of 50 is set by default.
- *Batch Root*: This is the directory where the batch control files are located. This directory also includes a license directory where you can find the network license file. When multiple instances of Runtime Server and Verifier are running they will check this directory for the network license file. If you use different batch root directories simultaneously, you can set up a specific path for the network license. However, you can only configure this path manually for Runtime Server and Verifier. This option is not available when using the Perceptive Intelligent Capture database.
  Note: When using Database, batch root directory licensing cannot be used, and a license share is required.
- *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 intent to use client-specific variables. In Perceptive Intelligent Capture 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.
  Note: This feature is disabled when the Use Database option is enabled. Any project that wants to use this feature is recommended to use the File System instead of Database.
- 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 additional menu entries. 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 column items are indicated by a blue rectangle without icons while valid / invalid column items are indicated by rectangles with a “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

Note: 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.

Perceptive Intelligent Capture 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 a Perceptive Intelligent Capture Verifier who corrects the errors and supplies missing results. Therefore, you need to know the output states used by your Runtime Server installation as they determine the input states used by the Perceptive Intelligent Capture Verifier stations.

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

Example

Note: Let’s assume that Runtime Server uses the input and output states depicted below to process batches.

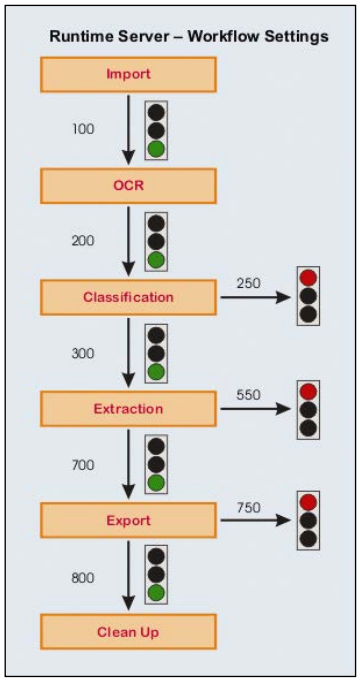


Figure 6-2: Sample Runtime server workflow configuration

In this case, a Perceptive Intelligent Capture Verifier station conducting both classification and extraction verification should use the settings shown to verify whether this is correct.

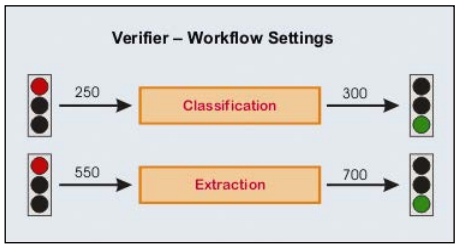


Figure 6-3: Verifier workflow Configuration

This produces the following combined workflow:

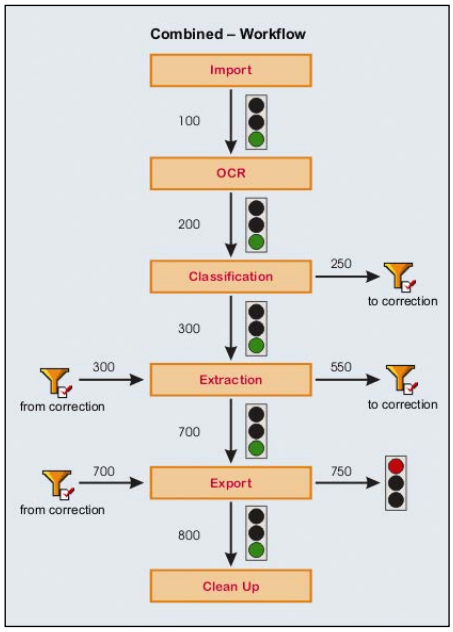


Figure 6-4: Combined workflow resulting from matching Runtime server and Verifier workflow configurations above.

Normally, the workflow should be set up like this:

- Manual indexing / extraction correction after manual classification

Note: If a document could not be classified, automatic extraction probably will not work.

- 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. There is the option however 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 Perceptive Intelligent Capture Verifier station, select the *Workflow* tab.

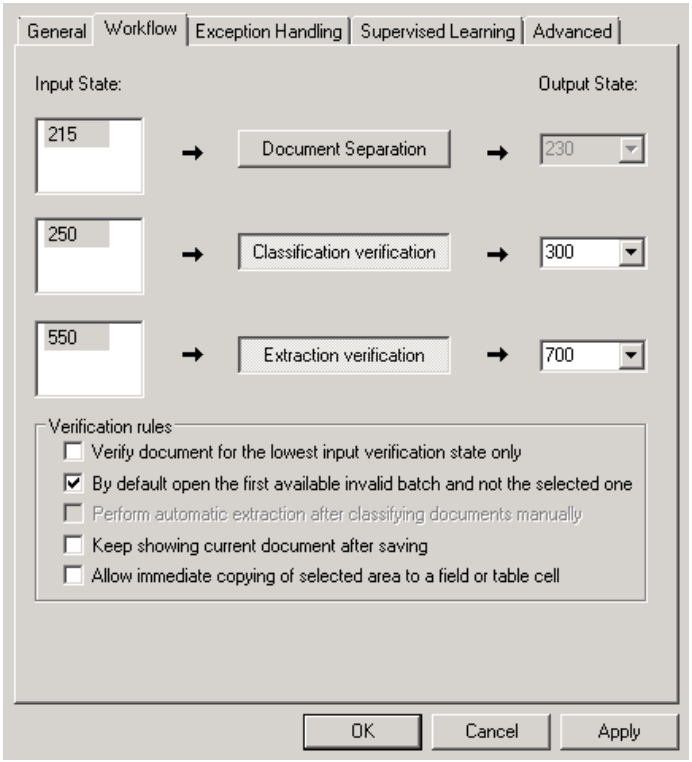


Figure 6-5: Workflow tab

- To configure document separation, click *Document Separation*.
- To configure classification verification at this workstation, click *Classification Verification*. Likewise, you can configure *Extraction verification* by clicking on it. One or both steps can be performed at the workstation.
- After you have selected the steps to perform, establish values for input and output states. To add an input value, right-click on the *Input* option box and select *Add State* on the shortcut menu. (You can also change states and delete states this way.) To set an output value, select it from the dropdown box to the right of the classification or extraction button.
- Next, configure the 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 form 1 and afterwards form 2 for Document 1 is 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)

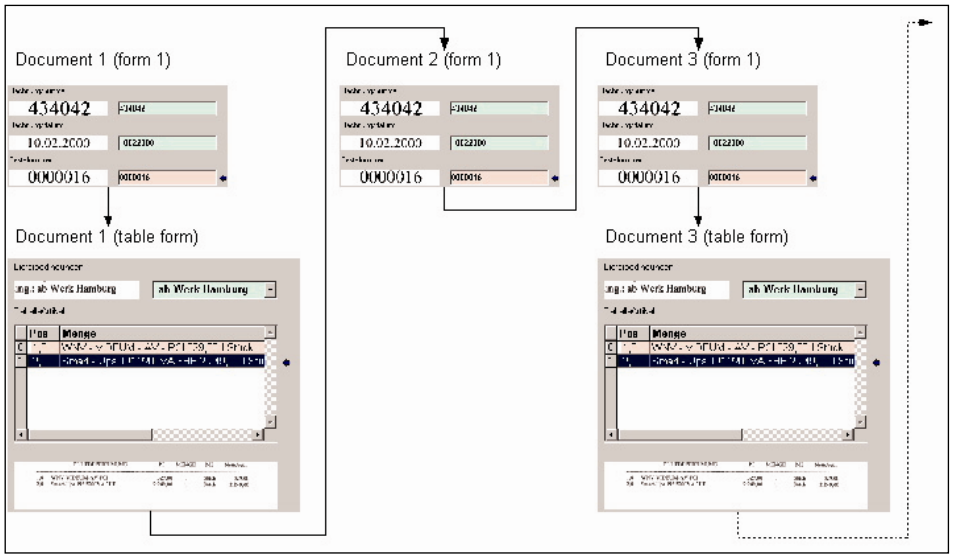


Figure 6-6: Default form sequence for extraction (check box not selected)

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 which form is shown for which input state. (FIGURE 6-8)

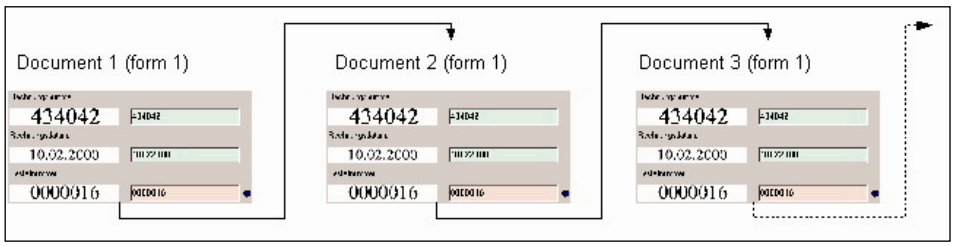


Figure 6-7: Sequence of verification forms depending on input states

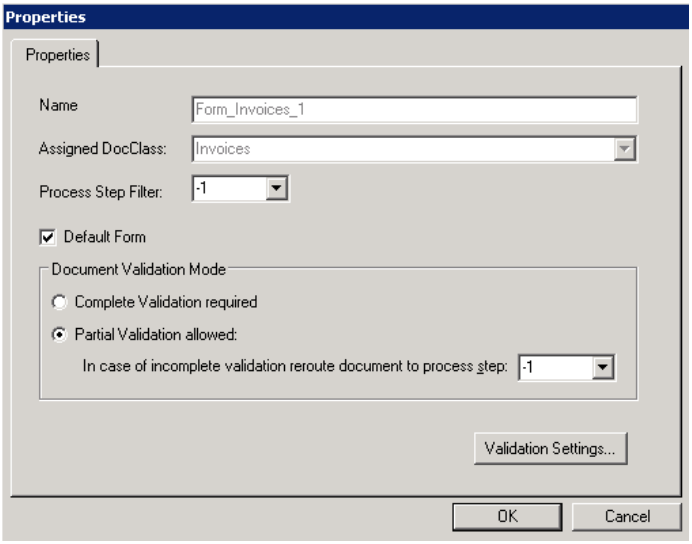


Figure 6-8: Perceptive Intelligent Capture Designer - Verifier Design Mode Verification Form Settings

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) which 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.

- *By default open the first available invalid batch and not the selected one:* By default, the first available invalid batch is opened and not the selected one. The first invalid batch is selected based on Priority (higher first), user's custom filter, sort settings and the Batch ID. This is for projects with large amount of batches and simultaneous Verifier users. It decreases time delay of project verification. This option is selected by default. And under Verification Mode, option "Verify first invalid batch" is also selected by default.
  If the user changed the option from the Verification Mode, this option will only remain valid for the current application session. As soon as the user quits the application, the default behavior will be re-initialized via the above mentioned "[x] By default open the first available invalid batch and not the selected one" settings option.
- *Perform automatic extraction after classifying documents manually:* Forces Perceptive Intelligent Capture 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 Perceptive Intelligent Capture 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.

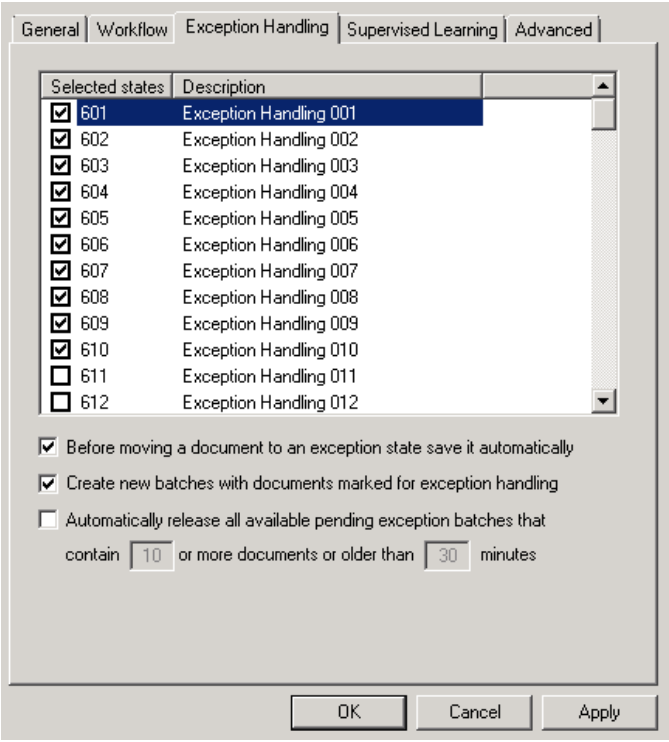


Figure 6-9: Exception Handling tab

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 new label name into the corresponding field and confirm. The *'New Description'* field allows 128 characters.

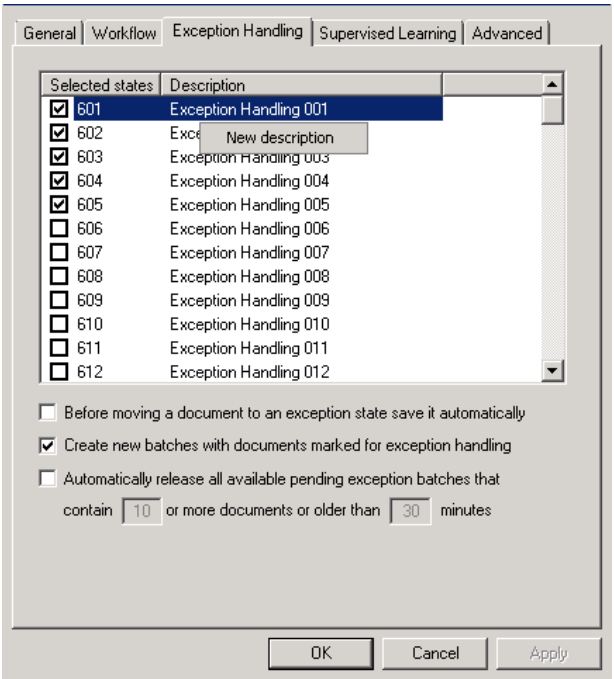


Figure 6-10: Editing Description

6.3.3. Additional Settings for Exception Handling

- *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 and for each user. Documents from all verified batches are moved to the same user specific "exception batches" in the *Batch View*. These exception batches receive new batch ID`s. Release your exception batches manually by pressing [Ctrl] + [E] (see section 7.1.1 MENU COMMANDS AND KEYBOARD SHORTCUTS - BATCH VIEW, *Table 8-3*) or automatically (see option below). Your exception batches will be active and locked for other users until you log out or release the exception batches.
- *When the option Create new batches … option is deselected the document marked for exception handling remains in its batch. This batch turns into an exception batch. The batch ID remains unchanged.*
- *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. Your exception batches will also be released manually by pressing [Ctrl] + [E] or when you exit the application.

6.4 Settings – Supervised Learning (Advanced Verifier)

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

Steps to Configure Supervised Learning in Verifier (With Perceptive Intelligent Capture **database):**

1. Open the *Settings* dialog (from *Verifier Options* menu>*Settings* or using the *Settings* icon) and select the *General* Tab.

2. Select the *Project* file, *Batch Root* and *Image Root* path along with the *Database* checkbox. (*Project* and *Batch* root need to be configured with a database job in RTS).
3. Set the *Input* states for various workflow stages appropriately in *Workflow* tab.
4. Go to the *Supervised Learning* Tab and check the option for *Activate Supervised Learning Workflow*.
5. Select the *Local Project* from *Verifier* workstation directory. (Local project must be configured with its own base directory (*Local Learnset*) and *Batch* root)
6. Select the *Knowledge Base directory* (Common Learnset folder).
7. Check the option *Use Database as knowledge source*.

   Note: To use the option "Use database as knowledge source" you need to have a job for a Common Learnset in the database. If this Common Learnset job is not available, a database 'job' must be created from RTS with Common Learnset folder as the Batch Root.
8. Check the option *Distribute Local Learnset to Knowledge base* only if you want to push the Local Learnset to Common Learnset. (This is required to verify the documents in Learnset Manager (LSM))
9. Select the option *Apply local classification and extraction automatically.*

Now the settings can be saved, and the SLW user should be able to perform classification and automatic extraction in Verifier.

An 'SLM' user can now launch LSM to verify the Learnsets from the *Common Learnset* in the *Accumulated Documents Browsing* mode.

Notes:

- *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.
  - *Knowledge Base Directory*: The file and pathname of the Common Learnset. The Common Learnset will be updated whenever the Local Learnset is migrated to it.
  - *Distribute Local Learnset to the knowledge base*: Automatically adds any documents added to the Local Learnset into a queue for the Learn Set Manager to review if the documents are appropriate to be added to the Global Learnset. The knowledge base is often referred to as a queue of accumulated documents or Common Learnset pending review by the Learnset Manager for improvements into the project file.
  - *Nominate for the Learnset but never train locally*: This option enables you to prevent the Learnset from being trained locally.
  - *Use database as knowledge source*: Here you are able to select the desired job from the dropdown list. This drop down list shows all batch jobs in the database, when the *Use Database* option is selected.
  - *Always show state of all field locations after opening a document*: not available in this version.
  - *Apply local classification and extraction automatically*: New classes will be created using the supplier's name. A Learnset should also be created if you select this setting. When no local project/ Learnset are used, Global Project and Global Learnset will be used instead.

  - *Prompt if script forces or rejects insertion to Learnset:* Will notify you if there is a discrepancy between script and your commands regarding the population of the Learnset.
- *Put document to Local Learnset*:
  - *Only if adding activated by a user.* If **selected,** a document will be added to a Learnset 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.* If **cleared,** a document will be added to the Learnset 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.* If **selected,** documents will automatically be added to the Learnset 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.* If **cleared,** documents will automatically be added to the Learnset if the threshold you set is exceeded. No confirmation will occur.
    Always prompt before adding option is not available.
- *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 Learnset.

6.5 Settings – Advanced

The Advanced tab is used for additional features.

