Scan, PDF and OCR Add-in 2018

for SharePoint 2016

Administrator and User Guide

Version:
Server 16.0.x; Client 2018.1.2.x

Contents

Note

Features marked by sign ℙ *are available for the Professional version only*

Important

This version is packaged as SharePoint Farm Solution (wsp). You can install it on SharePoint "on-premises" only. If you use SharePoint Online or SharePoint "on-premises" with configured Apps management please, download and install "SharePoint Scanner App" from our site: http://websio.com/spscanocr.aspx

Overview

SharePoint Scan, PDF and OCR Add-in is a powerful tool for scanning and composing documents from existing images or PDF files directly into a document library.

With easy-to-use interface of the Plug-in, your employees will be able to scan documents right away allowing you to save on training your staff.

The Add-in supports:

1. OCR allows to convert image to text for more than 50 languages* ℙ

2. Properties recognition - Zonal OCR ℙ **

3. Barcode recognition ℙ

4. Automatic upload of files from the folders on the local computer to SharePoint ℙ ***

5. Scanning multiple documents in a single batch using document separator sheets or barcode separation ℙ

6. Saving pages as:

    o Single document

    o Document per page

    o Multiple documents using document separator sheets ℙ

7. Document composing from the existing image or PDF files

8. Regular and Quick Scan modes

9. PDF, TIFF, JPEG, PNG, BMP and GIF output formats.

10. Advanced compression technologies, enabling enable best document quality with smallest file size

11. All types of Windows-compatible scanners: simple desktop and professional scanners with feeders, local-connected (USB) and network-connected (IP) scanners.

12. All major browsers (Internet Explorer, Edge, Chrome, Firefox).

13. Multilingual user interface (English, Arabic, Czech, German, Hebrew, Russian)

* Supported OCR languages: ₽

English, Afrikaans, Albanian, Arabic, Azerbaijani, Basque, Belarusian, Bengali, Bulgarian, Catalan, Cherokee, Chinese Simplified, Chinese Traditional, Croatian, Czech, Danish, Dutch, Esperanto, Estonian, Finnish, Frankish, French, Galician, German, Greek, Hebrew, Hindi, Hungarian, Icelandic, Indonesian, Italian, Japanese, Kannada, Korean, Latvian, Lithuanian, Macedonian, Malay, Malayalam, Maltese, Norwegian, Polish, Portuguese, Romanian, Russian, Serbian, Slovakian, Slovenian, Spanish, Swahili, Swedish, Tagalog, Tamil, Telugu, Thai, Turkish, Ukrainian, Vietnamese
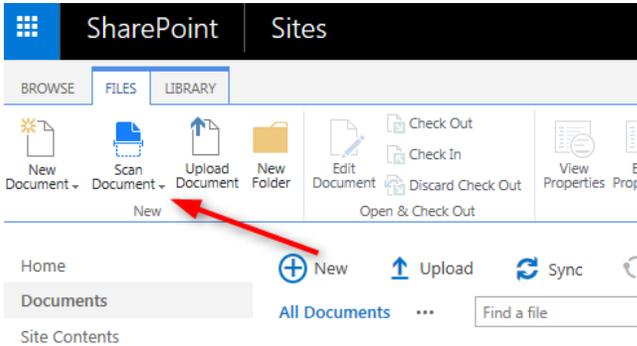

**   License for Zonal OCR Extension is required
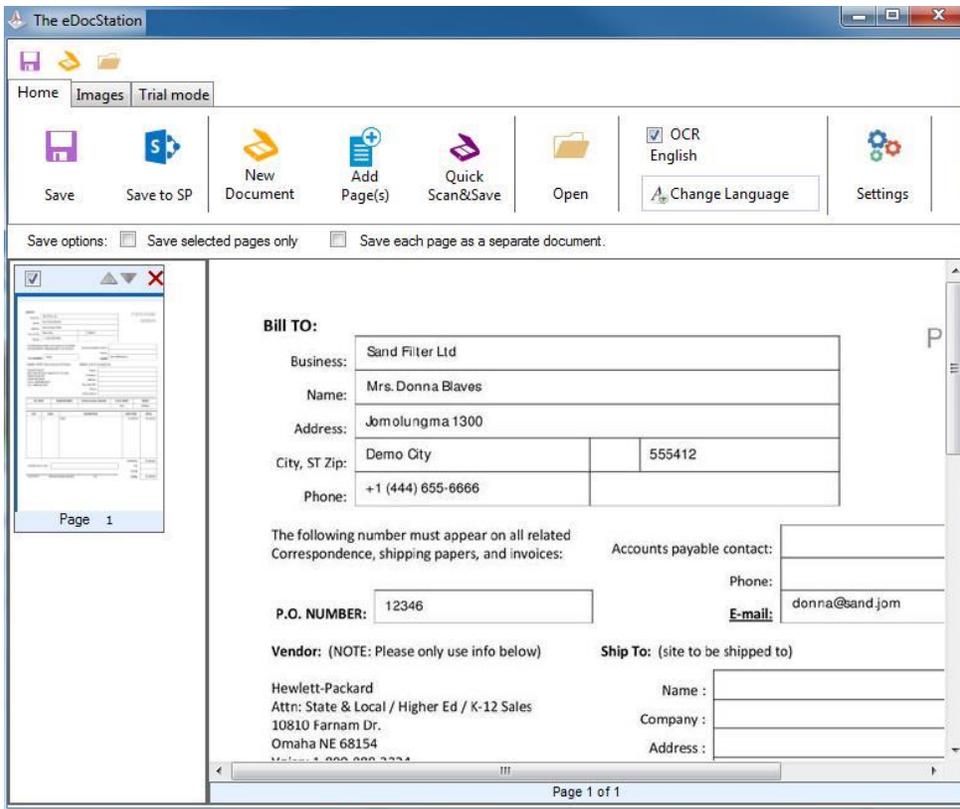*** License for Auto-Upload Extension is required

Picture 1

Scanner Add-in has native integration with SharePoint Document Library Ribbon (Optional. You can save documents to SharePoint from the eDocStation itself)

Picture 2

SharePoint Scan, PDF and OCR Add-in Client - The eDocStation

Requirements

Server Requirements

- Microsoft SharePoint Server 2016 "on-premises"

Client Software Requirements

- Windows 7 / 8 / 10

- Internet Explorer / Microsoft Edge / Chrome / Firefox

- Microsoft .NET Framework 4.5.2 +

Scanner Requirements

Local-connected and network-connected scanners have to be correctly defined on the workstation. Usually it is enough to install correct scanner drivers. Please refer to your scanner documentation.

