



DOCUMENTS 7.2

New Feature Guide

June 2012 – ReadSoft AB

Copyright ReadSoft AB (publ). All rights reserved. The contents of this document are subject to change without notice. ReadSoft is a registered trademark of ReadSoft AB. Other product and company names herein may be the trademarks or registered trademarks of their respective owners.

Questions or comments about this document may be emailed to documentation@readsoft.com.

27062012.0337

Contents

Target audience	4
A review of DOCUMENTS 7.1	5
New in DOCUMENTS 7.2	8
Database Configuration Tools	9
Adjusted licensing.....	10
Windows authentication	12
Improved document separation.....	14
Invoice functionality	15
REPORTER integration	18
Scan Information.....	19
Knowledge Store additions.....	21
COLLECTOR integration	23
And more.....	23

Target audience

This documents is intended for:

- Those who need to create learning materials that cover this version of DOCUMENTS.
- Those who are to deliver DOCUMENTS technical training to partners and customers.
- Those who are to participate in the analysis and design of a DOCUMENTS solution (including customization development), installation and setup, configuration, tuning, or support.

Users of this document are expected to be familiar with previous versions of DOCUMENTS.

A review of DOCUMENTS 7.1

The most important new functionality introduced in DOCUMENTS 7.1 is described below.

Licensing

A new type of licensing, based on credits, is now in place. How many credits a document costs depends on the classification and extraction methods needed to process the document. *DOCUMENTS 7 Licensing Procedures* describes in detail how a DOCUMENTS license works, how to set up the license server, how to troubleshoot license issues, and more.

User and database information	
Windows user:	<Not used>
DOCUMENTS user:	Administrator
Database Server:	STH-QADB02
Database name:	fusion_rel_12145_pst_20120524
Database user name:	READSOFT\Eileen Palsson

License information	
License server:	STH-LICENSE
HASP ID:	123198008924
License type:	Normal
License clock:	2012-06-05 15:05:43
License number:	9999
Version:	71
Expiration date:	2014-12-31
Annual credits:	<div style="width: 30%; background-color: green; height: 10px;"></div>
Reserve credits:	<div style="width: 100%; background-color: green; height: 10px;"></div>
Annual refill date:	2013-02-23

You can now use a VMware workstation (or any other form of virtual workstation) as a license server.

For production managers who want to know how many license credits are used per solution, task, document type, etc., the document cost information is now included in the XML output. It is also available via the DOCUMENTS Advanced API (on the document instance level) so that you can develop add-ons that calculate the cost for different objects.

Knowledge Store

If you set up the system to extract table fields, for example in a customer order solution, Knowledge Store Service is recommended.

Using the Knowledge Store Service, the results of work done while verifying table fields is saved to a separate database and reused when verifying similar documents. This is done automatically. The only difference that Verify operators notice is that more table field data is found automatically, so that less manual work is required in the Table view.

Templates & setup

There are now packaged solutions using free-text functionality for capturing data from Danish and Chinese customer orders.

You can disable unnecessary columns in table field templates found in pre-packaged customer order solutions in order to streamline your table fields.

Working with and setting up queues has been simplified and improved.

Enhance

In the **Image** settings of a production task, you can choose to have the Enhance activity delete blank pages found in a document – either *all* blank pages in a document or just back pages.

Inspect

New ways to split and separate documents: If pages belonging to two different documents are incorrectly grouped together as one large multi-page document, an Inspect operator can split the document into two new documents by using new Split before this page/Split after this page commands.

Document boundaries are now displayed in both Thumbnail and Document views. The two-page view makes the most of this new feature by displaying a full image of the thumbnail in focus (and the next page). With document boundaries now visible in the Document view, Inspect operators working on manual classification can quickly and easily see if two pages belong to the same document.