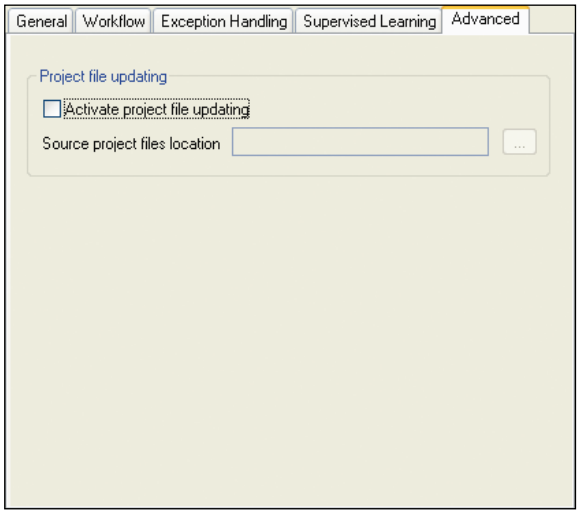


Figure 6-11: Advanced Settings tab

- *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.

6.6 Batch Filter

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

The filter tab is located outside the settings dialog so that users without a SET role but with the FLT role are able to filter batches. Only users having the FLT role assigned (managed in Designer by the Project Administrator) will be able to configure filter conditions.

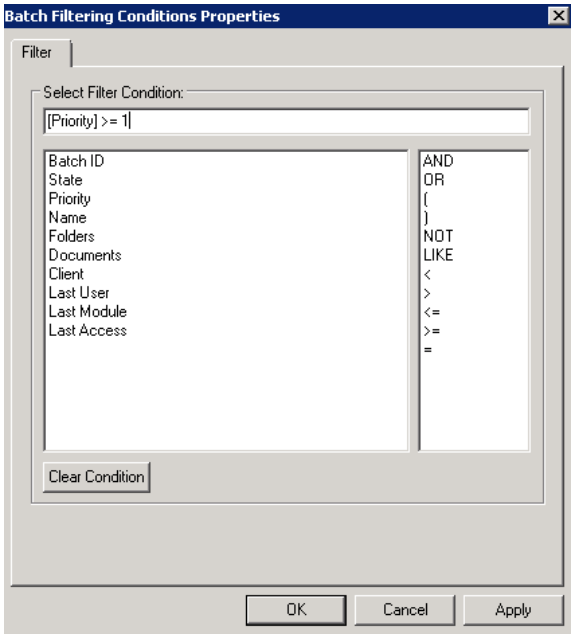


Figure 6-12: Batch Filtering Properties

To configure your filter settings:

- Select *Filtering* from the list under Options, or click on the *Batch filter* icon.
- Double click on an entry in the left pane to select a batch attribute, and then double click on a filter condition in the right pane. E.g. in the example above, the batches with *Priority* equal or greater than 1 would be displayed.

Use the following formats to specify the filter conditions:
Batch ID -> string, numeric: see note below
State -> numeric
Priority -> numeric
Name -> string
Folders -> numeric
Documents -> numeric
Client -> string
Last User -> string
Last Module -> string
Last Access -> date

- Click *Clear Condition* to clear the filter condition setting.

Note: All string related operations (LIKE, '%', etc) can be used with [Batch ID] value. Use numeric queries only with numeric Batch IDs. When batches were imported from the filesystem, then their IDs have a string format. If the batches were imported only with a RTS import instance and the IDs were not changed through script, then the IDs are numeric.

Chapter 7 Getting Familiar with the User Interface

7.1 The Batch View

The first window displayed after starting Perceptive Intelligent Capture 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: 

Note: The list of batches will be empty if Perceptive Intelligent Capture Verifier has not yet been configured. (See CHAPTER 6 - CONFIGURING VERIFIER AND ADVANCED VERIFIER for details)

The *Batch View* looks like this:

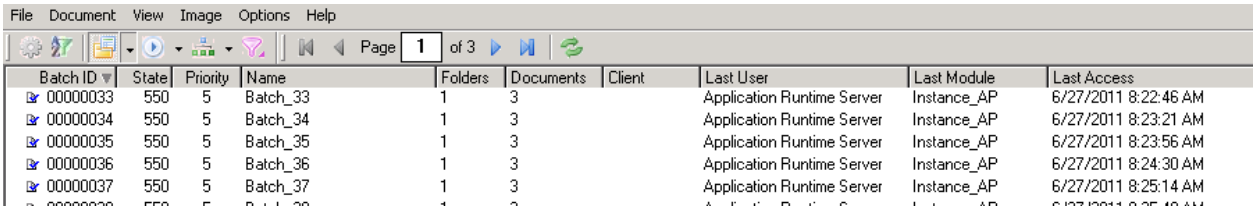


Figure 7-1: Batch View

Menu Bar (See section 7.1.1 MENU COMMANDS AND KEYBOARD SHORTCUTS - BATCH VIEW)

Toolbar (See section 7.1.2 TOOLBAR BUTTONS)

Table of Batches (See section 7.1.4 TABLE OF BATCHES)

7.1.1. Menu Commands and Keyboard Shortcuts - Batch View

The following commands can be accessed from the menu bar:

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.
	Print Setup	n/a	n/a	Adjusts print output
	Print	n/a	n/a	Not available in Batch View
	Exit	n/a	n/a	Quits Perceptive Intelligent Capture Verifier.

Table 7-1: 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	Displays the status bar at the bottom of the Batch View.
	Batch List View	n/a	[Ctrl] + [1]	Displays the Batch View. (See section THE BATCH VIEW)
	Verification Mode	n/a	[Ctrl] + [2]	Starts the verification of the selected, next or first invalid batch. Depending on the status of the batch, the Verification View is either displayed in Classification Mode (See section THE VERIFICATION VIEW - CLASSIFICATION WINDOW) or in Indexing Mode. (See section THE VERIFICATION VIEW - INDEXING WINDOW)
		Verify first invalid batch	n/a	To verify the first invalid batch.
		Verify selected/ next batch	n/a	To verify the selected or the next batch.
	Document Separation Mode	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, Document Separation Only	n/a	Only batches waiting for document separation.
		Batches to Verify, Classification Only	n/a	Only batches waiting for classification verification.
		Batches to Verify, Indexing Only	n/a	Only batches waiting for indexing verification.
		Batches to Verify	n/a	Only batches waiting for verification.
	Document Filter	n/a	n/a	n/a
	Restore Focus		[Ctrl] + [N]	Manually restores input focus without using the mouse
	Refresh	n/a	[F5]	Checks the directory that contains the batches for

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
				changes and updates the table of batches accordingly.
	Multi-Page view	n/a		Not available in Batch View

Table 7-2: Batch view menu commands (Document & View)

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
Image				
	No commands are available in the Batch View context.			
Options				
	Settings	n/a		Displays a dialog box where you can configure Perceptive Intelligent Capture Verifier. (See CHAPTER 6 CONFIGURING VERIFIER AND ADVANCED VERIFIER)
	Filtering	n/a		Displays a dialog box where you can configure the batch filtering conditions.
	License	n/a		The path to the license file can be specified here.
	Release Exception Batches	[Ctrl] + [E]		Releases your exception batches. This menu item is only activated, if any documents have already been moved to exception state.
	Learnset Manager	n/a	n/a	Starts the Learnset Manager (if licensing and permissions are configured.)
	Change Password	n/a	n/a	Displays a dialog box for project authentication properties where you can change your password.

Table 7-3: 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 Perceptive Intelligent Capture Verifier. (See CHAPTER 6)
	Displays a property dialog where you can configure the batch filtering conditions. (See 6.6 BATCH FILTER
	If you clicked on the arrow to the right of this button, the available filters for the list of batches will be displayed. You can select either: • All Batches • Batches to verify, Document Separation Only • Batches to Verify, Classification only, • Batches to Verify, Indexing only • Batches to Verify

Button	Description
	Two options are available: Verify first invalid batch or Verify selected/ next 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 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 Learnset Manager. (Chapter 10 Working with the Learnset Manager)

Table 7-4: Batch view controls

7.1.3. Navigation Toolbar Buttons

The navigation toolbar enables you to easily navigate through a big amount of batches. It is also possible to configure a number of batches to appear per batch page.

Button	Description
	Go to the first batch page
	Go to the previous batch page
	Go to the next batch page
	Go to the last batch page
	Refresh

7.1.4. 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.

Symbol	Description
	Batch is locked and displayed dimmed as it is in use of another application. Therefore it cannot be opened for correction.
	Batch contains documents with exception statuses. When it is dimmed, it needs to be released before you can work on it again.

Table 7-5: 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)

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 Perceptive Intelligent Capture 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.
- *Last Access:*
  Displays the date when the batch was last processed.
- *External Group ID:*

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

- *External Batch ID:*

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

- *Transaction ID:*

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

- *Transaction Type:*

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

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

7.1.4.1. Sorting and Navigating in the Batch View

You can sort any column in the Batch View. To sort any items, 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 in 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 view 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 *Show selected batch*: 

The *Document View* looks like this:

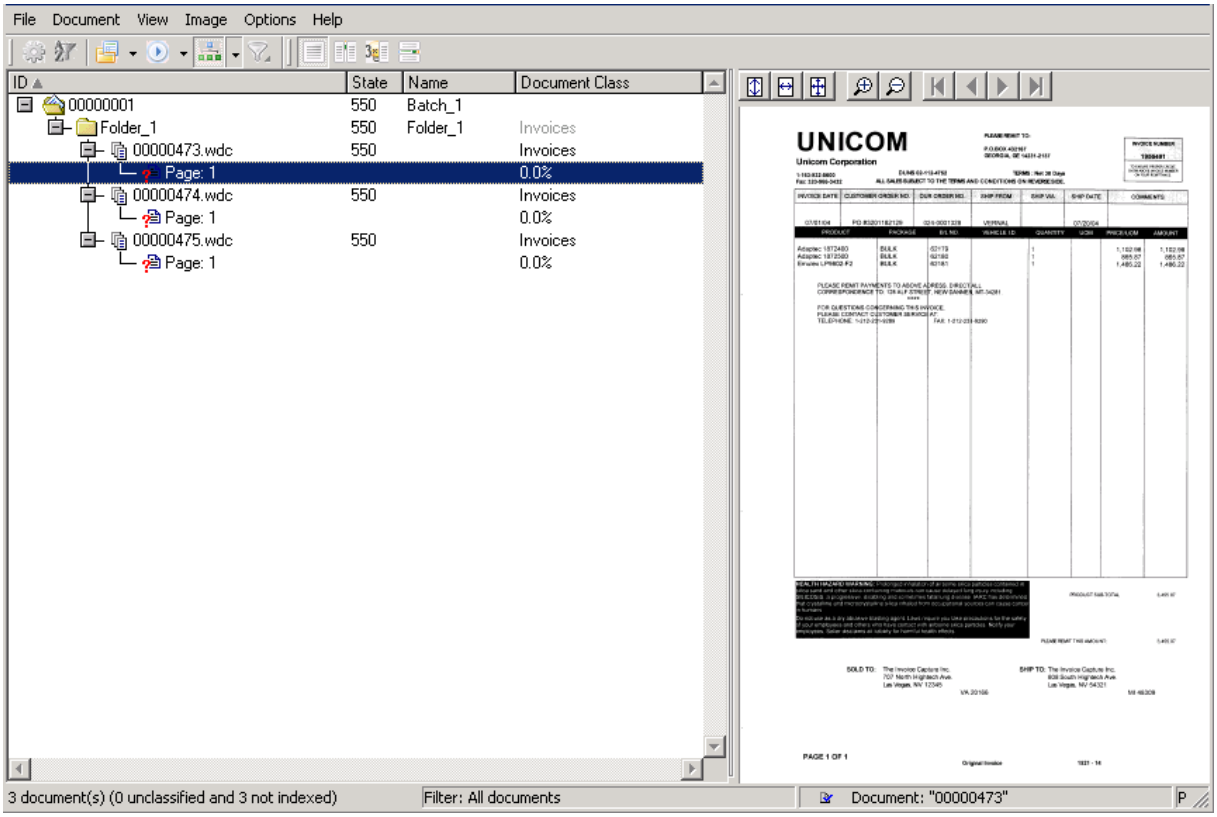


Figure 7-2: Document View

Menu Bar (See section 7.2.1 MENU COMMANDS AND KEYBOARD SHORTCUTS)

Main Toolbar (See section 7.2.2 MAIN TOOLBAR BUTTONS)

Batch Structure (See section 7.2.3 BATCH STRUCTURE AREA)

Viewer Toolbar (See section 7.2.4 VIEWER TOOLBAR BUTTONS)

Document (See section 7.2.5 DOCUMENT AREA)

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				
	Print Setup	n/a	n/a	Adjusts print output.
	Print	n/a	[Ctrl] + [P]	Starts printing dialog.
	Exit	n/a	n/a	Quits Perceptive Intelligent Capture 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.

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

Table 7-6: Document view menu commands (File & Document)

Note: The availability of the Document menu options depends on whether the document itself or one of its pages is selected:

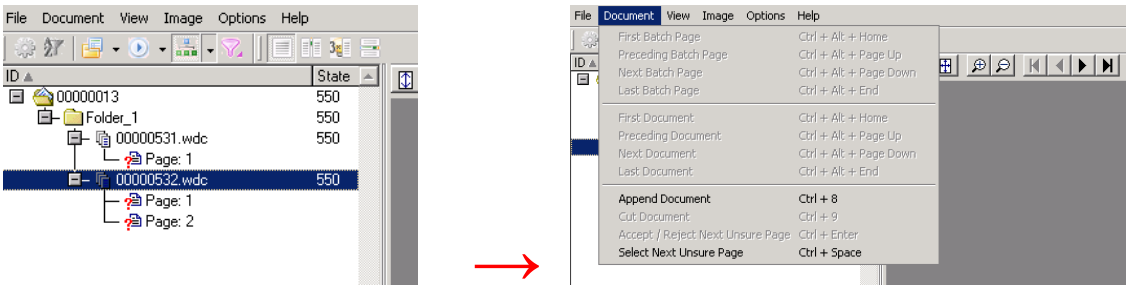


Figure 7-3: Document menu for the document selected

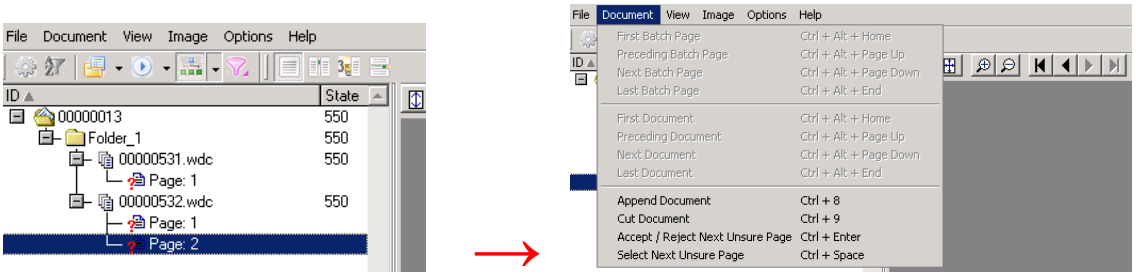


Figure 7-4: Document menu for the page selected

Menü	Submenu/ Command	Command	Keyboard Shortcut	Description
View				
	Toolbar	n/a	n/a	Switch that can be used to show or hide the toolbar.

Menü	Submenu/ Command	Command	Keyboard Shortcut	Description
	Status Bar	n/a	n/a	Displays status bar at the bottom of Document view page.
	Batch List View	n/a	[Ctrl] + [1]	Displays the Batch View. (See section THE BATCH VIEW) The check mark shows that this view is selected.
	Verification Mode	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 THE VERIFICATION VIEW - CLASSIFICATION WINDOW) or in Indexing Mode (See section THE VERIFICATION VIEW - INDEXING WINDOW)
	Document Separation Mode	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	No filter is applied. All of the documents are listed.
		Documents to classify or index	n/a	Shows documents that have to bbe classified or indexed.
		Documents to classify	n/a	Shows documents that require classification.
		Documents to index	n/a	Shows documents that require indexing.
	Restore Focus	n/a	[Ctrl] + [N]	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.
	Multi-Page View			Display options for multi-page documents.
		Single Page	n/a	Only one page is displayed. Use the navigation buttons to browse through further pages.
		Two Pages Horizontally	n/a	Two pages of the document are placed side by side.
		Three Pages	n/a	Three pages of the document are placed

Menü	Submenu/ Command	Command	Keyboard Shortcut	Description
		Horizontally		side by side.
		Two Pages Vertically	n/a	Two pages of the document are placed one below the other.

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

Menu	Submenu/ Command	Command	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. *Note that this setting will be kept for all further documents until next modification.*
	Show Selection Context Menu	n/a	[Ctrl] + [M]	Invokes the shortcut menu applicable for the selection.
	Fit to Height	n/a	n/a	Fits the document to the height of the viewer.
	Fit to Width	n/a	n/a	Fits the document to the width of the viewer.
	Best fit	n/a	n/a	Forces the document to fit into the viewer so all of it is displayed.

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
	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.
	Increase image area	n/a	[Ctrl] + [J]	Not available in Document View.
	Decrease image area	n/a	[Ctrl] + [K]	Not available in Document View.

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
Options				
	Undo	n/a	[Ctrl] + [Z]	Can be used in Verification Extraction View.
	Settings	n/a	n/a	Displays a property sheet where you can configure Perceptive Intelligent Capture Verifier. (See CHAPTER 6 CONFIGURING VERIFIER AND ADVANCED VERIFIER)
	Filtering	n/a	n/a	Displays a property dialog where you can configure the batch filtering conditions. Not available within Document view.
	License	n/a	n/a	Modify license connections: Use specified path or use License directory.
	Change client	n/a	n/a	Not supported in this application.
	Reclassify manually	n/a	[F7]	Option available in Advanced Verifier.
	Show last verified document	n/a	[F3]	Option available in Advanced Verifier.
	Get last value for selected field	n/a	[F8]	Option available in Advanced Verifier.
	Move document to exception state	n/a	[F9]	Moves current document to default exception state. Accessibility: Classification view, Indexing view.
	Release exception batches	n/a	[Ctrl]+ [E]	Releases all exception batches. This menu item is only activated, if any documents have already been moved to exception state.
	Apply local extraction	n/a	[Ctrl] + [L]	Option available in Advanced Verifier.
	Add document to Learnset	n/a	[Ctrl] + [A]	Option available in Advanced Verifier.
	Learn Documents	n/a	n/a	Option available in Advanced Verifier.