To test that your scanner is correctly defined, try to scan from one of the standard Microsoft programs like Paint or Microsoft Clip Organizer

Installation

What do I need to install?

Installation set contains two parts: server-side and client side (the eDocStation). The eDocStation is required.

You can use the eDocStation as a standalone application without server-part and save documents into a SharePoint library:
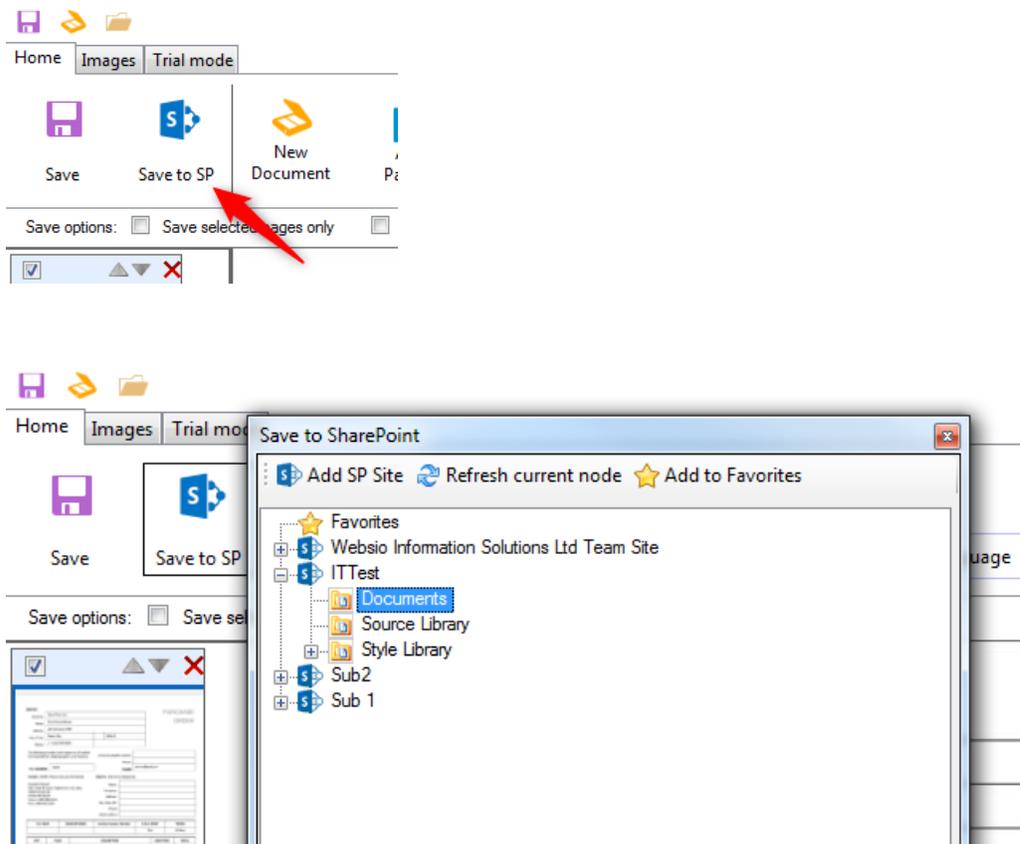
Server-side adds following functionality:

1. Start the eDocStation from the SharePoint document library and store final document to the same library

2. Use Zonal OCR (properties recognition). See section "Properties Recognition - Zonal OCR" for details

Server-side installation

3. **Important!** You must be logged on to the server by using an account that is a member of the Administrators group and SharePoint Farm Administrators

4. Unzip directory "server" from the ScannerAddin.zip file on your SharePoint 2016 front-end server (WFE).

5. **Important!** Right-click the "setup.exe" and choose "Run as Administrator" to begin the plug-in installation. Otherwise the error dialog box will alert you of the error that has occurred::

6. Repeat steps 1-3 on each SharePoint WFE server.

7. After installation is complete, open <u>top-level site</u> of the site collection where you will use the Add-in.

8. Open "Site actions" menu and select option "Site Settings"

9. Select option "Site collection features"

10. In the list of installed features, find "Websio Scan, PDF and OCR Add-in" and click "Activate".

	Websio Scan, PDF and OCR Add-in	Deactivate	Active
	Websio Send Mail	Deactivate	Active
	Workflows Aggregated set of out-of-box workflow features provided by SharePoint.	Activate	

Special case: If your SharePoint Server uses Forms Authentication,
see section "CRM & SharePoint".

Client-side installation

1. Unzip directory "eDocStation" from the ScannerAdd-in.zip file on your client machine
2. **Important!** Right-click the "setup.exe" and select "Run as Administrator" in order to begin the installation.
3. **Please pay attention!** SharePoint Scan, PDF and OCR Add-in *out of the box* arrives with English OCR language pack only. Download and install additional OCR language packs from our site: http://websio.com/ocrlangpacks.aspx
4. Windows Firewall alert may appear during installation. Approve network access for eDocLoader.exe. The eDocLoader does not access any external internet resource. It is responsible for communication with SharePoint only.