Extract

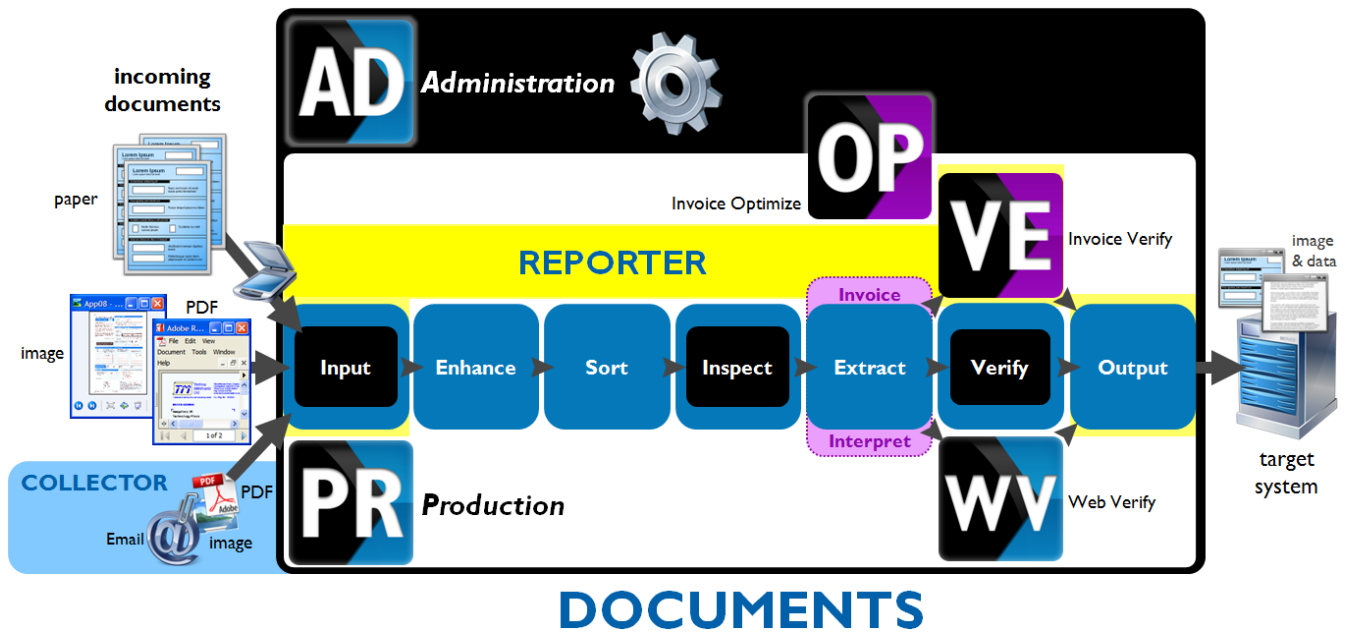
Barcode interpretation is only done one time per page/document. Also, format strings are used when attempting to classify or extract barcodes. If two barcodes are found on the same document, the system uses a format string and checks the barcode type, in order to differentiate between the barcodes, and chooses the correct one.

In addition, you can now extract barcode fields during the Input activity, making them immediately available to an add-on.

Verify

In Verify, you can now send an entire batch to a queue, as long as you specify a production task that only runs the Verify activity. Additionally, if a document belongs to a queue, the name of the queue is included in the XML output for that document.

You can also set documents to a new status, *Rejected*, by using an add-on specifically created to reject documents that do not meet the needs of the receiving system.



New in DOCUMENTS 7.2

This guide describes new features in DOCUMENTS 7.2:

- Database Configuration Tools
- Adjusted licensing
- Windows authentication
- Improved document separation
- Invoice functionality
- REPORTER integration (beta)
- Scan Information (beta)
- Knowledge Store additions
- COLLECTOR integration
- And more...

The features that are actually available in a specific system are determined by the type of license and solution purchased.

- ❗ **Important:** Limitations and known issues concerning DOCUMENTS functionality are not described in this document. Please refer to the Readme file for your DOCUMENTS 7.2 release for that information. Readme also describes minor changes in functionality that are not mentioned in this *New Feature Guide*.

Database Configuration Tools

New tools are available to create the DOCUMENTS databases and the knowledge store database. These are installed separately.



You can also use the tools to switch between databases. (Click **Save** instead of **Create**.) The tools are described in *DOCUMENTS 7 Installation Guide*.

Click to see video demonstrations:



Adjusted licensing

The licensing procedures that were introduced in DOCUMENTS 7.1 (see page 5) are adjusted in DOCUMENTS 7.2.

As before, each document specification and each layout has an associated cost that is based on its classification and extraction methods.

- To see the cost associated with a document specification, view the **Cost** settings in the **Document specification** dialog (properties):
- To see the cost associated with a document layout, open the document specification, right-click the layout, and select **Properties**. If there are page classifiers (this does not include barcode classifiers) or anchor points in the layout, these incur the same cost as **Sort – Standard**. Any classifiers in the document specification do not incur any cost in the layout.
- The **Cost** settings in both cases now look like this:

SWE Låneansökan

General
Identity
Used by
Cost
Processing
Events
Countries and languages

Costs (credit):

Base:	<input type="text" value="1"/>
Sort - Standard:	<input type="text" value="4"/>
Sort - Advanced:	<input type="text" value="0"/>
Extraction - Light:	<input type="text" value="4"/>
Extraction - Standard:	<input type="text" value="0"/>
Extraction - Advanced:	<input type="text" value="0"/>
Invoice with header data:	<input type="text" value="0"/>
Invoice line items:	<input type="text" value="0"/>
Reporting:	<input type="text" value="0"/>
Total:	<input type="text" value="9"/>

The cost, in number of credits, of each document processed using this specification. The cost is incurred at Output.