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
	Correct Tables	n/a	[Ctrl] + [T]	Option available in both Verifier and Advanced Verifier, but enhanced functionality only available in Advanced Verifier.
	Switch table highlighting	n/a	[Ctrl] + [Q]	Option available in Advanced Verifier.
	Learnset Manager	n/a	n/a	Starts the Learnset Manager.
	Move all exception documents	n/a	n/a	n/a
	Change Password	n/a	n/a	Allows you to change your password.

Table 7-9: 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 Perceptive Intelligent Capture Verifier. (See CHAPTER 6)
	Displays a property dialog where you can configure the batch filtering conditions.
	Displays the Batch View.
	Starts the verification of the currently selected batch. Depending on the batch state, the batch is either displayed in the classification window (Section 7.3 THE VERIFICATION VIEW - CLASSIFICATION WINDOW) or in the indexing window. (Section 7.4 THE VERIFICATION VIEW - INDEXING WINDOW) Clicking on the arrow calls a dropdownlist with the options *Verify selected/next batch* and *Verify next invalid batch*. This *Verify next invalid batch* option only works with database jobs.
	Displays the available filters for the batch structure. You can select from among the following options: • All documents • Documents to Classify or Index • Documents to Classify • Documents to Index
	Starts Learnset Manager (Available only if configured in Designer and properly licensed.) CHAPTER 10 WORKING WITH THE LEARNSET MANAGER
	Display the first page of the selected batch or the single page of the selected document.
	Display the first two pages of the selected batch horizontally or the first two pages of the selected document.
	Display the first three pages of the selected batch horizontally or the first three pages of the selected document.

Button	Description
	Display the first two pages of the selected batch vertically or the first two pages of the selected document.

Table 7-10: 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 of the first page which cannot be split. You can also merge consecutive documents into one with multiple pages. (See FIGURE 7-5 and FIGURE 7-6)

Splitting Multipage documents

To split a multipage document do the following:

1) Select *View, Show Selected Batch, All Documents* from the main menu, or click *Show Selected Batch, and then All Documents*.
2) In the document list, click on the desired multiple pages document.
3) Right click on the second page.
4) Select *Cut pages into a new document*. The document is now split into two documents. The second document will have the same file name as the first document, but will have an underline and 1 after the number or name. For example, a document initially called invoiceabc would now be invoiceabc and invoiceabc_1. Corresponding TIFFs will also be created.
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.).
   Note: This option will not available until you have marked a page as a cover page. You can do this by right clicking on the first page of the document and selecting Mark as cover page from the emerging menu.

Figure 7-5: Splitting pages into a new document

Appending Two Documents

To append two documents:

1) Select *View, Show Selected Batch, All Documents* from the main menu, or click *Show Selected Batch, and then All Documents*.
2) Select the document to append to the previous document.
3) Right click on the document.
4) Select *Append this document to previous one*. The document will now appear in the list as a multi-page document.

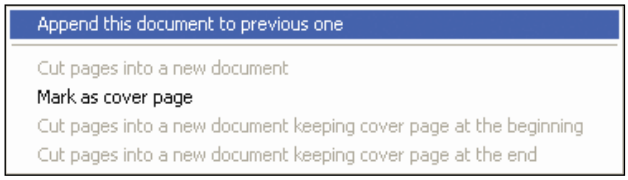


Figure 7-6: Appending pages to a document

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

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 7-11: 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.

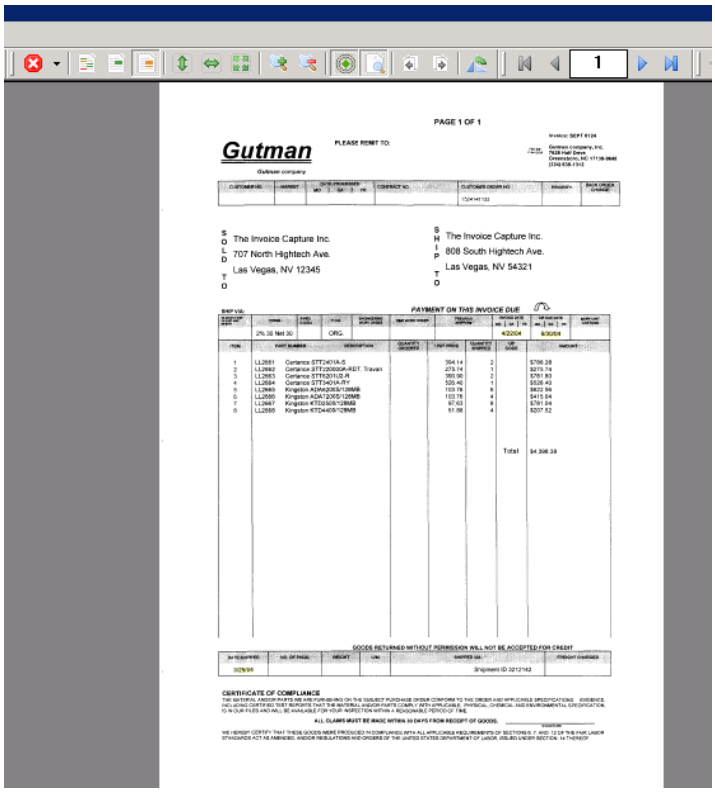


Figure 7-7: Document Area

It is also possible to default the Verifier to display a specific page of each document instead of the first one. Please refer to the Scripting Reference Guide, section on VerifierFormLoad event, for more information.

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 *Verify Selected Batch*.

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:

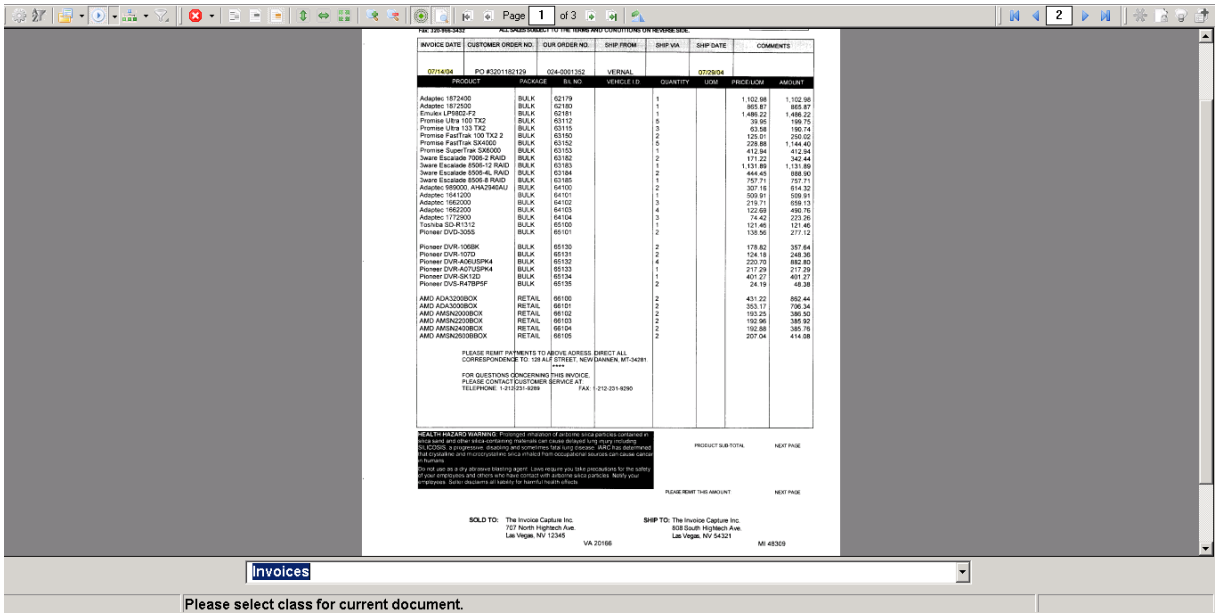


Figure 7-8: The Classification window

Menu Bar (See section 7.3.1 MENU COMMANDS AND KEYBOARD SHORTCUTS)

Toolbar (See section 7.3.2 TOOLBAR BUTTONS)

Document (See section 7.3.3 DOCUMENT AREA)

Class Selection (See section 7.3.4 CLASS SELECTION LIST)

7.3.1. Menu Commands and Keyboard Shortcuts

The following commands can be accessed from the menu bar:

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
File				
	Exit	n/a	n/a	Quits Perceptive Intelligent Capture Verifier.
Document				
	First Document	n/a	[Ctrl] + [Alt] + [Home]	Moves to the first document.
	Preceding	n/a	[Ctrl] + [Alt] +	Moves to the previous document.

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
	Document		[Page up]	
	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 7-12: Verification View - Classification 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 THE BATCH VIEW) 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 THE VERIFICATION VIEW - CLASSIFICATION WINDOW) or in Indexing Mode. (See section 7.4 THE VERIFICATION VIEW - INDEXING WINDOW)
	Show Selected Batch	n/a	[Ctrl] + [3]	Displays the selected batch in Document View.
	Highlight Mode	n/a	n/a	n/a
	Batch Filter	n/a	n/a	n/a
	Document Filter			Not available in Verification View
	Restore Focus	n/a	[Ctrl] + [N]	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.
	Multi-Page View			Display options for multi-page documents.
		Single Page	n/a	Only one page is displayed. Use the navigation buttons to browse through further pages.
		Two Pages Horizontally	n/a	Two pages of the document are placed side by side.

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
		Three Pages Horizontally	n/a	Three pages of the document are placed side by side.
		Two Pages Vertically	n/a	Two pages of the document are placed one below the other.

Table 7-13: Verification View - Classification Mode menu commands (View)

Menu	Submenu/ Command	Command	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. *Note that this setting will be kept for all further documents until next modification.*
	Show Selection Context Menu	n/a	[Ctrl] + [M]	Invokes the shortcut menu applicable for the selection.
	Fit to Height	n/a	n/a	Fits the document to the height of the viewer.

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
	Fit to Width	n/a	n/a	Fits the document to the width of the viewer.
	Best fit	n/a	n/a	Forces the document to fit into the viewer so all of it is displayed.
	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.
	Increase image area	n/a	[Ctrl] + [J]	Not available in Document View
	Decrease image area	n/a	[Ctrl] + [K]	Not available in Document View

Table 7-14: Verification View - Classification Mode menu commands (Image)

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
Options				
	Undo		[Ctrl] + [Z]	Can be used in Verification Extraction View.
	Settings	n/a	n/a	Displays a property sheet where you can configure Perceptive Intelligent Capture Verifier. (See CHAPTER 6 CONFIGURING VERIFIER AND ADVANCED VERIFIER)
	Filtering	n/a	n/a	Displays a property dialog where you can configure the batch filtering conditions. Not available in Verification View.
	License	n/a	n/a	Modify license connections: Use specified path or use License directory.
	Change client	n/a	n/a	Not supported in this application.
	Reclassify manually	n/a	[F7]	Accessibility: Indexing view.
	Show last verified document	n/a	[F3]	Option available in Advanced Verifier. Moves from the current document to the last document that has been verified, and switches from Verify Mode to Batch Mode.
	Get last value for selected field	n/a	[F8]	Option available in Advanced Verifier.
	Move document to exception state	n/a	[F9]	Moves current document to default exception state. Accessibility: Classification view, Indexing view.
	Release exception batches	n/a	[Ctrl]+ [E]	Release all exception document created batches.
	Apply local extraction	n/a	[Ctrl] + [L]	Option available in Advanced Verifier.
	Add document to Learnset	n/a	[Ctrl] + [A]	Option available in Advanced Verifier.
	Learn Documents	n/a	n/a	Option available in Advanced Verifier.
	Correct Tables	n/a	[Ctrl] + [T]	Option available in both Verifier and Advanced Verifier, but enhanced functionality only available in Advanced Verifier.
	Switch table highlighting	n/a	[Ctrl] + [Q]	Option available in Advanced Verifier.
	Learnset Manager	n/a	n/a	Starts the Learnset Manager.
	Move all exception documents	n/a	n/a	n/a

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
	Change Password	n/a	n/a	Allows you to change your password.

Table 7-15: 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 THE BATCH VIEW)
	Click this button to verify the selected batch. (See section THE VERIFICATION VIEW - INDEXING WINDOW)
	If you click this button, the Document View is displayed. (See section 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 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.

Button	Description
	This button is only enabled if the current document has more than one page. If you click it, the first page of the document is displayed.
	This button is only enabled if the current document has more than one page. If you click it, the previous page of the document is displayed.
	The text field allows you to enter a page number in order to directly navigate to it. All invalid entries (e.g. alphabetical characters, page numbers out of range, etc.) will be ignored, and the page number will be reset to the currently displayed page.
	This button is only enabled if the current document has more than one page. If you click it, the next page of the document is displayed.
	This button is only enabled if the current document has more than one page. If you click it, the last page of the document is displayed.
	Rotates the current document 90 degrees clockwise.
	Displays the first document in the batch, switches the application to Browsing Mode.
	Displays the previous document in the batch, switches the application to Browsing Mode.
	Displays the next document in the batch, switches the application to Browsing Mode.
	Displays the last document in the batch, switches the application to Browsing Mode.
	n/a
	n/a
	n/a
	n/a

Table 7-16: Verification View - Classification Mode Controls

7.3.3. Document Area

This area shows the current document.

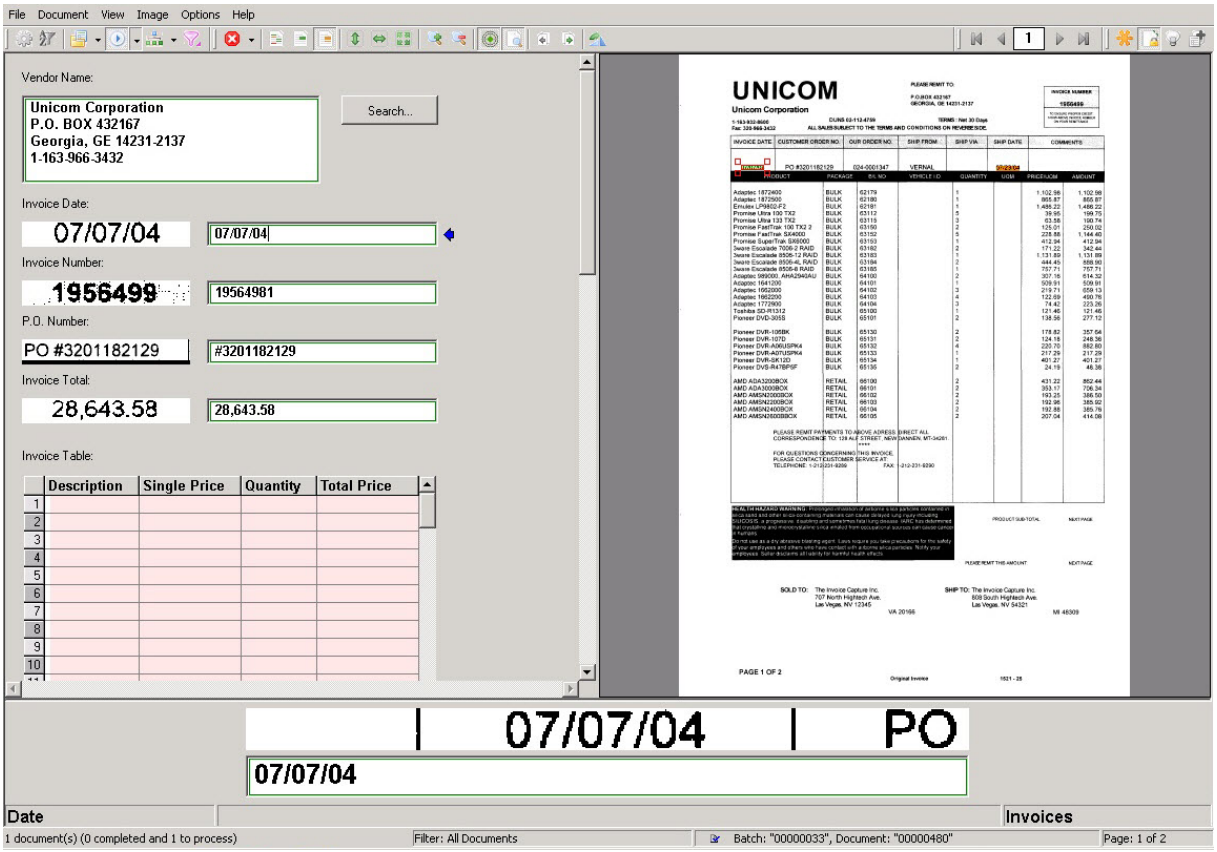


Figure 7-9: Document Area

7.3.4. Class Selection List

This box shows the classification result of the current document. If you open the list, you see the list of available classes.

The list entries represent the classes which have been assigned to the current project/user, and are controlled by the *VerifierClassify* script event.

If no result could be determined the box will show as empty. To set or change a classification result, make sure that you are not in Browsing Mode. Then either:

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

7.3.5. Advanced Classification

In the lower pane of the window on the left side, you will find the following button:

Click on this button will open a list containing either only one or more classes if the result could not be determined with 100% certainty.

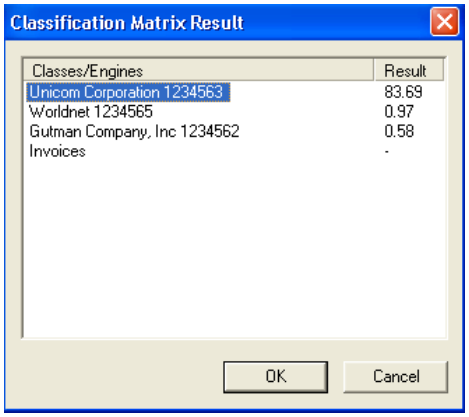


Figure 7-10: The Classification Matrix

If more than one class is in the list, the class entry is determined by probability. You will find the class with the highest probability on the top of the list. Select a class for the current document and then click *OK*.

Note: This button is not available until you have checked Enable advanced classification in Verifier on the Verifier Mode tab in Designer.

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 on the left.