After the installation has been completed:

1. Icon of the eDocClient appears in the system tray of the client workstations:
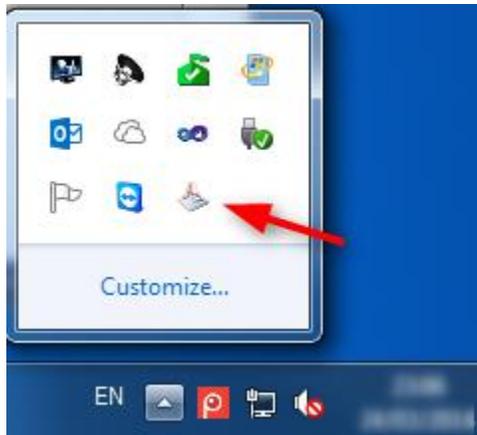
Please, do not close the eDocClient. This program is responsible for communication between client part of the Add-in and SharePoint.

2. New button appears in the Files Ribbon:
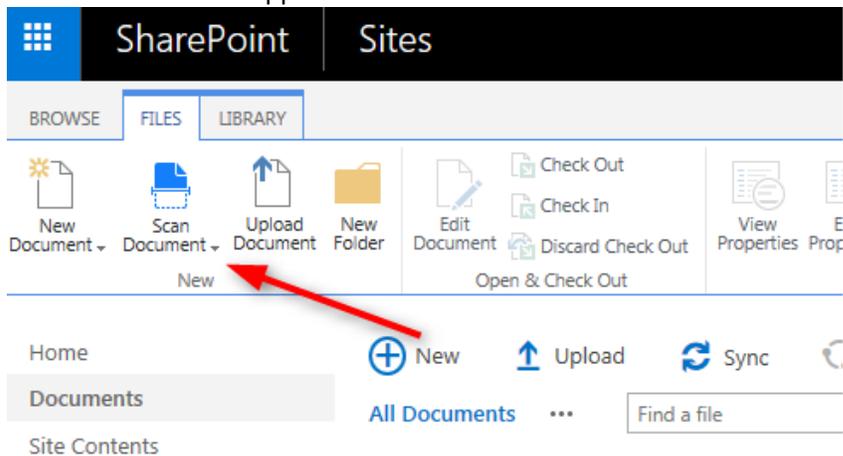
Trial version

You can switch between Standard and Professional modes during Trial period. Properties recognition may be activated in the Professional mode. See section "Properties Recognition - Zonal OCR" for details.
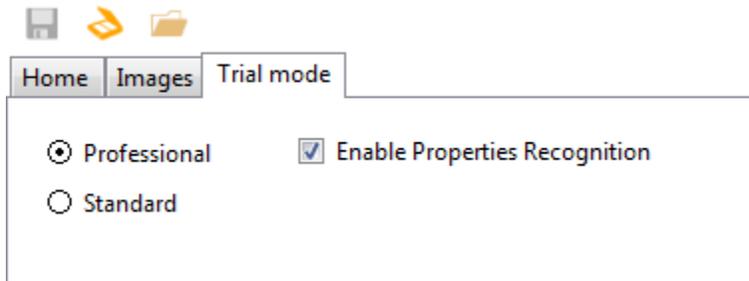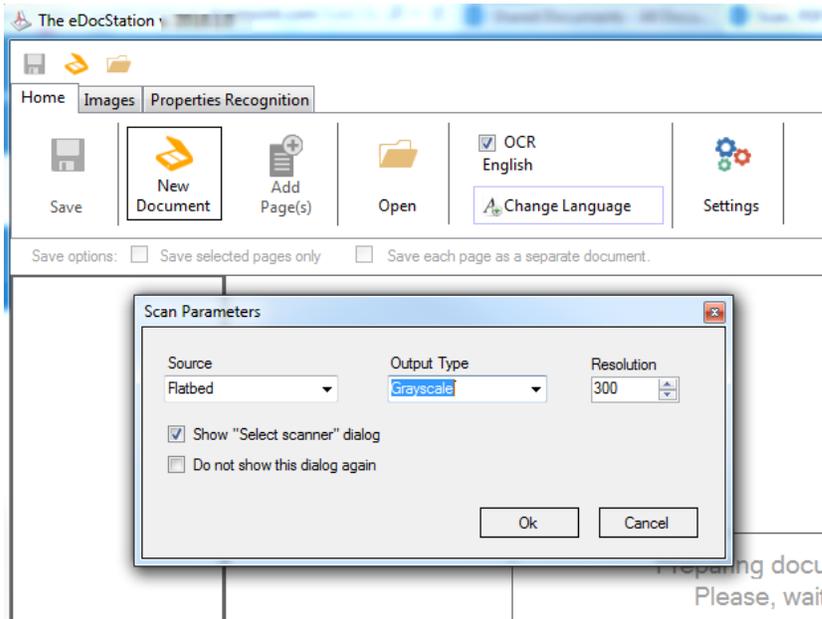
How to use

Option 1. Start process from SharePoint

Regular Scan Mode

1. Open document library view, select tab "Documents" and click on the button "Scan Document"
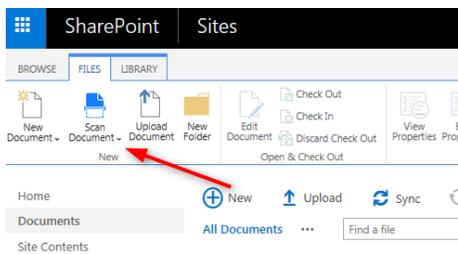
The eDocStation will be opened in the SharePoint-mode. Click on the "New Document", select scanner and scan parameters

Select default scanner and scan parameters in the eDocStation settings to avoid these dialogs.