OK Cancel Help

An explanation of the “Cost” settings

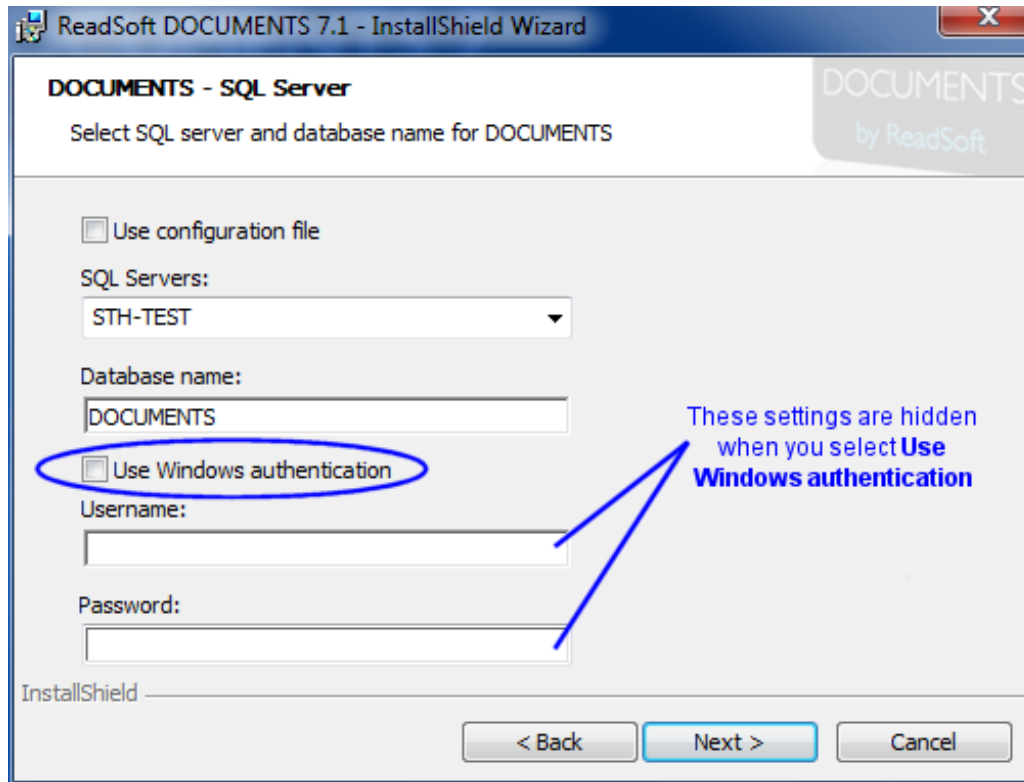
Base	The basic, minimum number of credits that any document costs passing through DOCUMENTS. Barcode separation and classification are included.
Sort – Standard	Additional “cost” (in credits) for ordinary classification (using automatic methods that are set up manually, such as anchor points, logotypes, or text classifiers). Exception: Barcode classifiers do not trigger this. (They are included in Base .) There is always a Sort – Standard cost if Sort – Advanced is used.
Sort – Advanced	Additional “cost” (in credits) for the use of self-learning text classifiers. (Use self-learning text classifier is a setting on the Classification tab of the Document specification editor .)
Extraction – Light	Additional “cost” (in credits) incurred by including at least one field to be extracted. Not incurred when all fields are entered manually. There is always an Extraction – Light cost if Extraction – Standard is used.
Extraction – Standard	Additional “cost” (in credits) incurred by including six or more fields to be extracted (that is, Enter data manually is not selected). There is always an Extraction – Standard cost if Extraction – Advanced is used.
Extraction – Advanced	“Cost” (in credits) incurred by including a table field.
Invoice with header data	“Cost” (in credits) incurred by invoice processing.* There is always an Invoice with header data cost if Invoice line items is used.
Invoice line items	Additional “cost” (in credits) incurred by invoice processing with line items.*
Reporting	If REPORTER (a separate ReadSoft product) is installed and activated, each type of <i>invoice</i> costs 5 more credits.
Total	The total number of credits that each document “costs” at the point of output. Documents rejected by the target system do not incur any cost.

* Invoice processing is triggered by using document templates that are supplied for this purpose.