The indexing window automatically displays the next document that requires a correction of the extraction result. The Indexing Mode typically looks like this:

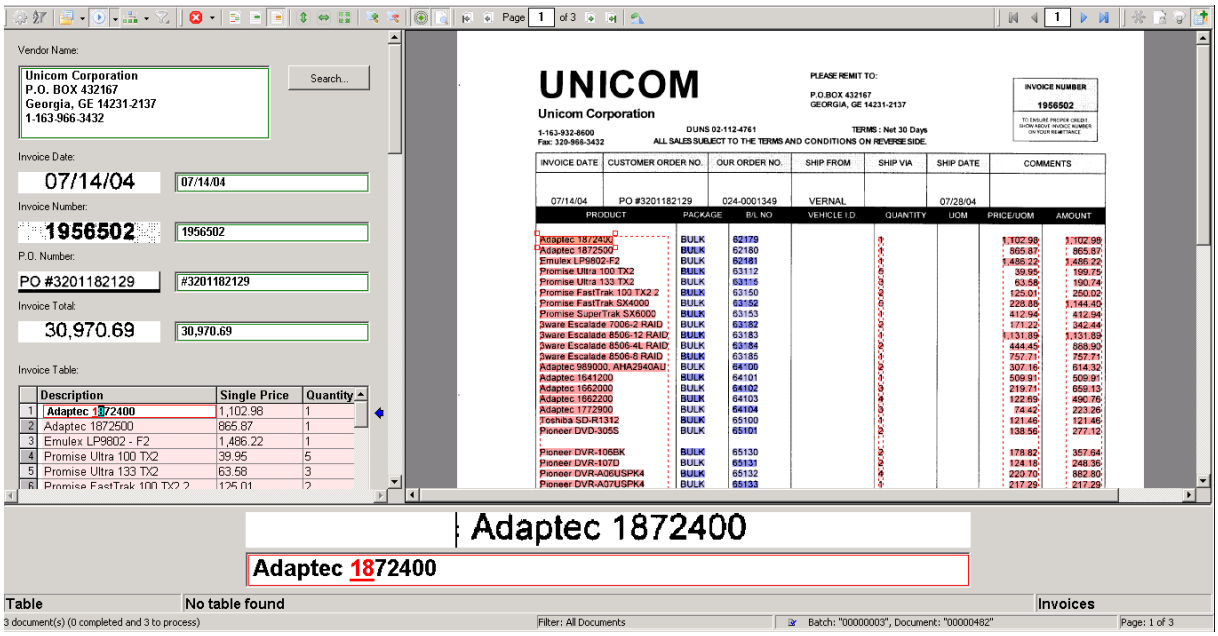


Figure 7-11: Indexing Mode

Menu Bar (See section 7.4.2 MENU COMMANDS AND KEYBOARD)

Document (See section 7.4.6 DOCUMENT AREA)

SHORTCUTS)

Toolbar (See section 7.4.3 TOOLBAR BUTTONS)

Current Area (See section 7.4.7 CURRENT INPUT AREA)

Field Area (See section 7.4.5 FIELD AREA)

User Info Area (See section 7.4.8 USER INFO AREA)

Your indexing window may look quite different than the one shown in FIGURE 7-11: 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 using the Designer application. Therefore, you might not be able to see all the window elements shown above as they may be arranged differently. For example, the document display could also be on the left side or there could be different window layouts for different document classes.

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

7.4.1. Increasing/Decreasing Image Area

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

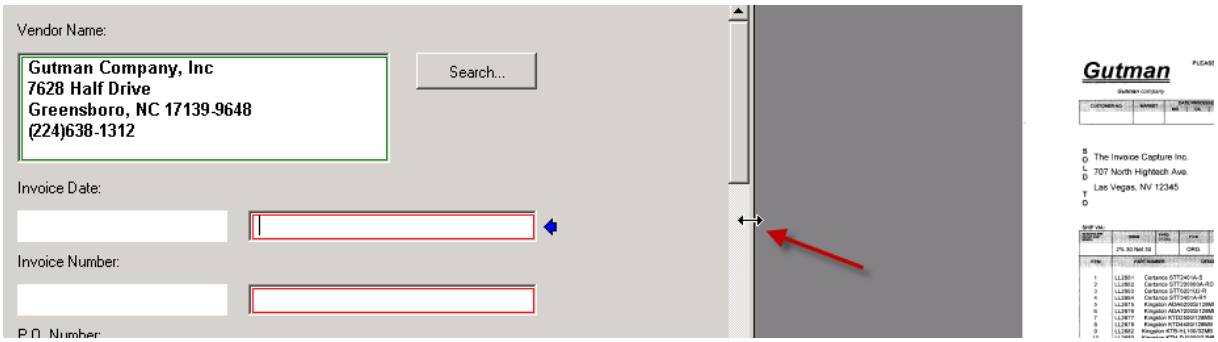


Figure 7-12: Indexing Mode – Split bar to increase/decrease image area

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

This option is only available in the extraction verification view.

7.4.2. Menu Commands and Keyboard Shortcuts

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
File				
	Print Setup	n/a	n/a	Adjust print output.
	Print	n/a	[Ctrl] + [P]	Print the document.
	Exit	n/a	n/a	Quits Perceptive Intelligent Capture Verifier.

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
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 7-17: 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.
	Batch List View	n/a	[Ctrl] + [1]	Displays the Batch View. (See section THE BATCH VIEW) The check mark shows that this view is selected.
	Verification Mode	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 THE VERIFICATION VIEW - CLASSIFICATION WINDOW) or in Indexing Mode. (See section THE VERIFICATION VIEW - INDEXING WINDOW)
	Document Separation Mode	n/a	[Ctrl] + [3]	Displays the selected batch in Document List.
	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, 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.

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
	Batch Filter	n/a	n/a	n/a
	Document Filter	n/a	n/a	Not available in Indexing Mode.
	Restore Focus	n/a	[Ctrl] + [N]	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.
	Multi-Page View	n/a	n/a	Display options for multi-page documents
		Single Page	n/a	Only one page is displayed. Use the navigation buttons to browse through further pages.
		Two Pages Horizontally	n/a	Two pages of the document are placed side by side.
		Three Pages Horizontally	n/a	Three pages of the document are placed side by side.
		Two Pages Vertically	n/a	Two pages of the document are placed one below the other.

Table 7-18: Verification View - Indexing Mode menu commands (View)

Menu	Submenu/ Command	Command	Description
Image (Navigates the document area specified below)			
	Zoom In	n/a	Zooms in.
	Zoom Out	n/a	Zooms out.
	Move Image to left	n/a	Scrolls the document area to the left.
	Move Image to right	n/a	Scrolls the document area to the right.
	Move Image upwards	n/a	Scrolls the document area upward.
	Move Image downward	n/a	Scrolls the document area downward.
	Rotate	n/a	Rotates the document clockwise.
	First page in document	n/a	Moves to the first page of a multipage document.
	Previous page in document	n/a	Moves to the previous page of a multipage document.
	Next page in document	n/a	Moves to the next page of a multipage document.
	Last page in document	n/a	Moves to the last page of a multipage document.
	Brightness & Contrast	n/a	Opens a dialog box with sliders to control the image's brightness and contrast. *Note: This setting will be kept for all further documents until next modification*
	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.
	Keep focus on field	n/a	Keeps the focus on the same field for each document you view in the batch.
	Keep zoom	n/a	Keeps the established zoom settings on each document you view in the batch.
	Increase image area	[Ctrl] + [J]	Increases the image area. (See section 7.4.1)
	Decrease image area	[Ctrl] + [K]	Decreases the image area. (See section 7.4.1)

Table 7-19: Verification View - Indexing Mode menu commands (Image)

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
Options				
	Undo	n/a	n/a	Retracts the last action.
	Settings	n/a	n/a	Open Settings dialog.
	Filtering	n/a	n/a	Displays a property dialog where you can configure the batch filtering conditions. Not available in Verification View.
	License	n/a	n/a	A dialog box is shown, to select the path for the license file.
	Change Client	n/a	n/a	Not supported in this application.
	Reclassify manually	n/a	n/a	Opens the classification window (See section 7.3 The Verification View - Classification Window) 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	n/a	[F9]	Moves current document to default exception state. Accessibility: Classification view, Indexing view.
	Release Exception Batches	n/a	[Ctrl] + [E]	Release all exception document created batches.
	Apply local extraction	n/a	[Ctrl] + [L]	Option available in Advanced Verifier.
	Add document to Learnset	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. This menu option will be available as soon as you place the cursor the table within field area. Clicking on it starts table correction.
	Switch table highlighting	n/a	n/a	This will change the line highlighting to show mapped lines and mapped columns.
	Learnset Manager	n/a	n/a	Starts the Learnset Manager.
	Change	n/a	n/a	Allows you to change your password.

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
	Password			

Table 7-20: Verification View - Indexing Mode menu commands (Options)

7.4.3. Toolbar Buttons

The toolbar provides quick access to some frequently used commands:

Button	Description
	Displays the Batch View. (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)
	Starts Learnset Manager.
	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.

Button	Description
	If this button appears pressed down, the application always displays the document area that is associated with the currently selected field.
	Keep zoom. Keeps the established zoom settings on each document you view in the batch.
	This button is only enabled if the current document has more than one page. If you click it, the first page of the document is displayed.
	This button is only enabled if the current document has more than one page. If you click it, the previous page of the document is displayed.
	The text field allows you to enter a page number in order to directly navigate to it. All invalid entries (e.g. alphabetical characters, page numbers out of range, etc.) will be ignored, and the page number will be reset to the currently displayed page.
	This button is only enabled if the current document has more than one page. If you click it, the next page of the document is displayed.
	This button is only enabled if the current document has more than one page. If you click it, the last page of the document is displayed.
	Rotates the current document 90 degrees clockwise.
	Displays the first document in the batch, switches the application to Browsing Mode.
	Displays the previous document in the batch, switches the application to Browsing Mode.
	Displays the next document in the batch, switches the application to Browsing Mode.
	Displays the last document in the batch, switches the application to Browsing Mode.
	Applies local classification and extraction to the current document.
	Adds current document to local Learnset.
	Starts document learning.
	Starts table correction.

Table 7-21: Verification View - Indexing Mode Controls

7.4.4. Support of Mouse Wheel

7.4.4.1. Description

Perceptive Intelligent Capture Verifier application supports the mouse wheel usage when validating documents in document verification mode. 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.

7.4.4.2. Usage

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

7.4.5. 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.

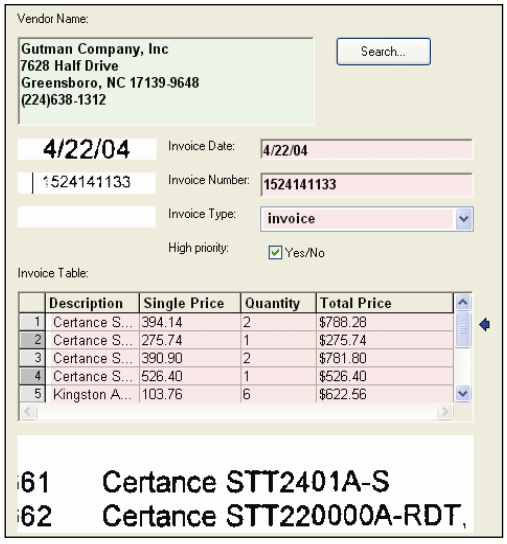


Figure 7-13: Field area example

Form fields

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

Check boxes

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

List boxes

A list of selection when verify an item on a document. Used during manual verification, this selection works with automatic completion.

Labels

Captions that help users to identify fields, viewers, and tables.

Viewer

Snippets of document areas normally 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 7-22: 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 the focus moves to the next field that requires correction. This field may also be within another document.

7.4.6. Document Area

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

- 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.

Note: In practice, red, green, and yellow areas never appear in the same document.

Table fields provide additional highlighting options, as shown in FIGURE 7-14. 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	Quantity	UOM	Unit Price	Price Uni
1	1372961008BOXSINGLE BL 2	2	BOX	2.50	1
2	POUNDS DRY ICE (REG PELLETS)	100	LB	0.58	1
3	HAZMAT FEE	1	EA	4.00	1

Figure 7-14: Highlighting for table fields

- Clicking the square in the upper-left corner of a table field will highlight the entire document table. In FIGURE 7-14, this square is above the 1 and to the left of Description.
- Clicking a column label of a table field highlights the appropriate document column.
- Clicking a row label of a table field highlights the appropriate document row.
- Clicking a cell of a table field highlights the appropriate document cell.
- Valid areas are green; invalid are red. These areas may also contain validity icons: green check marks for valid fields' or red Xs for invalid fields.

Note: 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.

For a large document with many line items, it is possible to detect and view the location of all the extracted line items that are currently shown within a table field. For more information please see section HIGHLIGHT EXTRACTED LINES WITHIN TABLES.

7.4.7. Current Input Area

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

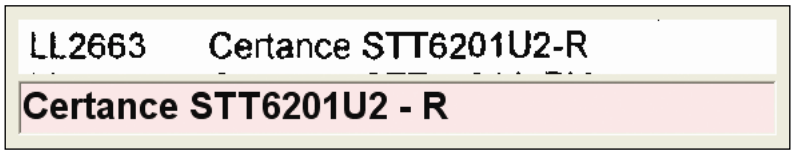


Figure 7-15: The input area

- 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 7.4.5 FIELD AREA) You can edit the data here.

7.4.8. User Info Area

The user info area 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

Perceptive Intelligent Capture Verifier now allows the currently opened image document to be printed. The user can do this by selecting *Print…* menu item of the *File* main menu option or by using Ctrl + P shortcut.

Note: 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 from the *Page Setup* dialog available in the *File* \ *Page Setup*… menu item of Perceptive Intelligent Capture Verifier application:

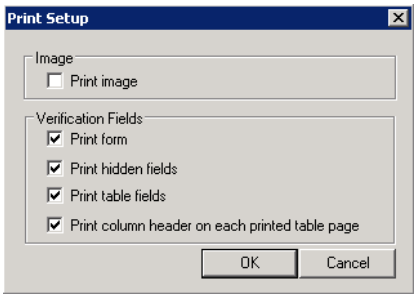


Figure 7-16: Print Setup

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	Effect (enabled) only if "Print table fields" option is turned on. When turned on (default), Verifier prints column header on each page (this option is useful for printing of long tables).

Table 7-23: Print Setup – available options

Save your Verifier Settings if you would like your current *Page Setup* preferences to be restored next time you start Perceptive Intelligent Capture Verifier application.

When you've 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.

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

In addition to the fields' content, Verifier also prints the document File 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 Perceptive Intelligent Capture Verifier. We recommend that you read CHAPTER 7 GETTING FAMILIAR WITH THE USER INTERFACE before you start working with this chapter.

In the Verifier, you can perform the following three main tasks:

- Correction of Page Separation – section 8.1.2
- Correction of Classification – section 8.2
- Correction of Extraction – section 8.4

The correction of extraction results comprises the extraction of document data into form fields as well as into tables. For table correction, please see sections 9.3.3 HOW BRAINWARE TABLE EXTRACTION LEARNS (STANDARD METHOD), section 9.3.4 ADVANCED LEARNING WITH BRAINWARE TABLE EXTRACTION, and section 9.3.5 ADVANCED LEARNING: ADDITIONAL FUNCTIONS.

In general, 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.

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 edifier project in that the *Project* and *Batch* root paths must be specified in the *General tab* of the *Verifier settings*.

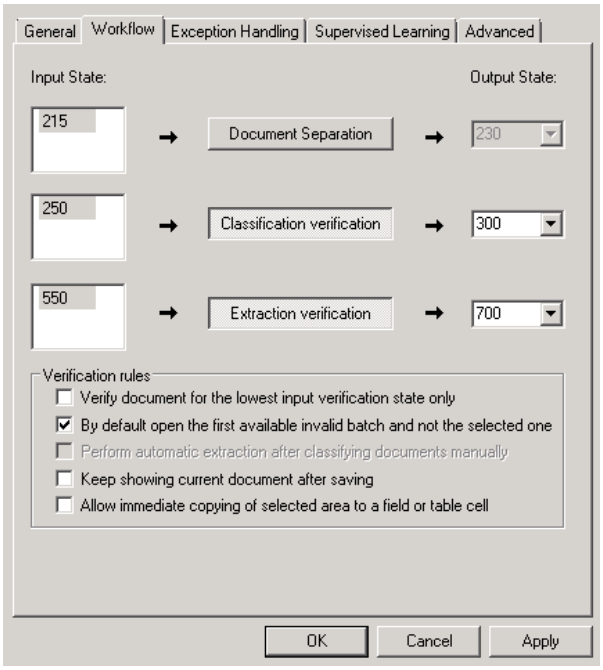


Figure 8-1: Settings for Page Separation

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

8.1.2. Manual Correction of Automatic Page Separation in Perceptive Intelligent Capture Verifier

If during the automatic page separation process 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 Perceptive Intelligent Capture Verifier application.

The special icon shown on the right indicates that this batch requires manual validation of the page separation's correctness.

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

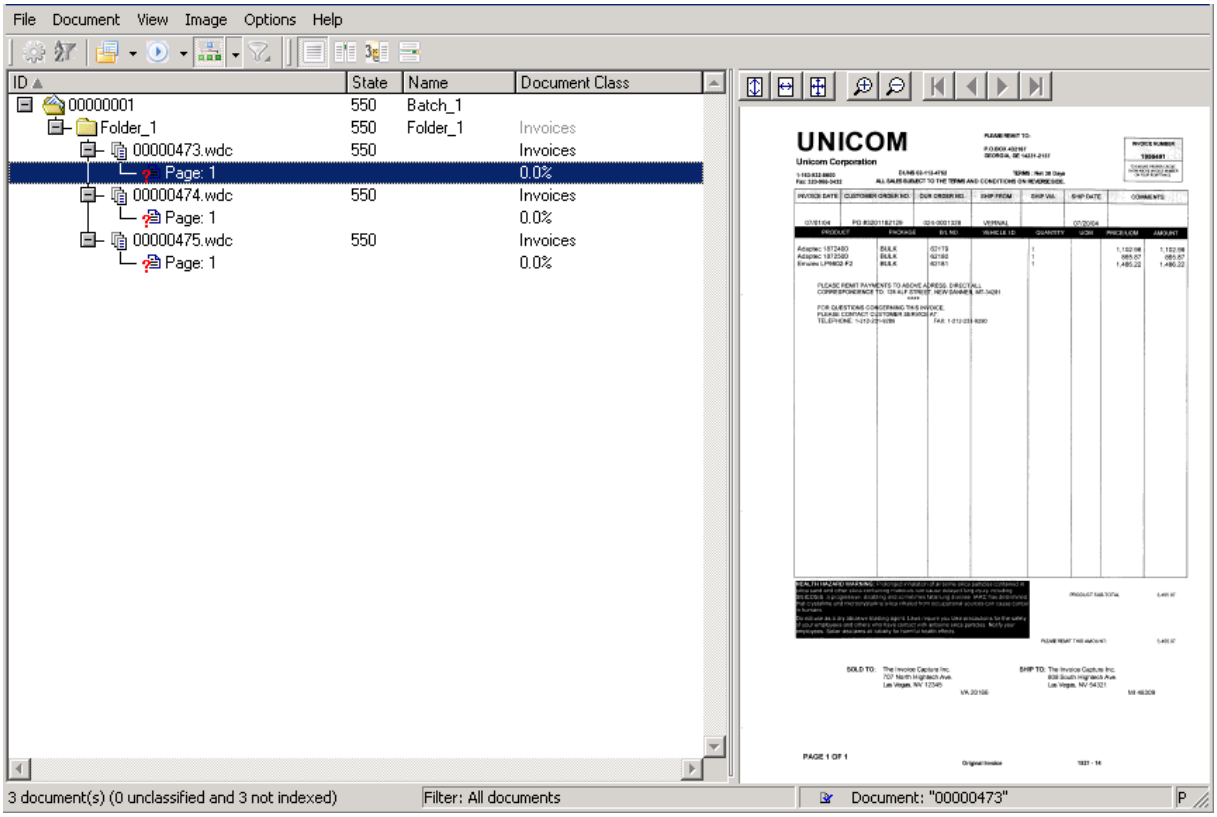


Figure 8-2: Document Browsing mode

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 document 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 gets manually accepted page correction status. Splitting previously merged documents will restore the original document names.
- *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 proceeding one automatically gets 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.