Scanned pages:
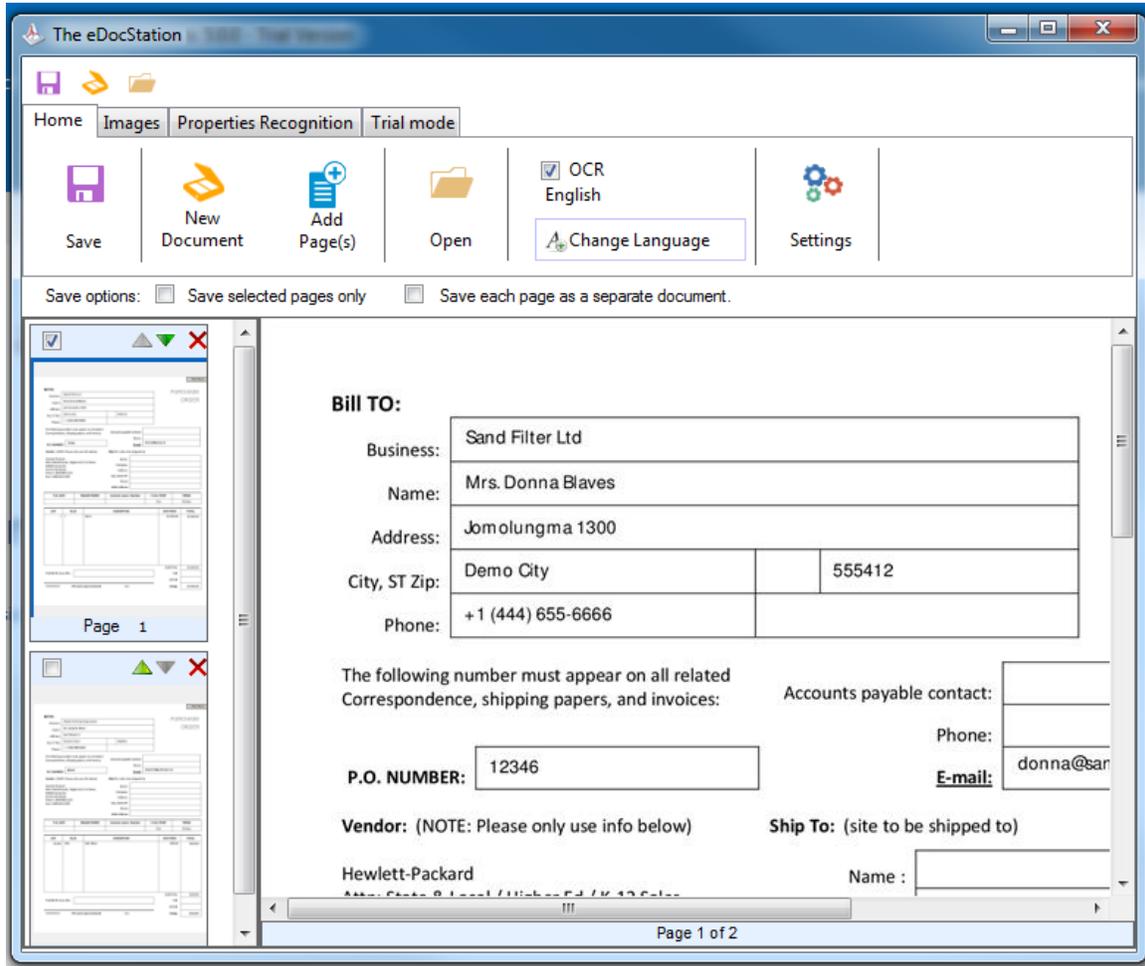
Click on the "Save" button, enter document name and select format
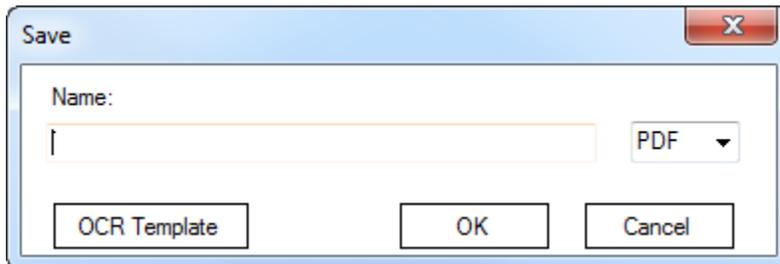
Select option "Always use auto name" in the eDocStation settings to avoid this step.

Created document will be saved in the SharePoint Document library.

Standard SharePoint properties window will be shown if document profile contains required properties.
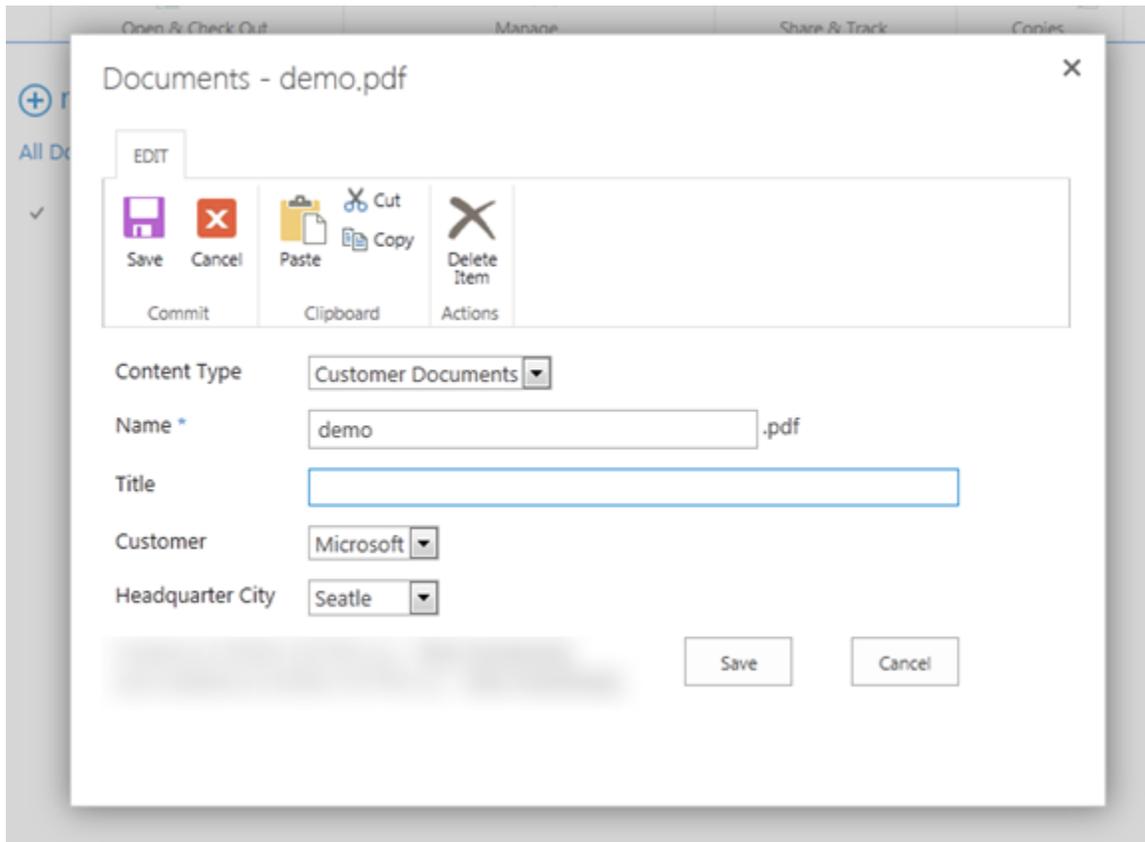
Additional save options:

1. Save selected pages only:
   1.1. Select checkbox "Save selected pages only"
   1.2. Select pages
   1.3. Click on the button "Save"
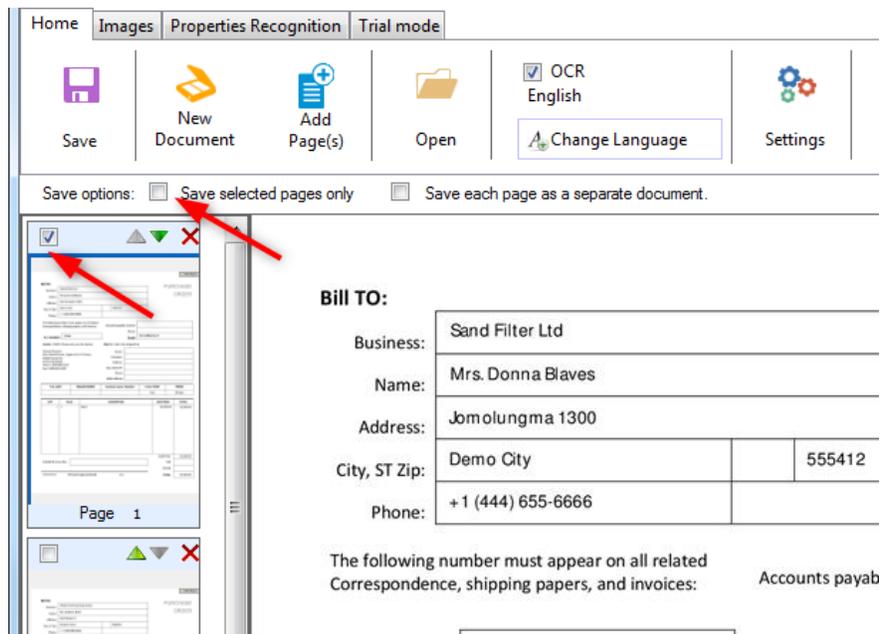
2. Save each page as a separate document:
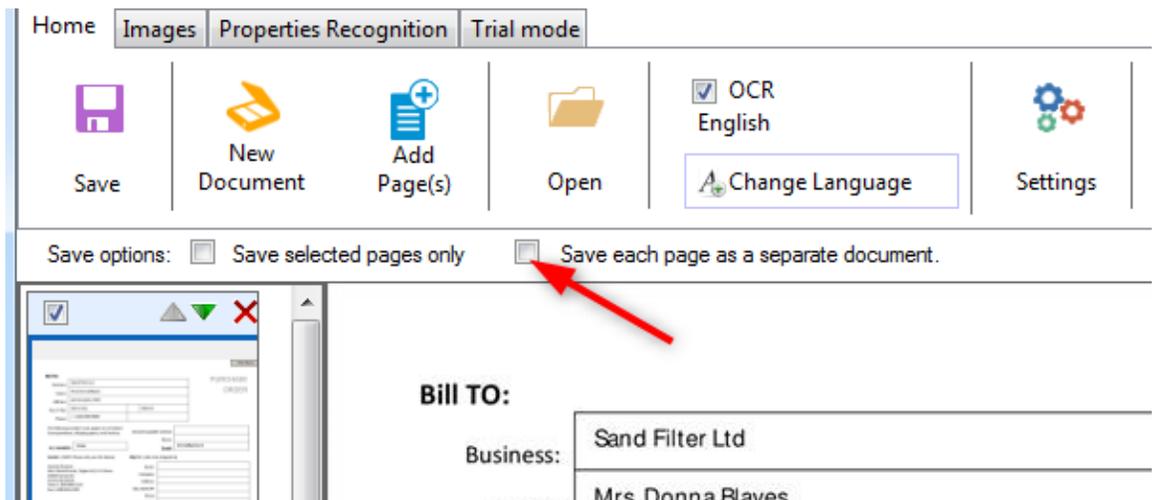
Quick Scan Mode

1. Select option "Quick Scan" from the drop-down menu.
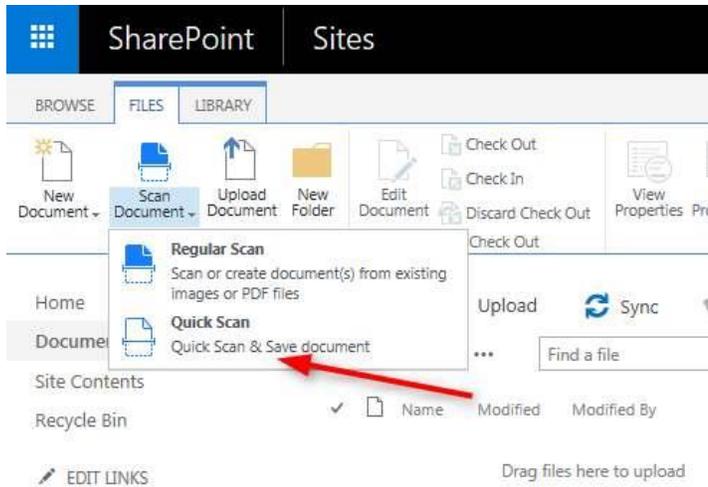
2. New document(s) will be scanned and saved automatically. The eDocStation window will not be displayed.