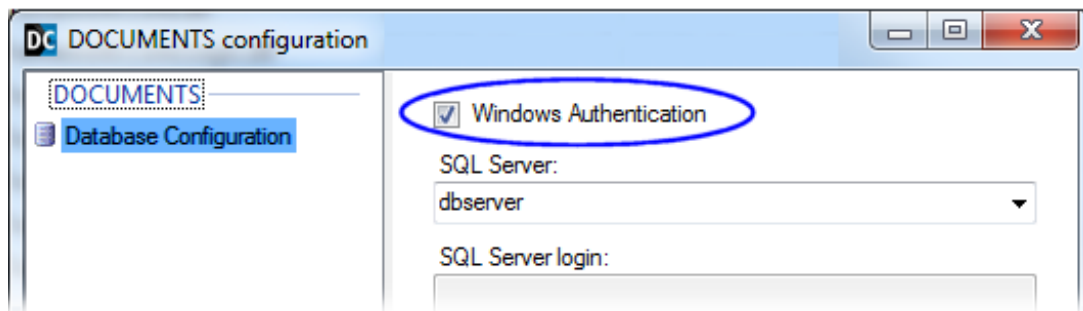
Windows authentication

You can now access DOCUMENTS databases using Windows authentication. This must be specified in three places:

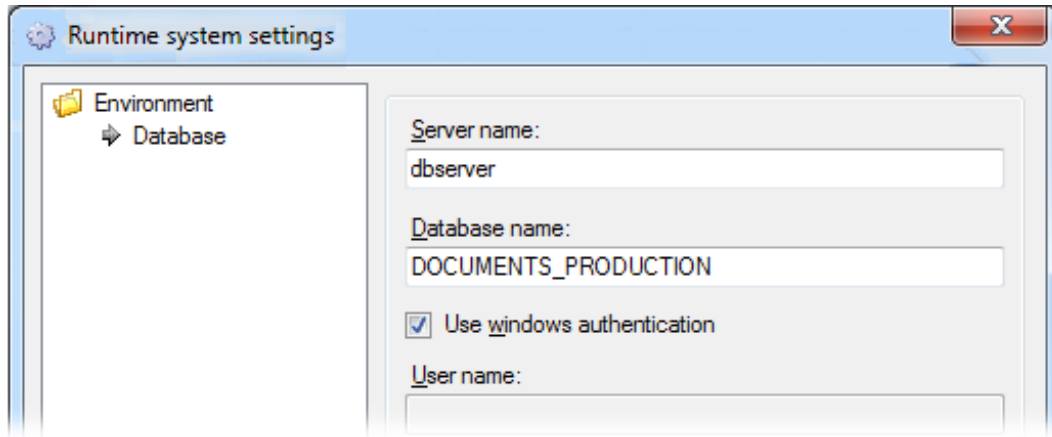
- During DOCUMENTS installation:



- When you set up the databases:



- In the Administration module, in the **Runtime system database settings** dialog:



(Select **Configuration** > **Runtime system settings** to access this dialog.)

- If you choose the other option (specifying a database login and database password, which was previously the *only* option), *the password is no longer encrypted in the Windows Registry. If you need that security, use Windows authentication.*

For more information, please refer to *DOCUMENTS 7 Installation Guide*.

Improved document separation

Documents can consist of these types of pages:



First page (green), additional pages (gold), last page (red), and appendix.

Document specifications have different types of page logic that tell the program how pages in documents are organized.

Now you can use the “fixed page logic” to separate documents with only one page or documents with a definite number of pages. You can specify classifiers and fields on every page.

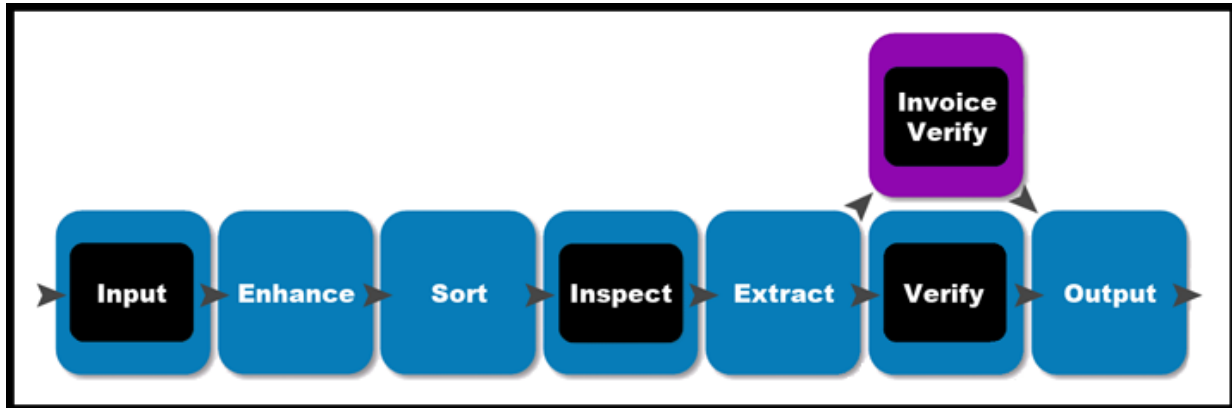
If a page classifier matches on the first or last page in a document specification with fixed page logic, a document with as many pages as the fixed-page document will be separated. This reduces manual work in Inspect when documents with a fixed number of pages are mixed with documents with an unknown number of pages.