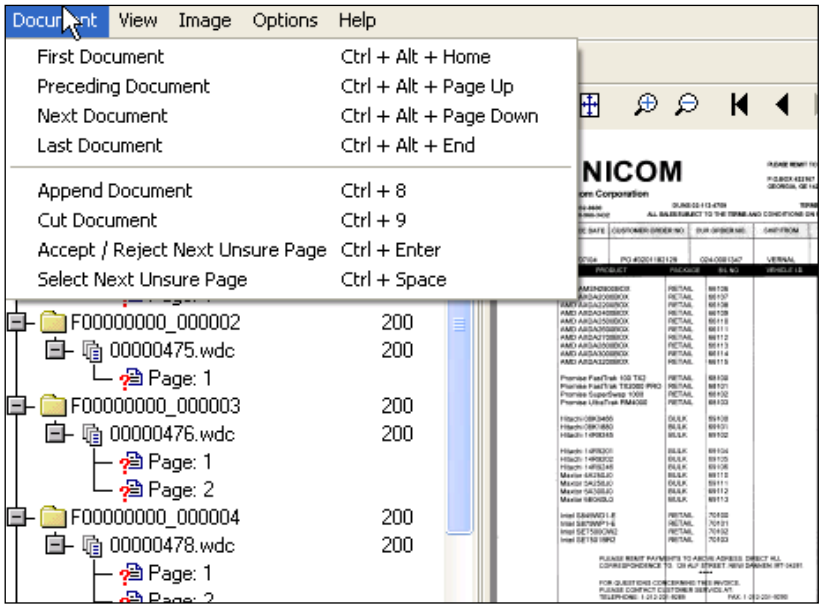


Figure 8-3: Select Next Unsure Page option

When verifying the correctness of automatic page separation in Perceptive Intelligent Capture Verifier, the user can now select between different convenient page view modes, for example having two consecutive pages displayed simultaneously:

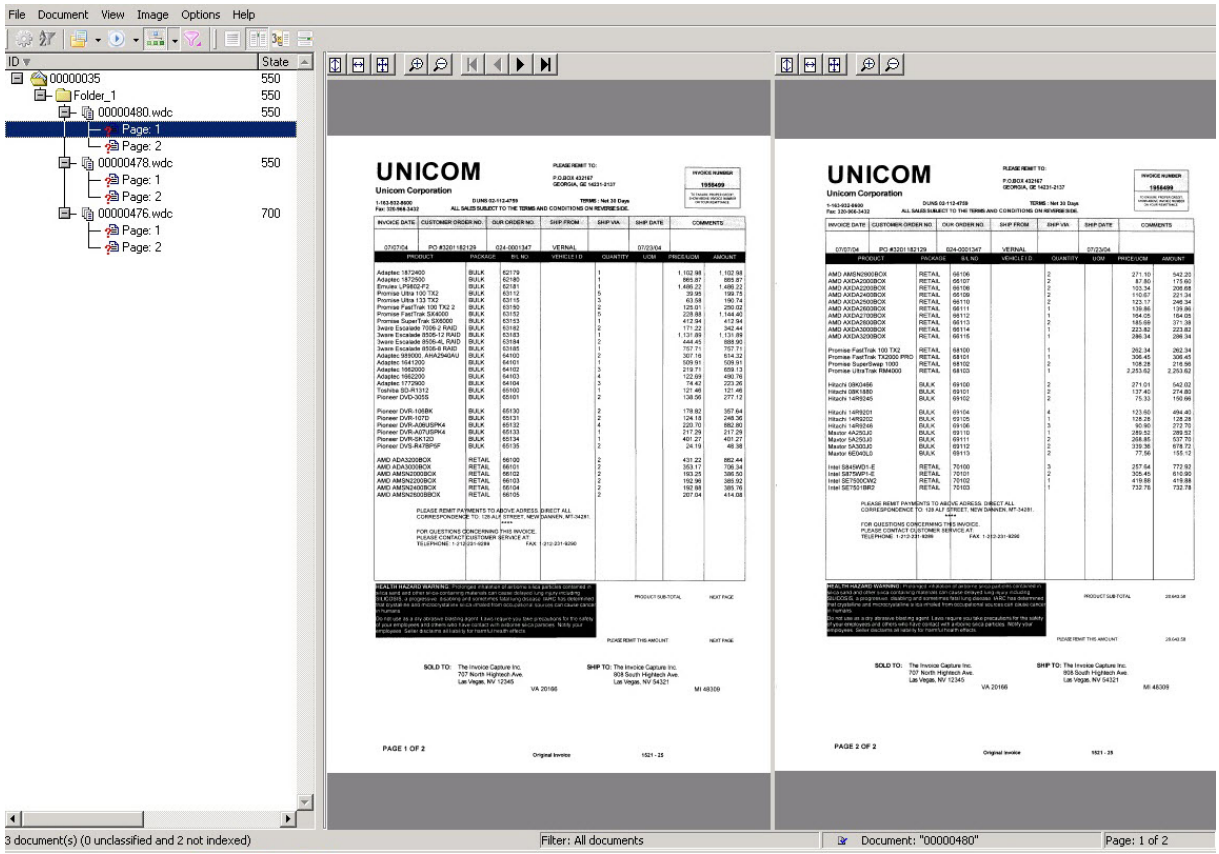


Figure 8-4: Verifying automatic page separation – View modes

There are 4 different page view modes available in Perceptive Intelligent Capture Verifier:

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

- Two pages displayed vertically view.

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

8.2 Manual Correction of Classification Results

Manual correction of classification results will be done if the Perceptive Intelligent Capture Verifier workstation is configured as follows:

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

To determine your settings, check the workflow tab of the Perceptive Intelligent Capture 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.
5) When all documents in the batch are validated, the application prompts you to select what you want to do next.

Figure 8-5: Finishing a batch

6) Select the next step by clicking *Yes, No*, or *Details*. Clicking on Details will reveal more options:

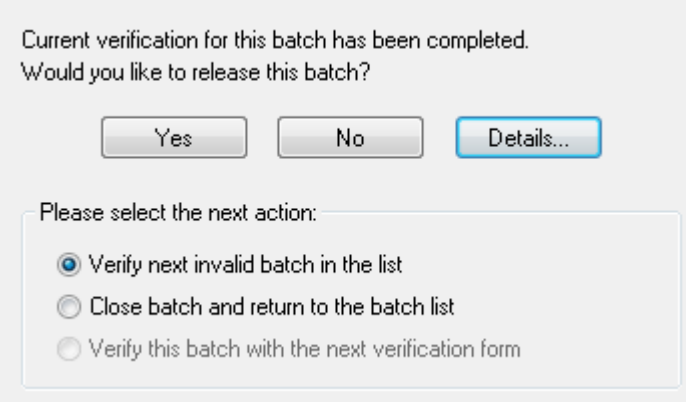


Figure 8-6: Actions after batch completion

- *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. 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 only be processed if their former parent class still exists in the project where the document is being processed.

8.3.2. Usage

The present feature is very useful in the context of Supervised Learning Workflow (using Advanced Verifier), where the Vendor class is often deleted (or not inserted) from the global project's configuration.

8.4 Manual Correction of Extraction Results

Manual correction of extraction results will be done if the Perceptive Intelligent Capture 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 be configured within the *Fields edit mode* on the *General Settings* tab.)

To determine your settings, check the *Workflow* tab of the Perceptive Intelligent Capture Verifier *Properties* dialog box. (See section SETTINGS – WORKFLOW)

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

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 (see SETTINGS – GENERAL, Fields edit mode). 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.

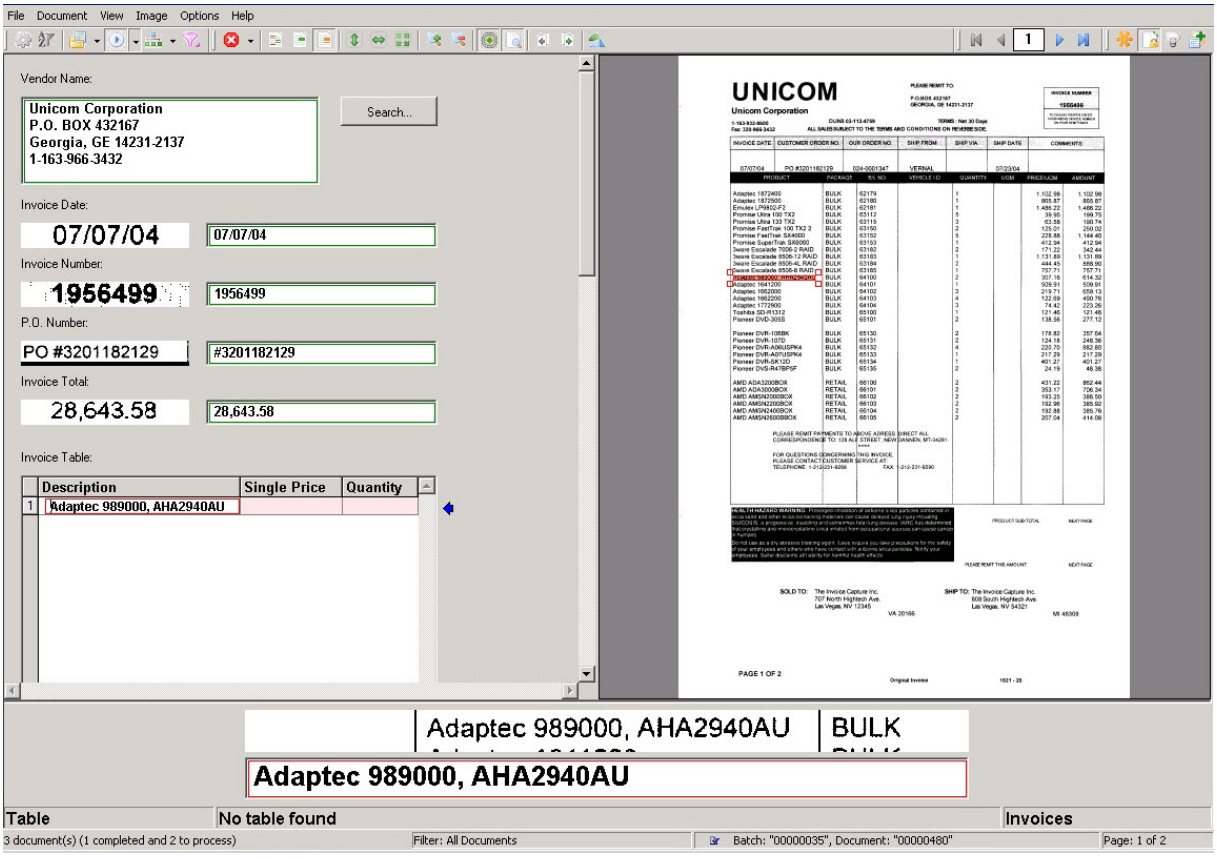
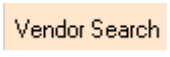


Figure 8-7: A document open in Verifier

Note: You'll notice a Vendor Search button in many of the illustrations in this document.



This button is created and programmed in Perceptive Intelligent Capture's scripting language either for the Associative Vendor search or for the Multi-column Associative search engine. In these illustrations, the button is used to quickly classify or reclassify as shown in FIGURE 8-8: VENDOR CANDIDATES LIST. To learn more about the scripting language, see the ***Scripting Documentation****.*

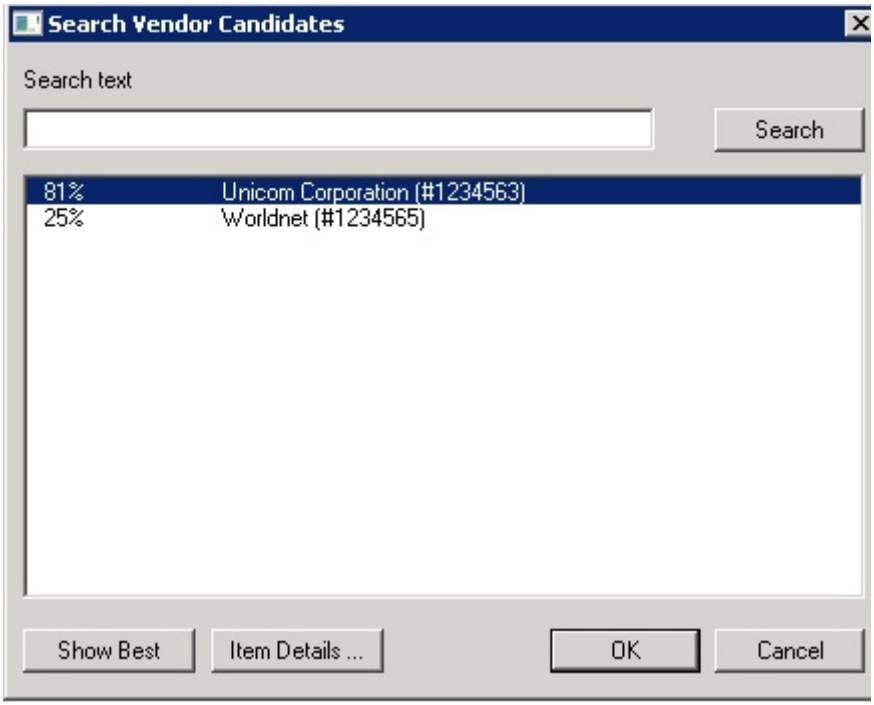


Figure 8-8: Vendor Candidates List

8.4.1.2. Editing Text Fields

Perceptive Intelligent Capture 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 the selected text.

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

Note: 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. Auto-completion feature for a header field is supposed to automatically select the best candidate from the available ones, which works

within *Highlight Candidates* mode. *However, the viewer will be updated only if the candidate appears once in the document; otherwise the viewer will be blank when auto-text completes the word for the 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.

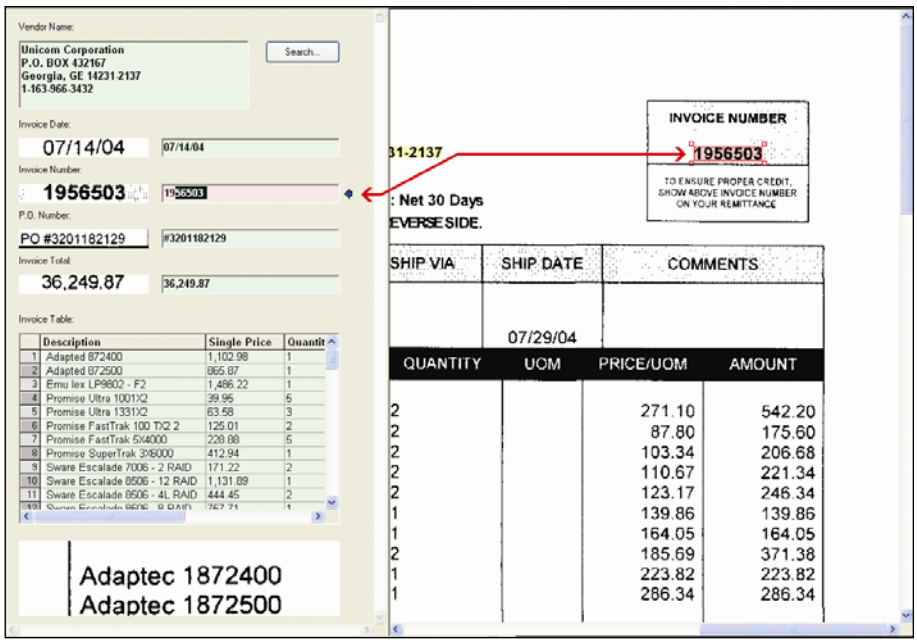


Figure 8-9: Example of auto-completion

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.

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

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

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

Make sure that this word fits the format analysis rules defined for that field. 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 6.2 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.

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

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 field proceed as described above. For Table Correction, please refer to CHAPTER 9.

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

Select the next step by clicking *Yes, No* or *Details*:

- *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 disabled.

Note: If this option is enabled, extraction will be carried out automatically by Perceptive Intelligent Capture *Runtime Server. In this case, classification verification and extraction verification are two separate steps.*

To determine your settings, check the workflow tab of the Perceptive Intelligent Capture Verifier *Properties* dialog box. (See section 6.2 SETTINGS – WORKFLOW)

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

8.6 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. There will rarely be a company without a customer database that contains addresses, contacts, and so on. Financial transactions are also recorded in databases. Databases can be an excellent support for indexing because they store related information that 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 a * to represent a sequence character or a ? 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.