3. Document names will be generated automatically using prefix defined in the Quick Scan settings (see paragraph "Settings").

4. To save each page as a separate document in the Quick Scan mode, select option "Create document per page" in the eDocStation Quick Scan settings
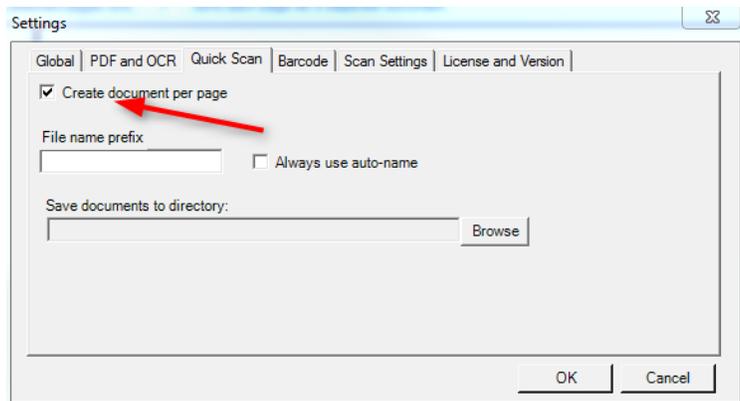
Option 2. Use the eDocStation as a standalone-application

1. Use desktop shortcut ![Websio eDocStation icon] in order to launch the eDocStation. Scan or compose document as described in the previous section "Option 1. Start process from SharePoint"

2. Press on "Save to SP"
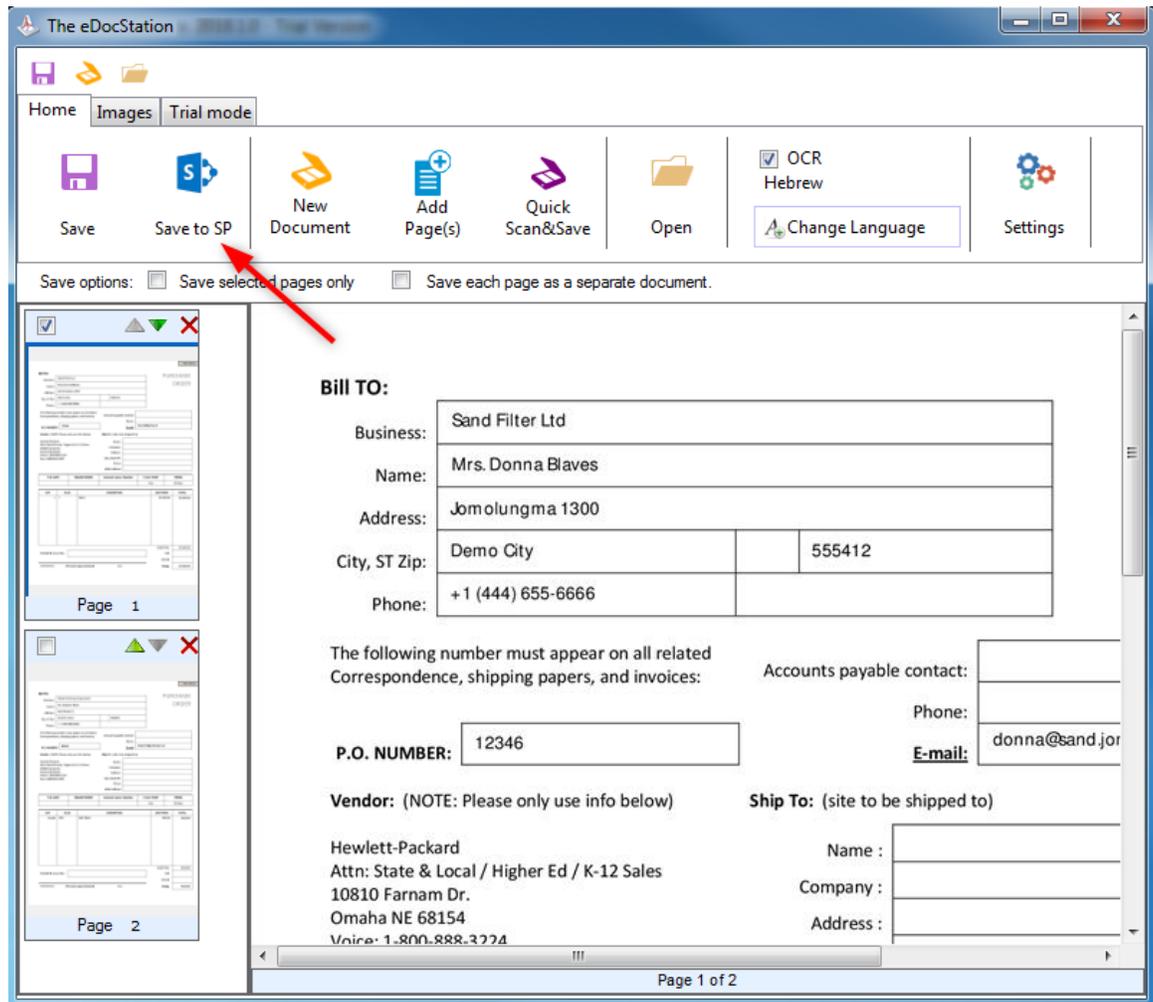
3. "SharePoint "dialog will be opened. If you're accessing this window for the first time, you'll be prompted to enter site url:
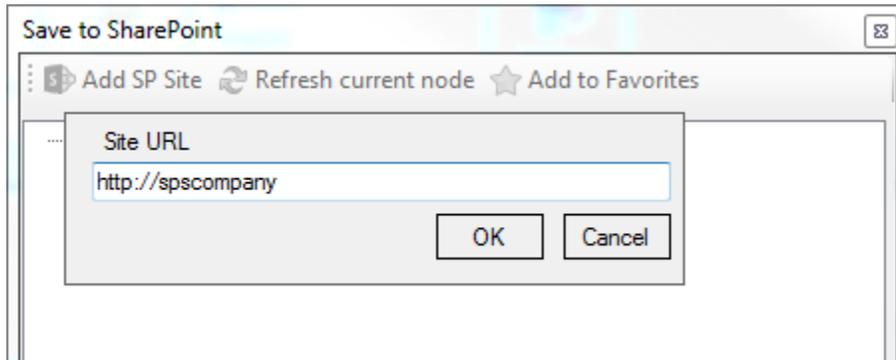
4. Wait until list of libraries will be displayed