For example, if you set **Number of pages** to 3, you can specify classifiers and fields on all three pages.



Invoice functionality

Invoice production (with Invoice modules from INVOICES) was introduced—as a beta version—in DOCUMENTS 7.1 R1. You can set up your system using the same invoice profiles as are available in INVOICES 5-6, and then use Invoice Verify to verify your invoices.



Document flow in DOCUMENTS 7.1.

Other parts of the process, such as input, extraction, and output are handled by the DOCUMENTS Production module. The Invoice Interpret module runs in the background during Extract.

What's changed in DOCUMENTS 7.2 is that:

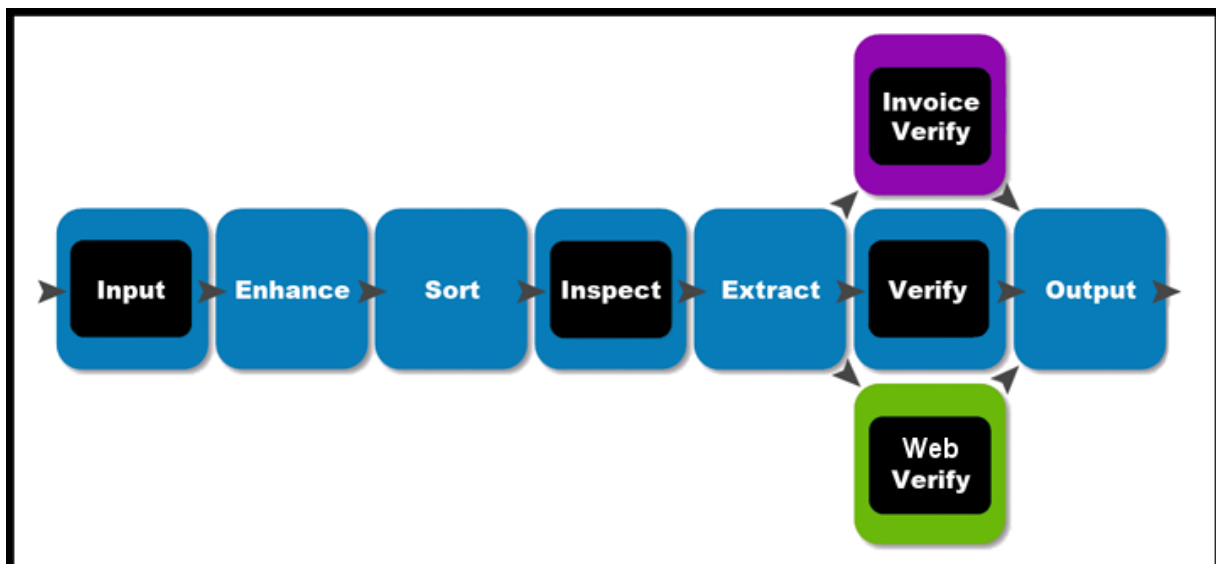
- Invoice functionality is now fully supported – with the exception of the limitations and known issues listed in the Readme file.
- Web Verify is now supported (beta – see separate section on page 16).
- Handling of invoice appendices has changed (see separate section on page 17).
- The **Process log** and **Production history** have been removed. Instead, the data is now displayed in the **System audit trail**, **Document audit trail**, and **Document searches** dialogs. This includes invoice ID, as well as user remarks that Invoice Verify operators added to invoices.
- If you need to create user remarks via the API, the new AddUserRemark method in the Invoice API replaces the now obsolete AddAPIProcessMessage method.
- If an Invoice Verify operator selects a new last page for an invoice, a dialog is now displayed that offers the option to re-extract the invoice. The new InvoiceReinterpretOnSetLastInvoicePage event in the Invoice API gives plug-ins access to suppress or modify this behavior.

- The Canadian country profile has been updated.
- There is a Chinese invoice template.
- There is a Hebrew solution template for processing Hebrew invoices (see page 23 for more information).
- Handling of invoice appendices can be adjusted by a new setting in `Eilocal.ini`, `DoPageIdentificationInInterpret`. (See description below.)

Web Verify

- Note:** Web Verify is to be considered a prototype, or “beta” functionality. It does not yet have “General Availability” status, as more early adopters are needed to secure the functionality.

Web Verify provides a complement to using the Invoice Verify module to verify invoices that are based on any of the document specification templates called "Invoice..." (for example "Invoice USA 6.0 PO"). Verify operators use a browser (Microsoft Internet Explorer is recommended) to log on and connect to the DOCUMENTS system.



Document flow in DOCUMENTS 7.2.

Web Verify must be installed separately. For requirements and installation instructions, refer to *DOCUMENTS 7 Installation Guide*. For help while using Web Verify, click the **Help** button.

If you use Web Verify:

- Do not depend entirely on Web Verify for production. Be prepared to use Invoice Verify, as well.