8.7 Checking Entire Batches

To browse through all documents in a batch:

1) In the *Batch View*, use the status value to determine a batch you can browse through. 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

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

In Perceptive Intelligent Capture, it is possible that any given table was either trained by the new extraction engine – Brainware Table Extraction or the traditional Table Analysis Engine. As a Verifier user you will not know which engine was being used and you will not be able to see a difference in Verifier. The process of table fields is similar regardless of whether you are using Verifier or Advanced Verifier.

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

9.1 Automatic Training and Extraction of Verified Table Data

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

Brainware Table extraction supports the Generic Table extraction. When using Generic Table Extraction table data is easily mapped with numerical correlation and Label matching. No extra learning is needed.

So far column and line extraction is not supported with the Generic Table Extraction. Extract and correct fields manually while working with the Generic Table Extraction (see section 9.3.3.5).

9.2 Manual Training and Correction Methods

In case that the automatic table extraction failed to recognize the line items properly, the Verifier application provides several ways for convenient manual table correction.

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.) This function only works with *Highlight Candidates* mode  .

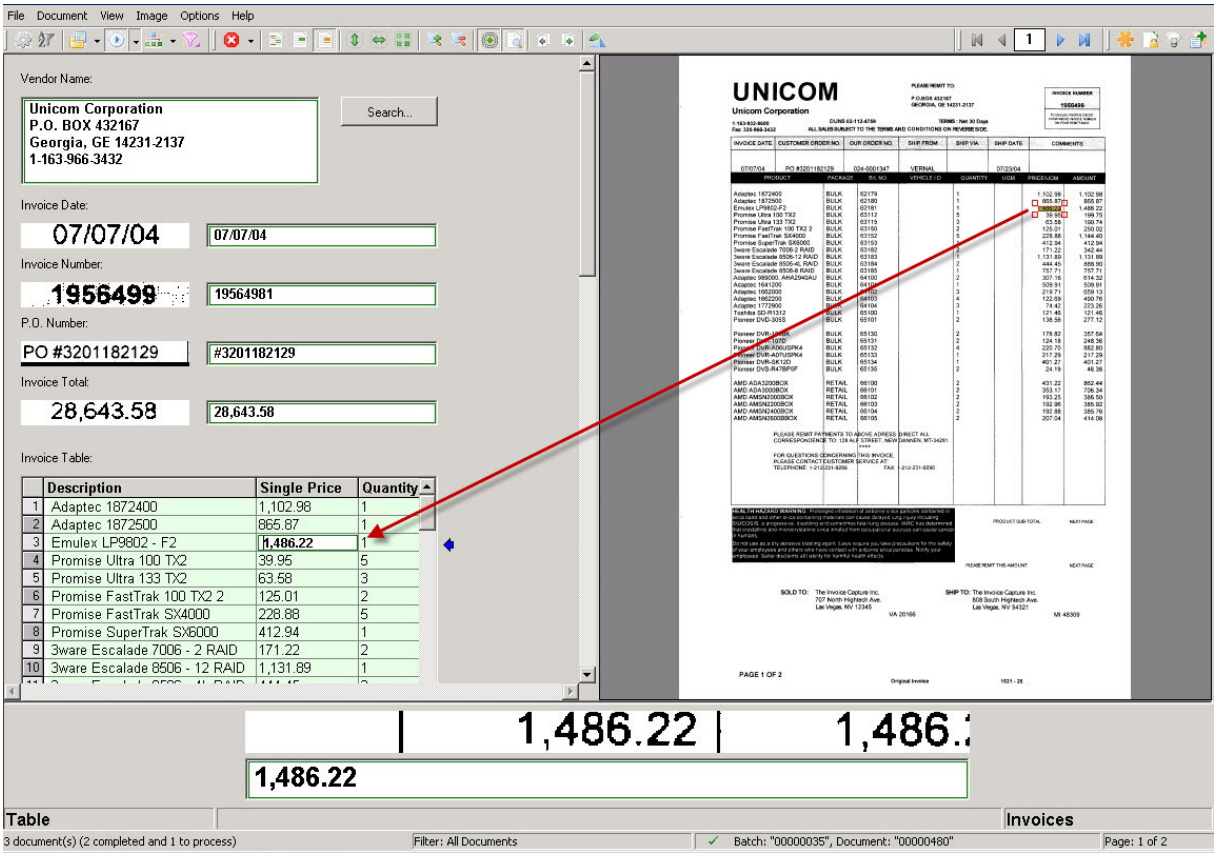


Figure 9-1: Table with text auto-completion

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, and then click *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.

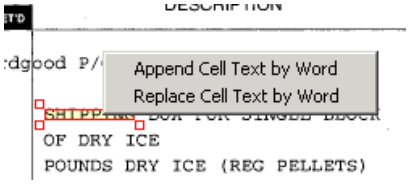


Figure 9-2: Table cell, word in cell area

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 *Align & Copy to Current Field*. To replace text, select the word, and then press CTRL + left click, or select *Copy to Current Field* in the shortcut menu.

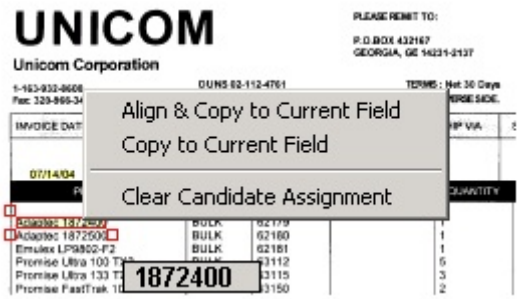


Figure 9-3: Table cell, word not in cell area

9.2.3. Correcting Table Structure

You may also need to correct the table structure. Table rows, cells and columns have shortcut menus with options for modifying the table structure. To invoke them, right click on the row, cell 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.
Cell		
	Insert Cell	Will insert an empty cell above the selected cell while shifting the cells below down.
	Delete Cell	Will delete the selected cell. The cells below will be shifted up.

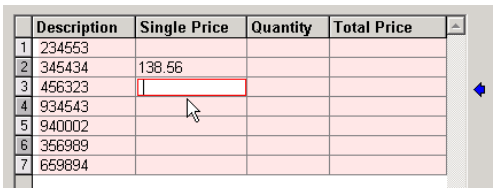
Table 9-1: Table shortcut menus

9.2.3.1. Rubber Banding Feature

The Rubber Banding Feature allows you to place a block of values at a particular point within the table.

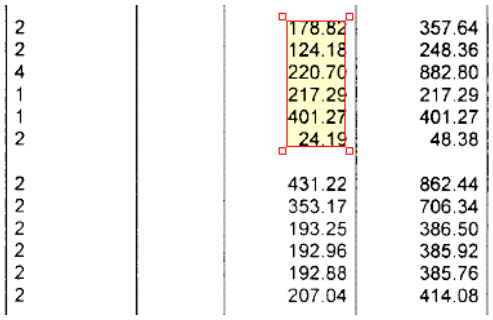
The basic steps are as following:

1. Place the cursor into the destination cell within the table.



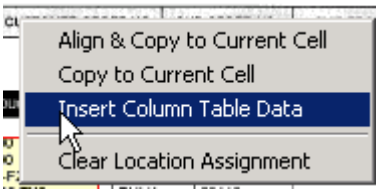
	Description	Single Price	Quantity	Total Price
1	234553			
2	345434	138.56		
3	456323			
4	934543			
5	940002			
6	356989			
7	659894			

2. Draw a rubber band rectangle around the desired column data within the document viewer.



2		178.82	357.64
2		124.18	248.36
4		220.70	882.80
1		217.29	217.29
1		401.27	401.27
2		24.19	48.38
2		431.22	862.44
2		353.17	706.34
2		193.25	386.50
2		192.96	385.92
2		192.88	385.76
2		207.04	414.08

3. Right-click the selection and click on *Insert Column Table Data* within the context menu.



Use cases

This is helpful for following scenarios:

1. If columns of one data type are split up and placed side-by-side, or stacked, on a page as shown in the example below.

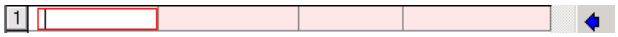
Test Company 113 Testing Bvd. Boston MA 02105

Invoice No.	Reference	Invoice Amt.	Total	Invoice No.	Reference	Invoice Amt.	Total
934567	A750405	13.55	13.55	456323	G455940	54.90	748.14
356454	C550403	355.78	369.33	934543	H845344	165.90	914.04
842039	D359403	33.56	402.89	940002	I345453	74.45	988.49
234553	E940589	23.45	426.34	356989	J892345	46.78	1035.27
345434	F453653	266.90	693.24	659894	K345655	245.90	1281.17

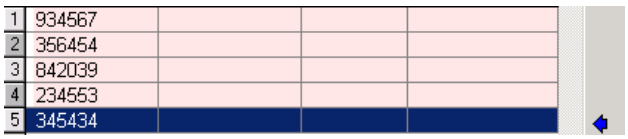
Figure 9-4: Invoice with stacked columns

To correct such a table (the table does not contain any data):

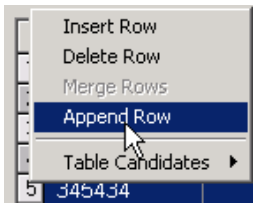
- Place the cursor into the first line of the desired column.



- Draw a rubber band rectangle around the first block of column data on the document.
- Right-click the selection and click *Insert Column Table Data*. Based on the document example above, the mapping result would look as following:



- Right-click the row's node of the last table row and select *Append Row*.



- Place the cursor into the next empty cell of the desired table column.
- Now draw a rubber band rectangle around the next data block of the same column type and select *Insert Column Table Data*.
- Proceed the same way with other columns to fill the table.

2. If column items have been extracted only partially.
Scenario A

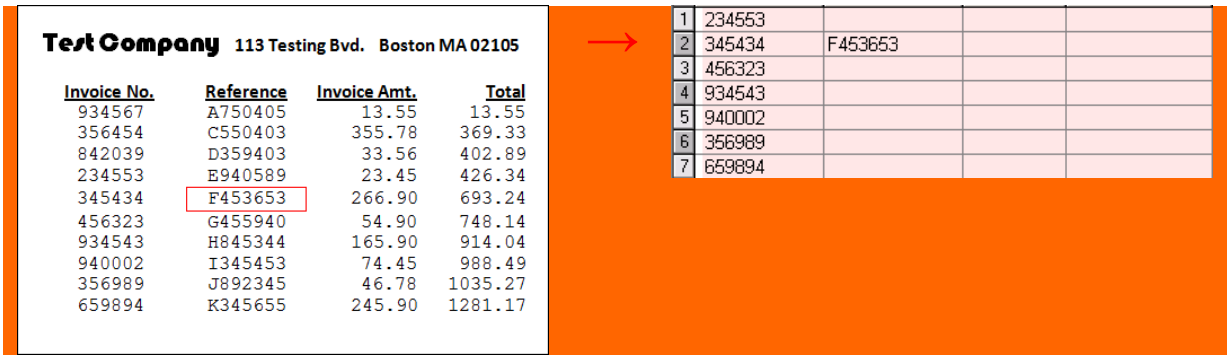
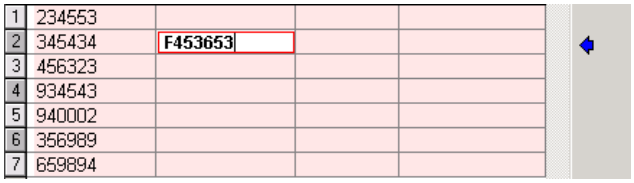


Figure 9-5: Partial extraction of line items

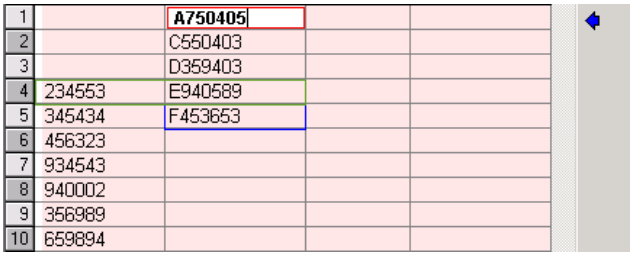
To correct such a table:

- Place the cursor into the already extracted cell to mark the starting point for the insertion.



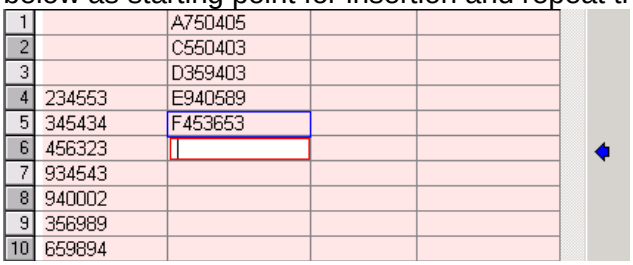
- Draw a rubber band rectangle around the desired column data on the document.
- Right-click the selection and select *Insert Column Table Data*.

This will insert the values above the previously selected cell. At the same time, the values will be automatically assigned to already extracted values of other columns if available (green frame in the screenshot below).



1		A750405
2		C550403
3		D359403
4	234553	E940589
5	345434	F453653
6	456323	
7	934543	
8	940002	
9	356989	
10	659894	

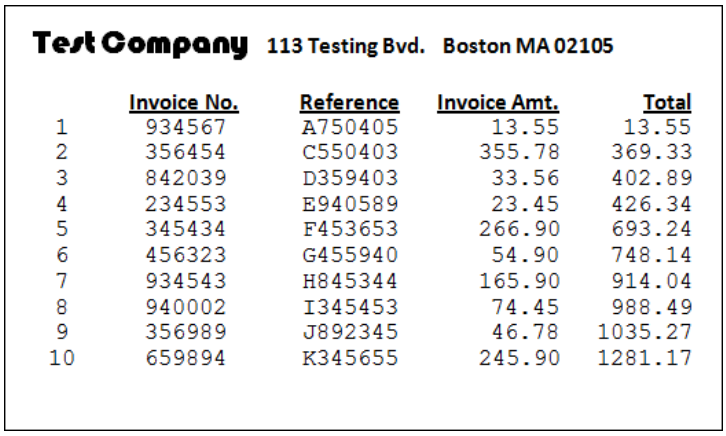
Now continue to map the other column data. To add the column data under the cell “F453653” (blue rectangle in the screenshot allow), select the very next empty cell below as starting point for insertion and repeat the steps from above.



1		A750405
2		C550403
3		D359403
4	234553	E940589
5	345434	F453653
6	456323	
7	934543	
8	940002	
9	356989	
10	659894	

Scenario B

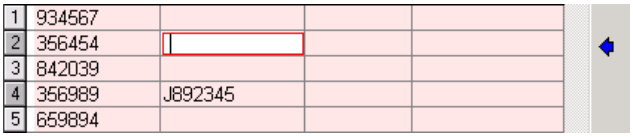
Missing column items in the neighbour column.



Test Company 113 Testing Bvd. Boston MA 02105

	Invoice No.	Reference	Invoice Amt.	Total
1	934567	A750405	13.55	13.55
2	356454	C550403	355.78	369.33
3	842039	D359403	33.56	402.89
4	234553	E940589	23.45	426.34
5	345434	F453653	266.90	693.24
6	456323	G455940	54.90	748.14
7	934543	H845344	165.90	914.04
8	940002	I345453	74.45	988.49
9	356989	J892345	46.78	1035.27
10	659894	K345655	245.90	1281.17

Taken the document example above, the column data has been exactred as follows:



1	934567	
2	356454	
3	842039	
4	356989	J892345
5	659894	

From the ‘Reference’ column, only the value from line 9 is extracted, and the ‘Invoice No.’ values from lines 4 – 8 are missing.

To insert the ‘Reference’ values from lines 2 – 8 as a block:

- Place your cursor into the ‘Reference’ cell in line 2.
- Perform the steps as described in previous scenario.

Please note, that the application observes the relationship between the ‘Invoice No.’ and ‘Reference’ columns and maps the values appropriately. ‘Reference’ values where the according ‘Invoice No’ values are missing, are inserted as rows while shifting the subsequent, already extracted cell “J892345” down.

1	934567			
2	356454	C550403		
3	842039	D359403		
4		E940589		
5		F453653		
6		G455940		
7		H845344		
8		I345453		
9	356989	J892345		
10	659894			

3. If you have documents where the data columns appear misaligned.

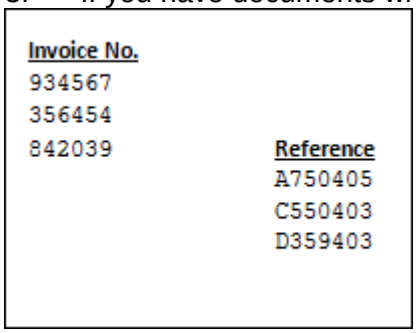


Figure 9-6: Invoice with misaligned data columns

Map the column data blockwise as described for previous use cases.

9.2.3.2. Correcting Single Cells

You may also need to correct the table structure if for instance an unnecessary cell has been mapped to the table or if a missing cell has to be added.

Use Case

As in the example below, you may have documents where one of the line items is missing. During extraction, the values from below might be shifted up to fill the empty space.

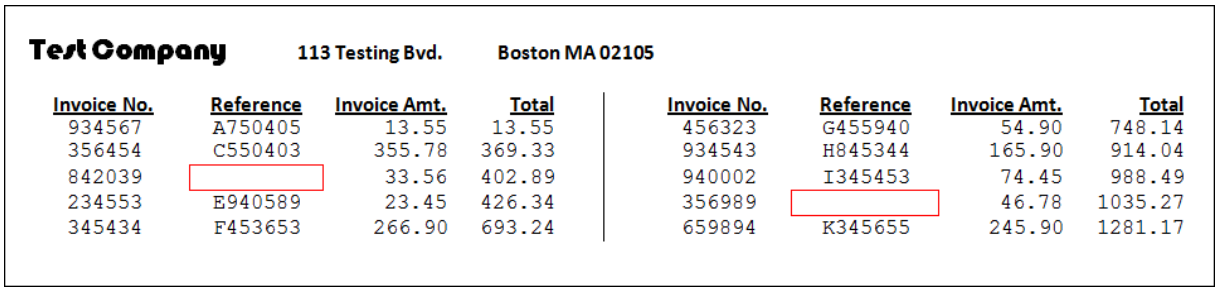


Figure 9-7: Invoice with missing line items

Solution

For this, you have the possibility to add or remove single cells.