5. Select library or folder, enter file name and press "OK" in order to save document to the selected location
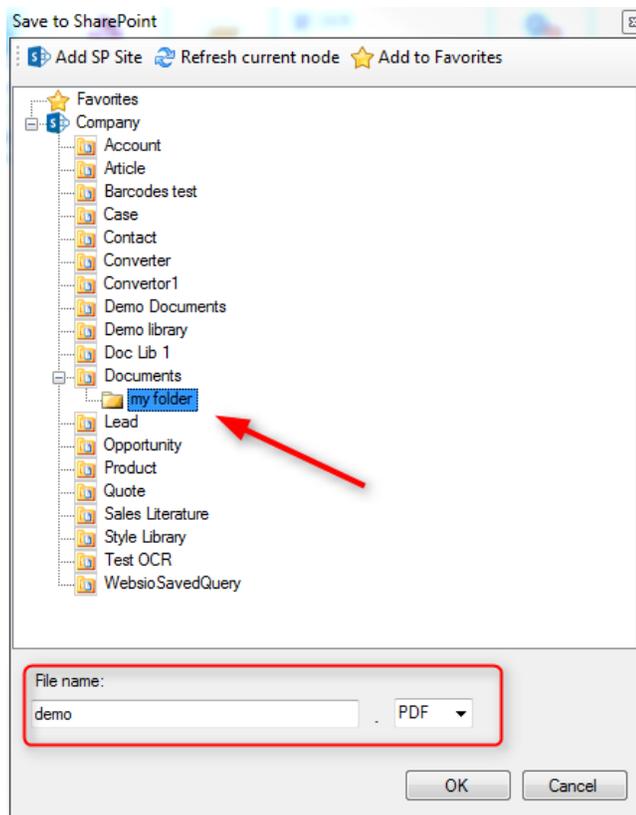
See section "Save to SharePoint Dialog" for additional information

6. Important! Select option "Display SharePoint Properties" to display document properties after upload.
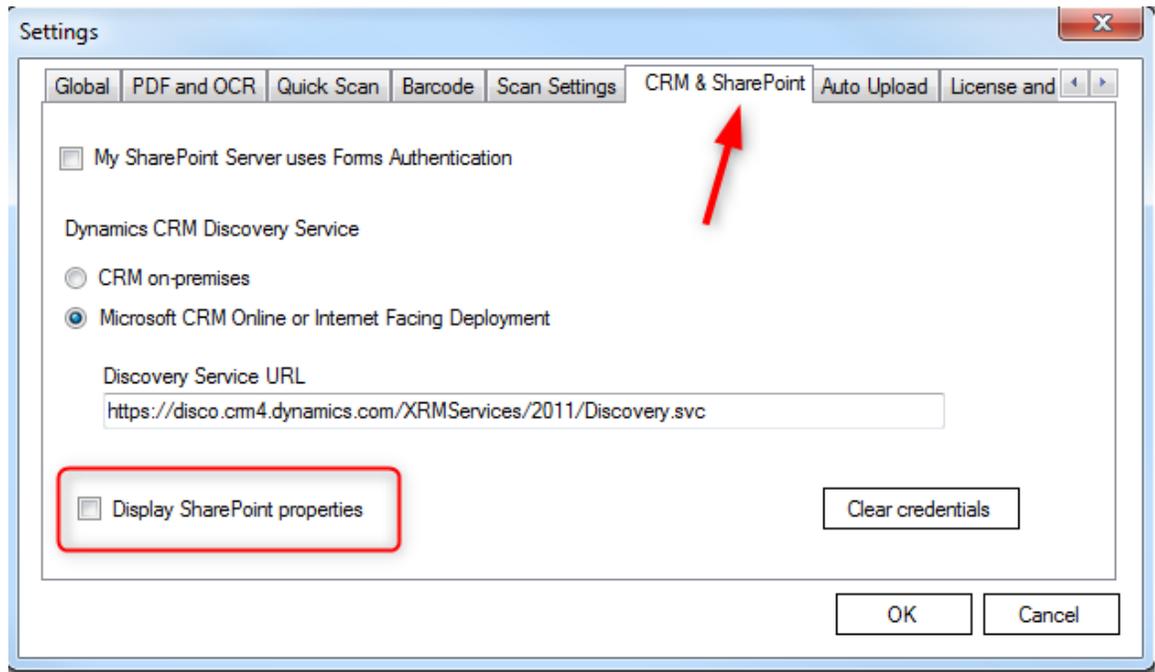
Document Separator Sheets ₽

A separator sheet allows the scan application to detect when a new document should be started. It helps the scan operators to increase productivity and organize a stack of paper into logical documents. That eliminates the need for individually scan of each document either.