- There are a number of known problems (for example those listed under “Web Verify issues” in the Readme file), but none that result in losing invoices.
- Web Verify has certain requirements and limitations. Please refer to *DOCUMENTS 7 Installation Guide* for details.

Handling of invoice appendices

A new setting in `Eilocal.ini`, `DoPageIdentificationInInterpret`, controls whether the system automatically detects appendices during the Extract activity.

If your DOCUMENTS 7 solution handles invoices with appendix pages, there are two strategies that can be used:

- All appendices are set during Sort / Inspect (`DoPageIdentificationInInterpret=0`).

This mainly involves right-clicking images in Inspect and selecting Set as appendix, to differentiate those pages from invoice pages. It is also possible to set up automatic detection of appendices using page logic settings.

If an Invoice Verify operator reroutes an invoice using the This is not an invoice command, the Inspect operator can change what pages are invoice pages. (This is also true of the strategy described below. However, in that case Extract then overrides the selection of appendix pages done in Inspect.)

- Appendices are automatically found and set during Extract using Invoice Interpret, which runs in the background (`DoPageIdentificationInInterpret=1`).

This is very similar to how INVOICES 5-6 handles appendices, and this is the default setting in DOCUMENTS 7.

As mentioned above, in this case Extract overrides the selection of appendix pages done by an operator in Inspect.

Regardless of which alternative is used, you can—in Invoice Verify—change what pages are appendix pages using the **Set as last invoice page** command. And you can always see in the output which pages were set as appendix pages, just as you could in DOCUMENTS 6-4.

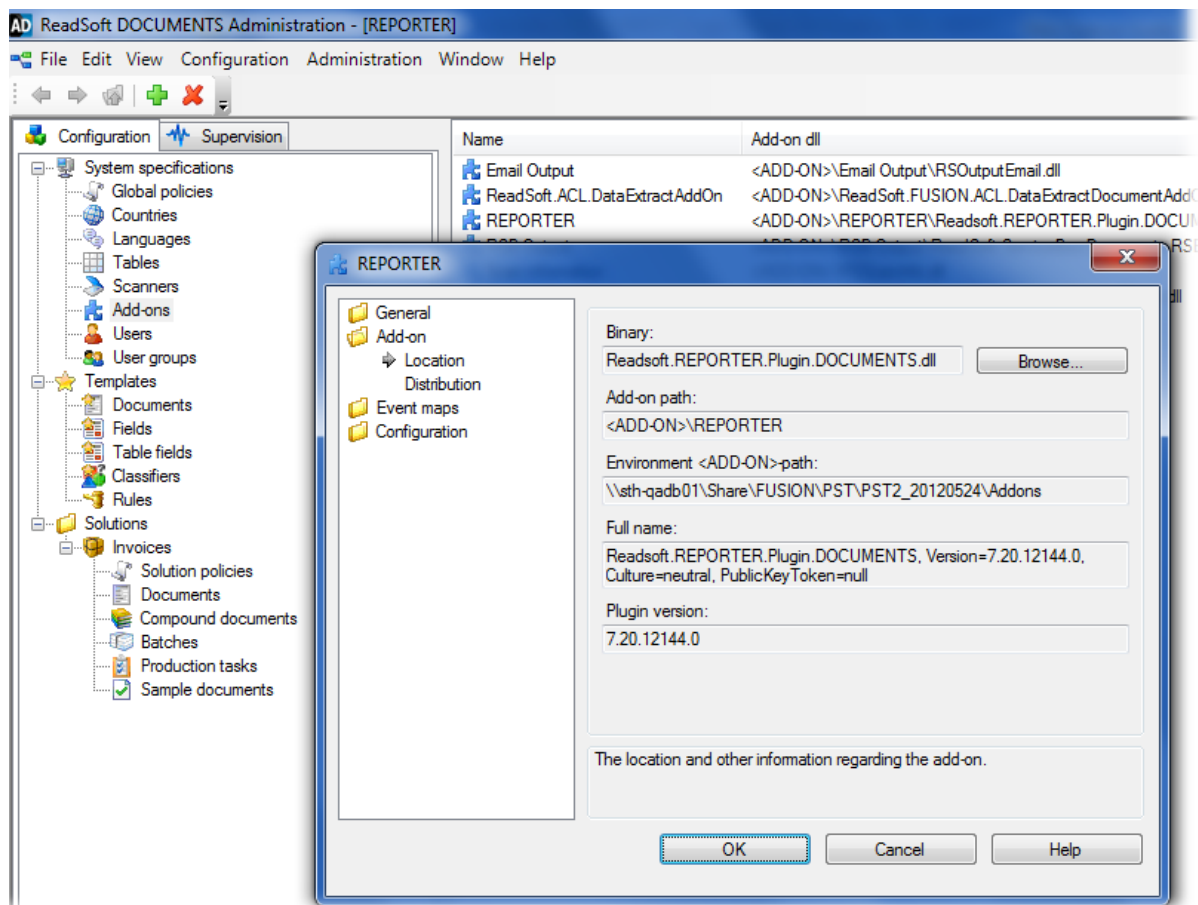
Please refer to *DOCUMENTS INI File Help* for further details about the `DoPageIdentificationInInterpret` setting.