For each of the table cells, the following context menu is available:

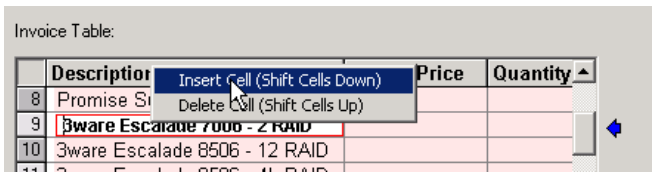


Figure 9-8: Inserting and deleting table cells

To delete an unnecessary cell:

1. Click the cell within the table and place the cursor in it.
2. Right-click and select *Delete cell (Shift Cells Up)*.

To insert a cell:

1. Click the cell within the table to be subsequent to the cell candidate and place the cursor in it.
2. Right-click and select *Insert cell (Shift Cells Down)*.

This will create an empty cell within the table above the selected cell. Now, you can copy the desired value from the document into the newly created cell.

9.3 Brainware Table Extraction and Correction

Note: This functionality is available for the Supervised Learning Verifiers. With the Generic Table Extraction no extra learning is needed.

9.3.1. About Brainware Table Extraction

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

- Learning lines
- Learning mappings of columns

These are discussed in detail in the following sections.

9.3.1.1. Learning Lines

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

- **Primary line**: A line that defines table structure. The BTE engine applies advanced and precise similarity analysis for all primary lines. It is important that all primary lines are well-structured and that they look similar in many of the rows to extract. The engine easily supports an unlimited number of types of primary lines for one table definition. The primary line must contain at least four words. Otherwise, the BTE engine will not learn it. Also, the primary line must be the first line in the table row.
- **Secondary line**: A line between primary lines. The engine applies smooth similarity analysis for these types of lines, which is possible because BTE 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 that 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.) Afterwards, all lines in the document considered by the engine to be members of the line class (similar to the learned line sample) will be extracted, or not extracted in the case of “wrong” lines.

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

Learning lines can be applied in lines learning (or lines highlighting) mode. Mapping of the column data in the lines can be done in column mapping learning (or columns highlighting) mode. The user can switch between learning (highlighting) modes from 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 Perceptive Intelligent Capture class the Learnset is shared between these tables. In other words, if you used interactive learning for one BTE table, cross-document learning (which happens if the system added the document to the Learnset after document validation) will be applied for all BTE tables in the document.

9.3.2. Correcting Fields in Tables Created with Brainware Table Extraction

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

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

Brainware 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 Brainware 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.

Figure 9-9: BTE Table

2) Click on 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.

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

Learning a New Line as Primary Line

- Right click on any line marked in gray in the TIFF.
- 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

- In the document viewer draw a rectangular selection over the primary lines
- Right click on the selection.
- 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 will be similarly extracted and displayed.

If some lines were not extracted (these lines will not 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 will be similarly extracted and displayed.

If some lines were not extracted (these lines will not be color-coded) repeat the procedure described 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 would train all different line samples now, you would need to learn the columns mapping separately for every line class. In order to reduce time to train the table, first learn the column mapping for the row you just learned. If you then want to learn another line sample, the engine will apply the existing mapping rules for the newly learned row automatically.

Note: Green highlighting indicates 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 shows 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 the 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 by double clicking on it.
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.

Note: Pressing CTRL+Q switches the highlighting 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 described before.
3) Repeat this step for all rows on all pages where something is missing. Go to the next step only after you are sure nothing is missing.

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 cannot 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.

Note: The BTE engine may determine the mapping automatically.

Repeat these steps until the data is completely extracted or cannot be learned correctly. (There is always a remaining risk 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.

Note: 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 the document's fields are validated, you may want to add this document to the Learnset 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 Local Learnset* toolbar button is not pushed,) but you still would like to learn your table, activate learning by clicking the *Add Current Document to Local Learnset* button.

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

9.3.3.7. Highlight Extracted Lines within Tables

For a large document with many line items, it is possible to detect and view the location of all the extracted line items that are currently shown within a table field.

In order to switch from cells correction highlighting mode to lines highlighting mode:

1) Open the *Options* menu
2) Select *Switch Table Highlighting* or press CTRL + Q

This will change the line highlighting to show mapped lines and mapped columns.

Non extracted lines are displayed in grey, and it is not possible to click on them. However, when clicking on an extracted line right next to it above or below, the appropriate table line will be highlighted by a blue arrow pointing to it. This will give an orientation where the missing line should be located in the table.

Note: This highlighting mode is not only available for Brainware Table Extraction tables, but is also applicable for e.g. script extracted tables.

If you want to insert the missing line into the table, a script Insert Line *button is required.*

9.3.4. Advanced Learning with Brainware 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.

Note: Never use the Learn as Row function in this case, as this will tell the engine that the first secondary line is actually a new sample of a 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 is not 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 use *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 will not 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, Brainware Table Extraction uses a line sample, searches for the line type and removes the line type from the Learnset.

1) Switch to lines learning mode and right click on the line you want to unlearn.
2) On the shortcut menu, click *Unlearn Line*. Unlearned lines change from green to gray.

9.3.5.3. Learn Line as Wrong Line

To Learn 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 Learnset Manager

10.1 What Does Supervised Learning Do?

Note: The Learnset Manager is a separately licensed add-in that can only be launched from Verifier. You have access to the Learnset 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 Learnset Manager.

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

In general, Learnset Management consists of:

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

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 Learnset Manager

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

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

10.3 Getting Familiar with the Learnset Manager User Interface

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

When you are working with Local Learnsets, 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 Browsing

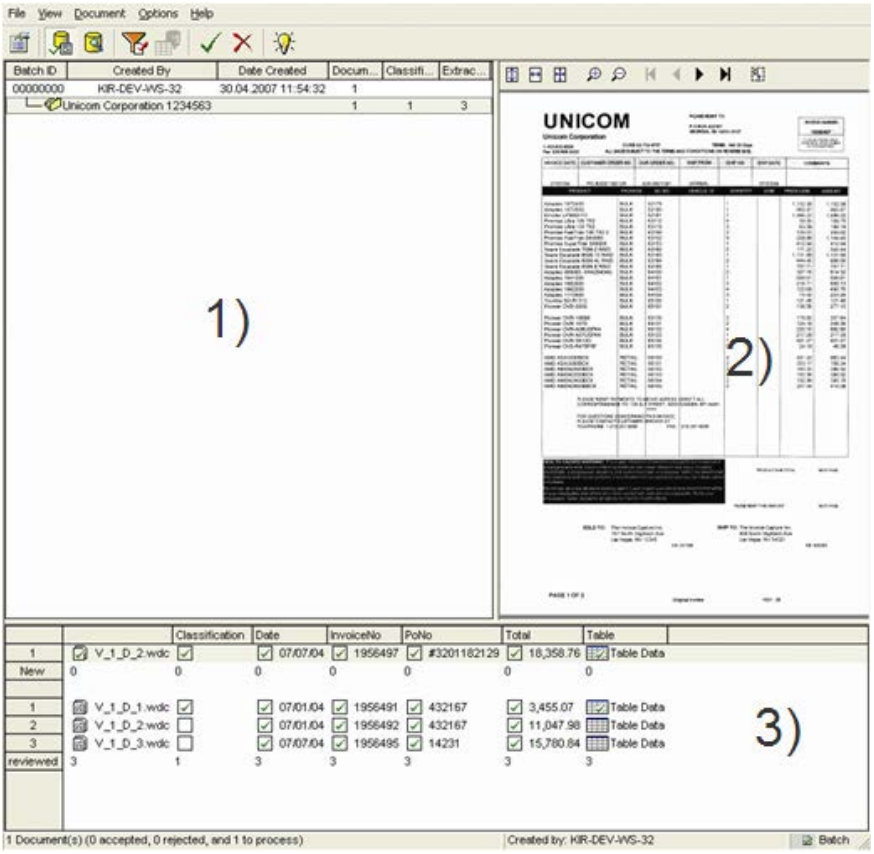


Figure 10-1: Accumulated Documents Browsing (Local Learnset) View

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 need 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 Perceptive. 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 Learnset Browsing

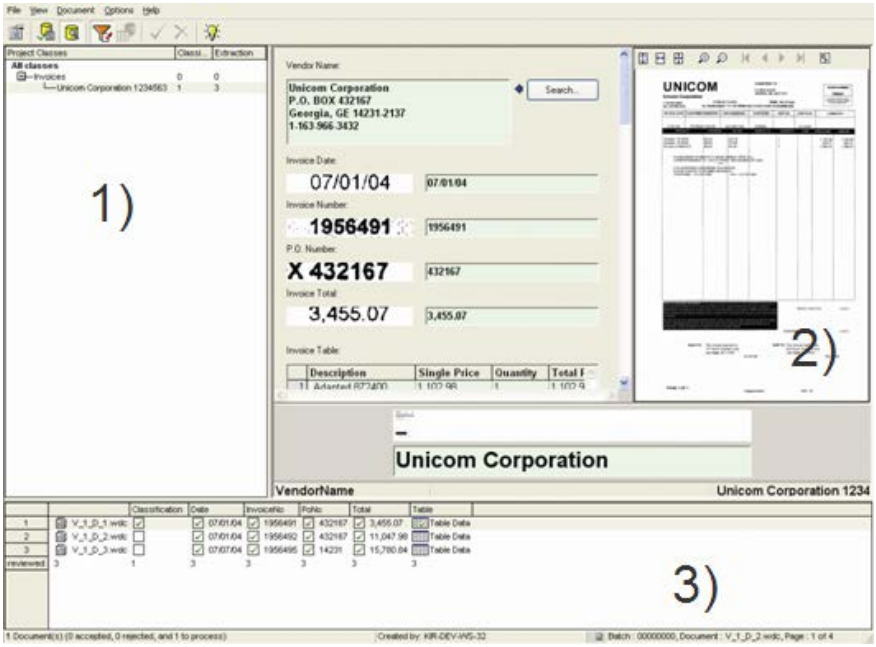


Figure 10-2: Global Learnset Viewer

1) The Batch viewer: Enables you to see each class in the batch you are working on. Shows the classes in the Global Learnset, 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 Perceptive. 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	Command	Keyboard Shortcut	Description
File				
	Exit	n/a	[Ctrl] + [E]	Closes the Learnset 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.
	Accumulated Documents Browsing	n/a	n/a	Switches to the Accumulated Documents Browsing mode (Local Learnset).
	Global Learnset Browsing	n/a	n/a	Switches to the Global Learnset Viewer.
	Verify Documents	n/a	[F5]	Verifies documents in Supervised Learning.

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
	Refresh	n/a	n/a	
	Sort Batches by Vendor...	n/a	n/a	Available to Verifiers and Learnset Managers. Not available in Global Learnset Browsing.
	Sort Batches by Date…	n/a	n/a	A sort option for batches

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

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
Document				
	Accept Document	n/a	n/a	Accepts the current document
	Reject Document	n/a	n/a	Rejects the current document
	Learn Accepted and Remove Rejected	n/a	n/a	Learns all accepted and removes all rejected documents
	Learn Accepted Only	n/a	n/a	Learns only the accepted documents
	Remove Rejected	n/a	n/a	Cleans up by removing all rejected documents
	Learn All New	n/a	n/a	Learns all new documents
	Relearn All in Global Learnset	n/a	n/a	Available only to Supervised Learning Managers.
	Relearn Changed in Global Learnset	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	Command	Keyboard Shortcut	Description

Menu	Submenu/ Command	Command	Keyboard Shortcut	Description
Options				
	Train Base Classes...	n/a	n/a	Available to Verifiers and Managers.
	Update Local Projects...	n/a	n/a	Available to Verifiers and Managers.
	Open with Perceptive Intelligent Capture Workdoc Browser	n/a	n/a	Opens the Workdoc Browser.
	Settings...	n/a	n/a	Shows settings for the Learnset Manager.
Help				
	Perceptive Intelligent Capture Verifier Help	n/a	n/a	Opens Verifier Help.
	Learnset Manager Help	n/a	[F1]	Opens Learnset Manager Help.
	About Perceptive Intelligent Capture Learnset Manager…	n/a		Shows information on licence file name, location and application's version

Table 10-3: 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 Browsing (The Local Learnset.)
	Switch to Global Learnset Processing.
	Verify Documents. Activates the Advanced Verifier mode.
	Correct tables. Allows you to correct data in the tables. You have to click a table field for this to be active.

Button	Description
	Accept documents.
	Reject documents.
	Learn documents. (Adds them to Global Learnset.)

Table 10-4: 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.

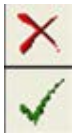
Button	Description
	Fits the document to window height.
	Fits the document to window width.
	Best fit.
	Zooms in. Alternatively, press [Ctrl] + [+].
	Zooms out. Alternatively, press [Ctrl] + [-].

Table 10-5: Viewer toolbar buttons for supervised learning

10.4 Using the Learnset Manager

10.4.1. Overview of the Process

Use the Learnset Manager to work with Local Learnsets: First, verify the documents, decide whether they belong in the Learnset, add them to the Common Learnset, and train the Learnset.



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

After that, you train the Global Learnset.

10.4.2. Getting Ready to Use the Learnset Manager

To use the Learnset Manager:

1) In Verifier, examine the properties for Learnset Manager. (FIGURE 10-4) Most were established in Designer. However, you need to ensure that Learnset Manager is enabled, that the *Activate Supervised Learning workflow* checkbox is checked. Also, ensure that the paths for Local Project Name, Local Learnset 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 Learnset Directory** should end with the Learnset Directory folder (it is probably called Learn).
   - **Knowledge Base Directory** should end with the Common Learnset folder. (it is probably called Common.)

   You can change these settings only if you are an Administrator, Supervised Learning Verifier, or Learnset Manager (see section 6.4 SETTINGS – SUPERVISED LEARNING (ADVANCED VERIFIER) for configuration).

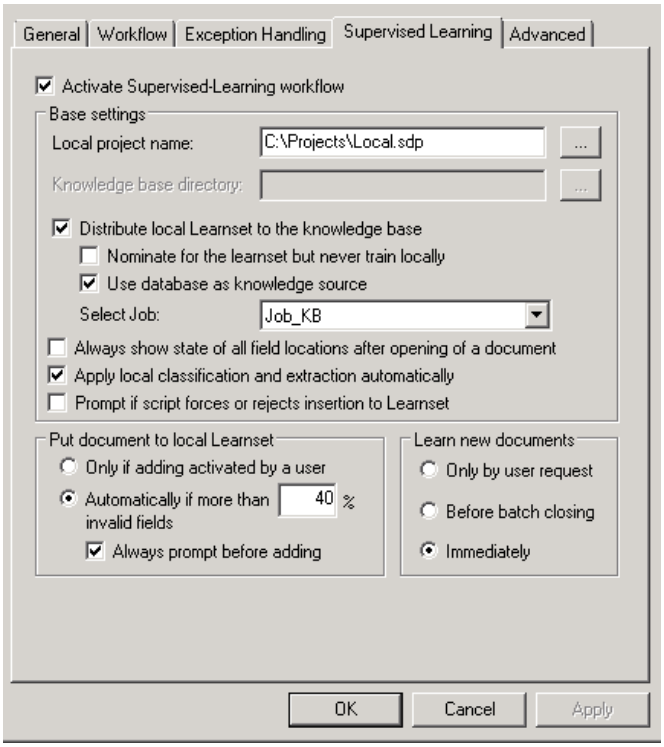


Figure 10-3: Supervised Learning Tab in Verifier Settings

2) Launch the Learnset Manager module from Verifier by clicking on  on the toolbar.

3) On the Learnset Manager Toolbar click the *Settings* button  .

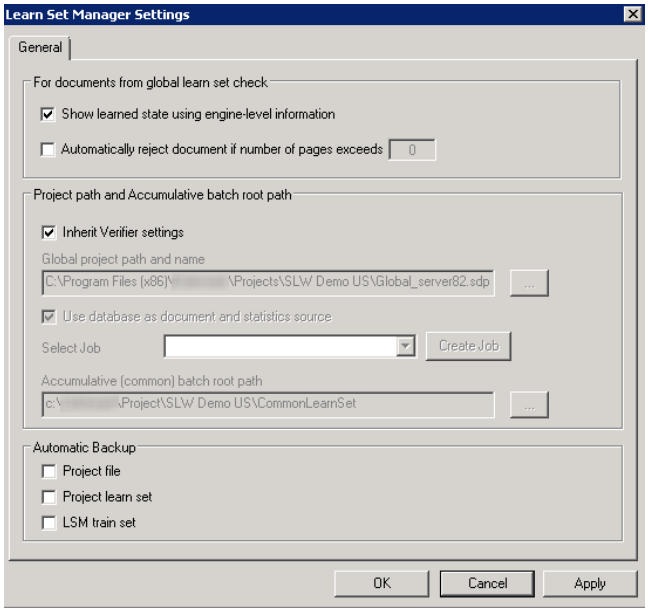


Figure 10-4: Learnset Manager Settings Tab

For document from global Learnset check:

Show learned state using engine-level information – 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.

Automatically reject document if number of pages exceeds… - Documents with more pages than specified in the appropriate field will be prevented from being added to the Learnset

Project path and Accumulative batch root path:

Inherit Verifier settings – If this option is checked, the settings made on the Verifier settings tab will be applied, and the options below will be grayed out.

If you want the LSM to us a different Global project, or a different job containing data, then uncheck *Inherit Verifier settings* option, and populate the options below as to your needs:

Use database as document and statistics source – Perceptive Intelligent Capture core information can be stored in the Perceptive Intelligent Capture database.