REPORTER integration

Note: REPORTER is to be considered a prototype, or “beta” functionality in connection with DOCUMENTS. It does not yet have “General Availability” status, as more early adopters are needed to secure the functionality.

You can connect your DOCUMENTS to REPORTER, which collects data about what happens in invoice solutions and makes it available to you in the form of reports. You can use REPORTER to monitor invoice processing efficiency and user performance.

The REPORTER Data Collection add-on is included in the DOCUMENTS installation. It collects data about what happens in invoice solutions and sends it to the REPORTER database. You configure it using the DOCUMENTS Administration module (see [DOCUMENTS Help](#) for instructions).



Data is collected from Invoice modules, as well as from Extract and Output. It is not collected from Web Verify, nor from any activities besides Extract and Output.

REPORTER itself must be installed and configured separately. The installation program is available on the DOCUMENTS product CD –refer to the REPORTER installation guide.

Licensing: When you use REPORTER, each invoice costs several more credits (currently 5 more credits) to process. The license counter deducts the 5 additional credits that it costs to process each document. You can see this in the **Cost** settings of each document specification and layout.

Scan Information

- Note:** Scan Information is to be considered a prototype, or “beta” functionality. It does not yet have “General Availability” status, as more early adopters are needed to secure the functionality.

The “Scan Info” add-on is included in DOCUMENTS for use with invoice solutions. It displays an additional dialog to operators before they start scanning a batch of invoices. In this **Additional information** dialog, operators can specify certain additional information for scanning (such as the buyer and the invoice type):

Additional information

Buyer selection

- Automatic detection of buyer
- Specify buyer manually

PO/GE Selection

- Automatic detection of invoice type
- Purchase order invoices
- General expense invoices

Invoice/Credit selection

- Automatic detection of credit
- Scanned documents are invoices
- Scanned documents are credit notes

Batch Information

Number:

Tag:

Description:

User defined variables

- ListUDV:
- ListUDVNoDefault:
- Computer name:
- ControlUDV:
- ControlUDVNoDefault:
- ControlUDVQueues:

OK Cancel

This information is then assigned to all the scanned invoices, and the values are used when processing the invoices.

This add-on is particularly useful if, prior to being scanned, invoices are presorted based on their type (debit or credit; general expense or purchase order) and/or buyer.

The variables used by this add-on (see table below) are included by default. Data associated with them is available when processing the invoices, for example:

- As an extra column in the invoice inbox.
- As custom properties on the batch in DOCUMENTS.

Variable	Setting in the <i>Additional information</i> dialog	Valid values
BUYERID	Automatic detection of buyer Specify buyer manually	AUTO (Long)
INVTYPE	Automatic detection of invoice type Purchase order invoices General expense invoices	AUTO PO GE
CREDIT	Automatic detection of credit Scanned documents are invoices Scanned documents are credit notes	AUTO INVOICES CREDITNOTES
BATCHNUMBER	Number	(Long)
BATCHTAG	Tag	(Integer)
BATCHDESCRIPTION	Description	(Integer)

You can also set up your own variables. These are not available until output, where they are available in both the Batch and Document nodes. (Exception: DOCUMENTS customizations can access these variables as Private Strings on the Batch and Document objects.)

- i **Advanced tip:** You can override the default BATCHTAG and BATCHDESCRIPTION variables by setting up your own variables with those same names. For example, you could set up BATCHTAG as a date control by creating your own BATCHTAG variable on the **Controls** tab of the **User defined variables** dialog. **Tag** and **Description** are then disabled (grayed out) in the left side of the dialog but are displayed separately in the right side of the dialog (or at the bottom, depending on how many variables you set up).

Detailed instructions for these procedures are available in *DOCUMENTS Help*:

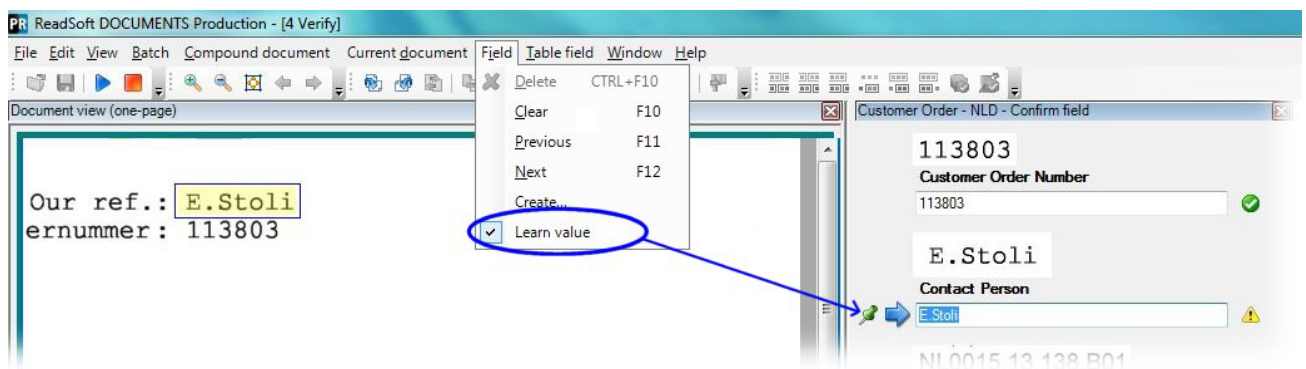
- [Configuring the Scan Information add-on](#)
- [Setting up variables for the Scan Information add-on](#)
- [Adding your Scan Information variables to the batch report template](#)
- [Exporting your Scan Information variables](#)
- [Importing Scan Information variables](#)
- [Scanning invoices: Providing the additional information required by Scan Info](#)

Knowledge Store additions

Self-learning of header fields

Running Knowledge Store results in better extraction of machine-printed character fields, date fields, amount fields, and table fields. It is recommended in customer order solutions, and in solutions that process invoices without using the Invoice modules.

You can also learn the value of a character field if you want it to be the same for all future documents that are the same variant as this one. Enter the value that you want the selected field to have for all future documents. Then select **Field > Learn value**. A pin icon shows that the field value is learned:



Please refer to *DOCUMENTS Help* topic [“Verifying fields using Knowledge Store”](#) for more information.

Enabling Knowledge Store on the solution level

You can now enable Knowledge Store either globally or on the solution level for your customer order solution. You do this in the **Persistence** settings on the global level or the solution level.

Click to see how to start the service, and how to enabled it on the global level:



Status icons

Icons at the bottom of the Production module show you the status of the Knowledge Store Service:

	<p>This combination is shown during a Verify task. It means that there is an active connection to the Knowledge Store Service and that good matching can be expected.</p>
	<p>This combination is shown during a Verify task. It means that the layout is regarded as new, normally because the displayed document is new to the system.</p>
	<p>This is what the icons should look when you run a non-Verify production task. If you see this combination in Verify, it is mostly likely because Extract was run without connection to Knowledge Store and the connection is now restored.</p> <p>However, even if the connection is restored, you cannot "learn" fields by drawing them. You have two options:</p> <ul style="list-style-type: none"> ▪ If you accept the fact that you cannot "learn" the fields, simply type in the data (or draw the fields). The fields can be learned the next time this type of document is processed. ▪ The other alternative is to delete the batch and process it again when the service is running.
	<p>Knowledge Store is not running as expected.</p>

Please refer to *DOCUMENTS Help* topic [“Checking how the Knowledge Store Service is working”](#) for more information.

COLLECTOR integration

COLLECTOR 6-2 can now be used to input the body of email messages, even if there are no attachments, into DOCUMENTS.

COLLECTOR 6-2 must be installed separately. The installation program is included on the DOCUMENTS product CD – please refer to the COLLECTOR installation guide.

The COLLECTOR installation includes the COLLECTOR add-on for DOCUMENTS, which you configure using the DOCUMENTS Administration module. Detailed instructions are available in [DOCUMENTS Help](#).

And more...

Customer order templates

A new customer order template is available for Dutch orders.

The Chinese customer order template has been adjusted and fine-tuned.

Hebrew solution template

You can now set up an invoice solution to process Hebrew invoices. These invoices follow the standard document flow and do not use Invoice modules.

When extracting Hebrew invoices, ABBYY FineReader version 10 or later is used. This is installed by default with DOCUMENTS 7.2. However, you must obtain a separate license for it from ReadSoft, and activate that license and start the service as described in *DOCUMENTS 7 Installation Guide*.

Extraction and classification with rotated text

Character fields, date fields, amount fields, and compound fields can extract text that is rotated 90, 180 and 270 degrees. This was previously only possible with layouts.

In addition, character classifiers and keyword classifiers can classify using text that is rotated 90, 180 and 270 degrees. This, too, was previously only possible with layouts.

Form functionality removed

Form functionality (Form Manager, Form Verify, and associated files and documentation) is no longer included in DOCUMENTS.

New scanners supported

New scanners are supported, and new scanning specifications are included in the installation. For details, please see [DOCUMENTS Certified Scanner List](#).

Better log files

For easier troubleshooting, DOCUMENTS log files now contain the name of the document specification and image file that were being processed when an application terminates abnormally.