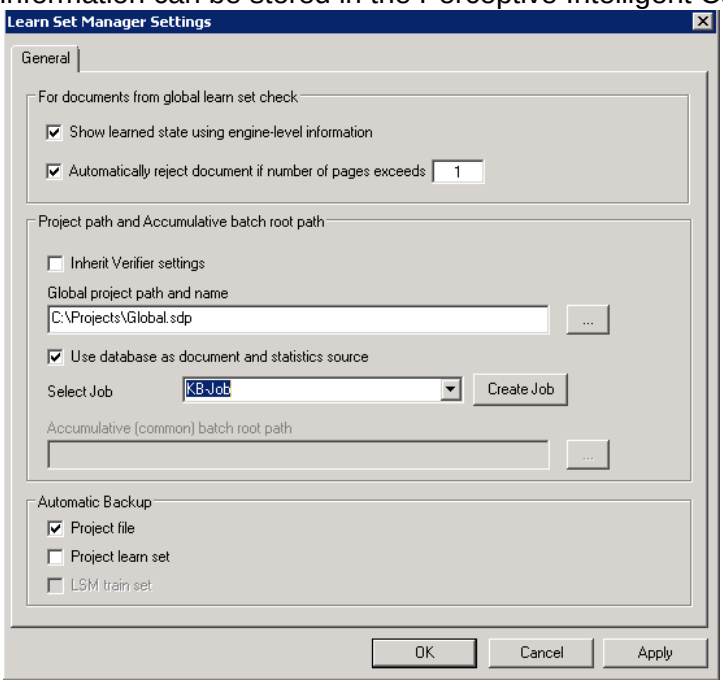


Figure 10-5: Configuration without inheriting Verifier settings

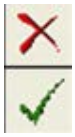
Note: The file system functionality is still supported.
Select Job: you are able to select the desired job from the *Select Job* dropdown list if you have selected the Database as your source.
Note: To use the option "Use database as knowledge source" you need to have a job for a Common Learnset in the database. If this Common Learnset job is not available, a database 'job' must be created from RTS with Common Learnset folder as the Batch Root.
Accumulative (common) batch root path – Here you can set the path to the batch root. This option is not available when using the Perceptive Intelligent Capture database.
Automatic backup: Select here the files you want to be backed up automatically:
Project file
Project learn set
LSM train set

10.4.3. Working with Common Learnsets

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

1. When you launch Supervised Learning from Verifier, the *Accumulated Documents Browsing Mode* should appear by default. The *Accumulated Documents Browsing Mode* is the one you use to work with Common Learnsets. The *Accumulated Documents Browsing Mode* is activated when you click the *Switch to Accumulated Documents Browsing* button or from the *View* Menu >*Accumulated Documents Browsing*.
2. In the *Accumulated Documents Browsing 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 to activate 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 *Accept*. This marks the document for learning as the first step for promotion into the Common Learnset (You could also click *Reject* to eliminate the document from being considered for the Common Learnset.).
6. Select another document from the batch by accessing the *Learn Statistics Panel* 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 have verified all the documents you need to verify, click *Learn*. This promotes all the accepted documents to the Common Learnset. Notice that the *Learn Statistics* window for the Local Learnset 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 Learnset Manager will proceed just like it would be in Verifier itself (See section 8.7 CHECKING ENTIRE BATCHES).

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 have selected the new class, press *ENTER.*

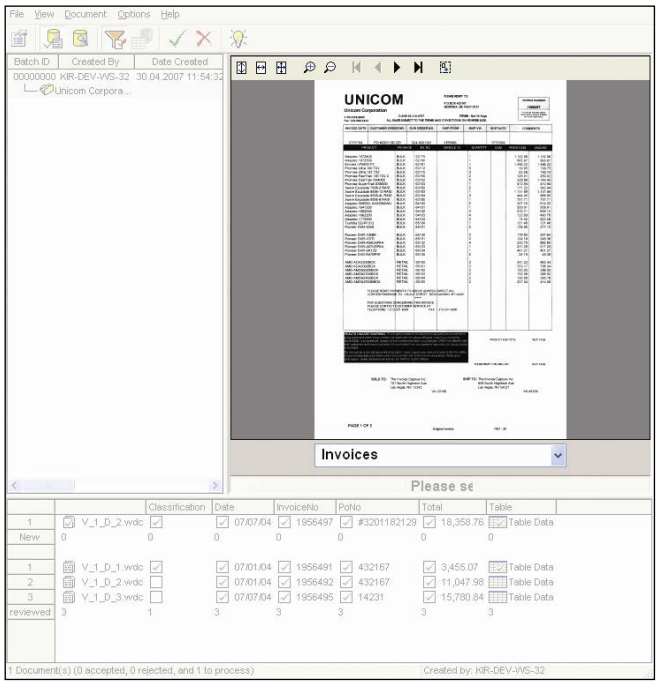


Figure 10-6: Assigning a document to a new class.

Note: 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 Learnset.

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 Learnset to the corresponding batch in the Global Learnset.

If you activate the "*Automatically reject document if number of pages exceeds…*" check box, documents with more pages than specified in the appropriate field will be prevented from being added to the Learnset. The user will be notified when a document exceeds the number of allowed pages, and the process will continue by choosing *YES* in the pop-up window.

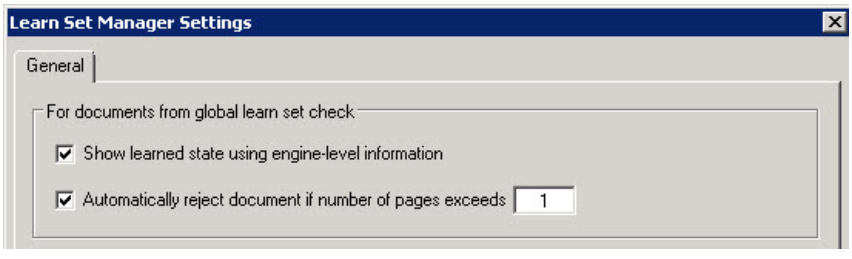


Figure 10-7 Lean Set Manager Settings

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

10.4.4.1. Description

The Learnset Manager tool of Perceptive Intelligent Capture Verifier can now sort by vendor name across multiple batches produced by different local supervised learning Verifiers.

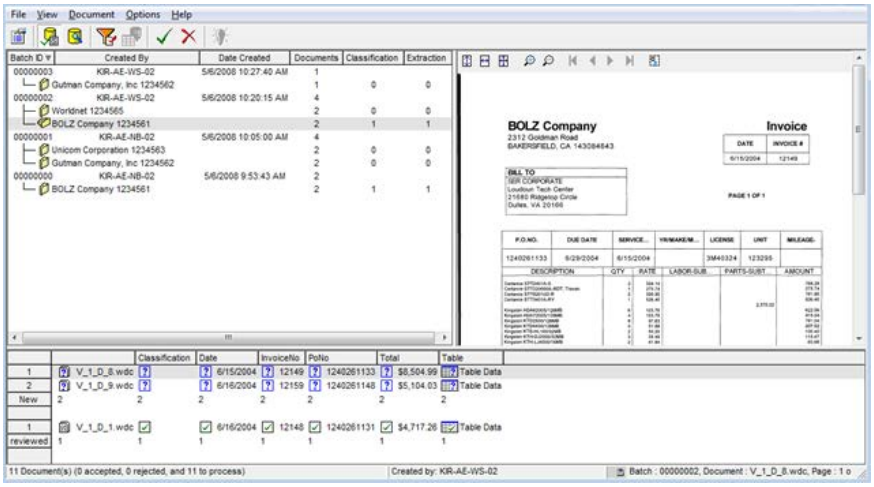


Figure 10-8: Learnset Manager tool – sorting by vendor name

By selecting the *View* \ *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.

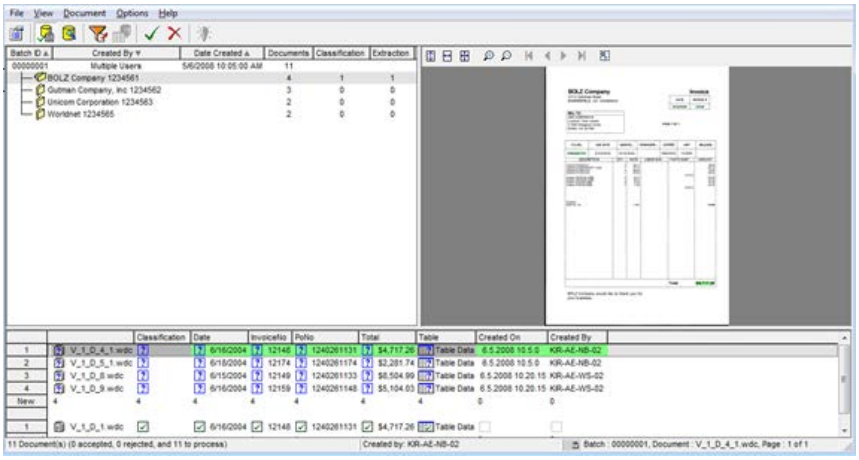


Figure 10-9: Learnset Manager screen – "Created On" and "Created By" data fields

Note: In the Global Learnset Browsing mode Sort batches by Vendor and Sort batches by Date options will be disabled. Both of the sorting options will be enabled only in Accumulated document browsing mode. However, in Global Learnset Browsing mode the Vendor classes will be sorted in alphabetical order. Furthermore, under Display Class Name column, class name will be shown only for the document classes that already exist in the global project.

10.4.4.2. Usage

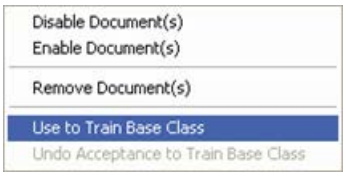
This simplifies SLW decisions as to which documents to train for a specific vendor class. With the help of the present feature, the user is now able to review all documents (for the same vendor class) created by different Advanced Verifier workstations at once.

10.4.5. Working with Global Learnsets

In the previous section, you learned how to work with Common Learnsets and to promote documents to the Global Learnset. The Global Learnset 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 Learnsets, click *Global Learnset Viewing*.

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. Right click the document if you are satisfied with it, and then select *Enable*. Select *Use to Train Base Classes*.



3) Click *Learn*.

4) Confirm that you do want to learn the document by clicking *Yes*.

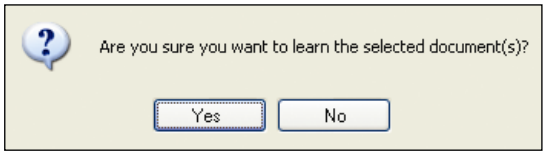


Figure 10-10: Confirming learning

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

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

10.4.6. Training Base Classes

The final milestone in creating or enhancing your Global Learnset 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 10-11) is important for keeping your Learnsets in sync with each other by enabling you to automatically synchronize them.

During the work with the Verifier application, the global project's Learnsets are constantly updated. An administrator may then wish to update the local projects (which are out of date) with a new global project. They use this functionality to add all the local projects they want to update into the main list. Then they point to the global project template (the project that will be used to overwrite all the out of date local projects, and press update to update them

To update local projects:

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

Note: Locked projects will not be updated.

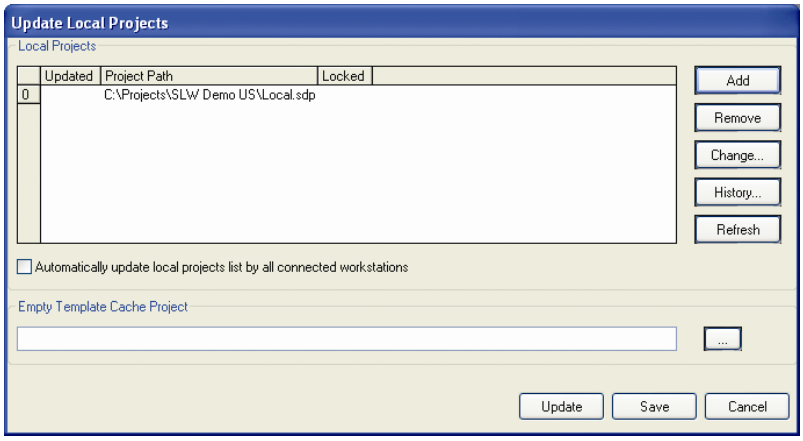


Figure 10-11: Updating Local Projects

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 *Add* and browsing to the project.)
- Remove paths from the Update list (by clicking *Remove* and browsing to the project.)
- Change existing project paths (by clicking *Change* and browsing to the project.)
- View a history of all Verifier workstations that have connected to the Common Learnset. 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 Learnset.
- Refresh the list of projects to see the most recent update information about them.
- Empty the template cache.
  The '*Empty Template Cache Project*' field is used to update the local project.
- 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 information in Section 10.6 USING LEARNSET MANAGER ON SEVERAL WORKSTATIONS), and whether the project is available.

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

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

10.6 Using Learnset Manager on Several Workstations

Learnset Manager can be used simultaneously on more than one workstation thanks to the application's ability to lock projects and files. Learnset Manager supports three levels of protection that facilitate this ability:

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

10.6.1. Batch-Level Locking

Batches are locked when they are in process; it prevents several users from updating the same batch at the same time. No one else can access the batch until processing is completed and the batch is closed.

10.6.2. Tracking changes made by Supervised Learning Managers

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

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

10.6.3. Project file and Learnset 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.

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 6.3 SETTINGS – EXCEPTION HANDLING.

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

You can use the Document Mode (See section 7.2.5 DOCUMENT AREA) to get to the document. Select the document and switch to Verify Mode. Make corrections and press Enter.

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

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 *keep focus* (APPENDIX A) or you could just use a different magnification ratio.

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

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

OK, I tried to start the Learnset 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 Learnset Manager mode.
- Learnset Manager might not be enabled for the project. Again, contact your project administrator.
- A third reason might be that the Learnset Manager mode is not properly licensed. The Learnset Manager

gets its license through a Runtime Server process. If you've been able to get into to the Learnset Manager mode before, but you can't now, it may be that Runtime Server has stopped. If restarting Runtime Server does not fix the problem, your project administrator should contact Perceptive Customer Support.

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

Perceptive Intelligent Capture 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**.

Appendix A Quick Reference

Main Controls	
	Batch View
	Start Verification
	Batch Structure
	Display Properties
	Display Batch Filter Option
	Start Learnset Manager
	First Batch Page
	Previous Batch Page
	Next Batch Page
	Last batch Page
	Refresh

Verification View	
	Exception State
	Highlight all Fields
	Highlight Selected Fields
	Highlight Candidates
	Keep Focus on Field
	Keep Zoom
	Previous Page in Document
	Next Page in Document
	Rotate Image (90 degrees)
	First Document in Batch
	Previous Document in Batch
	Next Document in Batch
	Last Document in Batch

Document View	
	Fit to Height
	Fit to Width
	Fit to Size
	Zoom In
	Zoom Out
	Single Page View
	Two Pages View horizontally
	Three Pages View horizontally
	Two Pages View vertically

Learnset Manager	
	Show Properties
	Accumulated Documents Browsing
	Global Learnset Browsing

			Correct Tables		Verify Document
			Add Current Document to Local Learnset		Accept Document(s)
			Classify and Analyze Current Document		Reject Document(s)
			Learn Documents		Learn Documents

Appendix B Index of Figures and Tables

Figures

Tables

Glossary

Accumulative Learnset	The Common Learnset.
Administrator	In Perceptive Intelligent Capture, 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.
Perceptive	Perceptive contains the core technology of Perceptive's knowledge management product suites and is a core component of Perceptive Intelligent Capture. Perceptive analyzes text from any media type. It uses artificial neural network techniques to automatically classify structured and unstructured documents and extract meaningful information from them. A neural network must be trained before you can use its ability to categorize at high speed – it must learn. The method of learning is similar to the way humans learn: It is purely sample-based. The major benefit is that after being trained, Perceptive can handle information that is similar to the samples without programming or extensive rule setting. Perceptive can operate at high speed and can be implemented on parallel hardware to further enhance performance.
Brainware Table Extraction	An extraction method that facilitates interactive table training.
Candidate	Set of possible values for a field.
Child class	A class spawned by a parent class. See also base class and parent class. Also called a sub-class.
Class	A set of documents that are grouped by common content. Each class usually has a mnemonic name that describes its contents from the user's point of view.
Classification	The process of assigning one or more classes and corresponding confidence values to one or more unknown documents.
Common Learnset	An accumulation of Local Learnsets.
DocClass	A parent document class.
Document	Any electronic file mainly consisting of ASCII text. 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 Perceptive Intelligent Capture, document export releases the documents so that they are no longer managed by the software.
Extraction	The process of automatically finding specified information within a document and writing the information to data fields associated with the document. Extraction is used for automatic indexing.
Folder	A logical structure inside a batch for coherent documents. For example, a folder may consist of all pages of a correspondence with many folders inside one batch.

Form	(1) A structured, standardized document that is used to support business processes. (2) A custom dialog box in a software application.
Global Learnset	A Global Learnset is the Learnset of the Global project file. It encompasses similar classes or projects. See also Local Learnset.
Importing	Bringing documents into Perceptive Intelligent Capture 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	In Perceptive context, this is a set of accumulated documents which are pending review to be added/rejected to the global Learnset. See also Common Learnset
Learnset	In classification, a Learnset is a set of documents whose class assignments are specified by the user. For each view and each class, the user must provide a sufficient number of representative documents. Similarly, in extraction, a Learnset is a set of documents whose field contents are selected by the user from a set of candidates.
Learnset Manager	A user who designs, modifies, and maintains Learnsets.
Learning	Given a view with a set of documents in vector representation and their class assignments, a neural network is created, so that the defined classes can be reproduced without error. This neural network is then used in all subsequent classification tasks.
Literal character	Normal alphanumeric characters that are not used as operators.
Local Learnset	Learnset specific to a document class.
Neural network	An artificial neural network is an application that in some ways works like a human brain. This includes the ability to learn. It consists of artificial neurons that are linked into a network of layers. The neural network can receive signals through an input layer, process it within the internal layers, and send signals through the output layer. During learning, a specified input (called a teacher signal, such as documents from a Learnset) and the desired output (such as the corresponding classes) are presented to the network together. Processing is then adjusted until the desired output can be produced from the teacher signal.
OCR	Optical Character Recognition. The reading and recognition of symbols of text from a piece of paper or a scanned image. OCR detects the symbols and converts them into characters and words that can be read electronically.
Parent class	A class with derived classes, called children.
Persistent	Permanent; something that is saved persistently is saved permanently, unless a user or process deletes it.
Project	Project files are used to persistently save custom settings for Perceptive Intelligent Capture applications. They are created in Perceptive Intelligent Capture Verifier and handed over to Perceptive Intelligent Capture 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	Perceptive Intelligent Capture's QA application.
View	A set of documents that represent at least two classes. A view is usually defined using a small set of documents that represent the domain of interest. In a view, classes compete for documents; that is, a document may only be assigned to one class within the view.

Web Verifier	Perceptive Intelligent Capture's web based extension of the Verifier thick client.
Workdoc	An internal structure representing the logical structure of a document. The Workdoc represents the data created during processing of a single document and is stored in a file with the extension *.wdc. Since the Workdoc includes all OCR and analysis results it may exceed the document file by size.

Index