1. Select option "Use Document Separator Sheets" in the eDocStation settings
2. Print Document Separator Sheet:

    a. Open Document Separator Sheet from the "Start->All Programs -> Websio eDocStation"
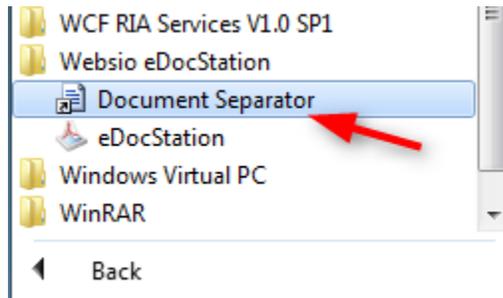
or from the eDocStation installation folder
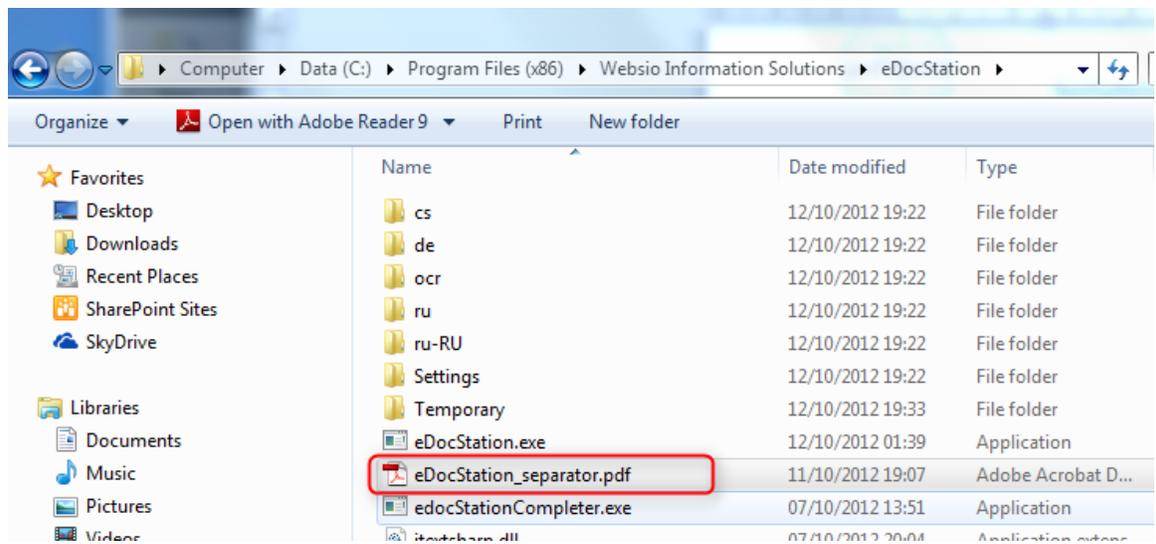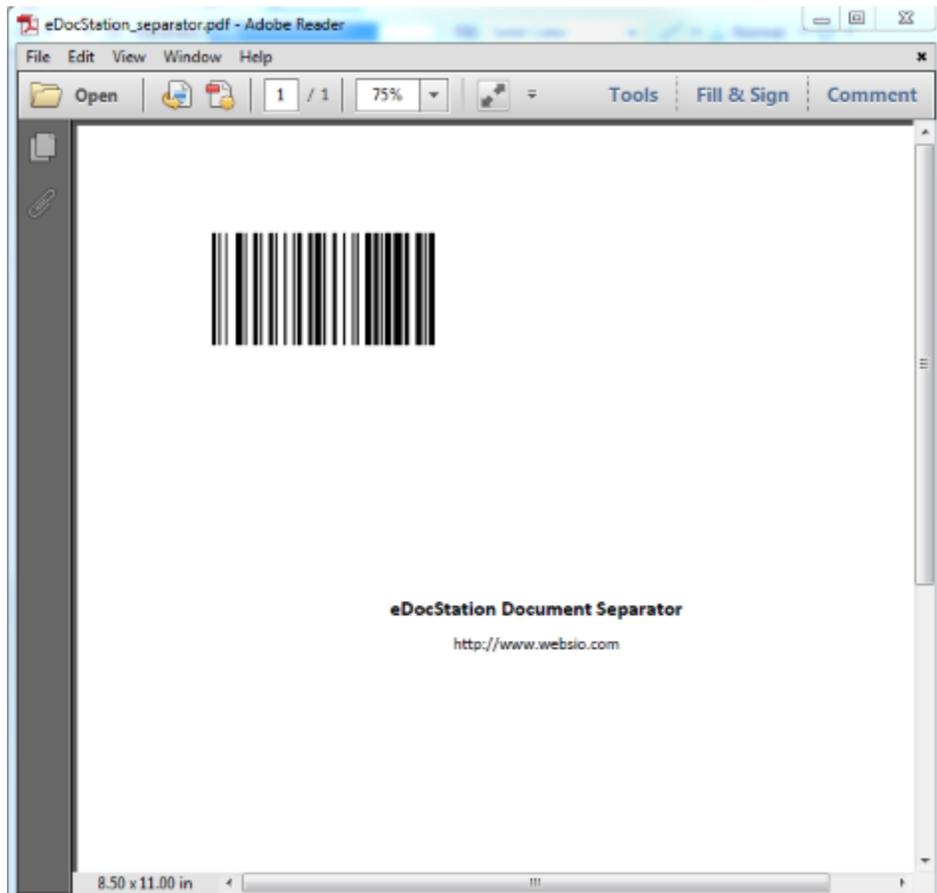
b. Print Separator Sheet in A4 format in a high quality.

3. Separate documents in the scanner feeder by previously printed sheets

4. Scan documents using SharePoint Scanner Add-in. Quick Scan mode is recommended but not required.

OCR – convert images to searchable PDF ℗

Searchable PDF (document created using OCR process) advantages:

    a.   Search documents by content using standard SharePoint search engine

    b.   Search text in the document:
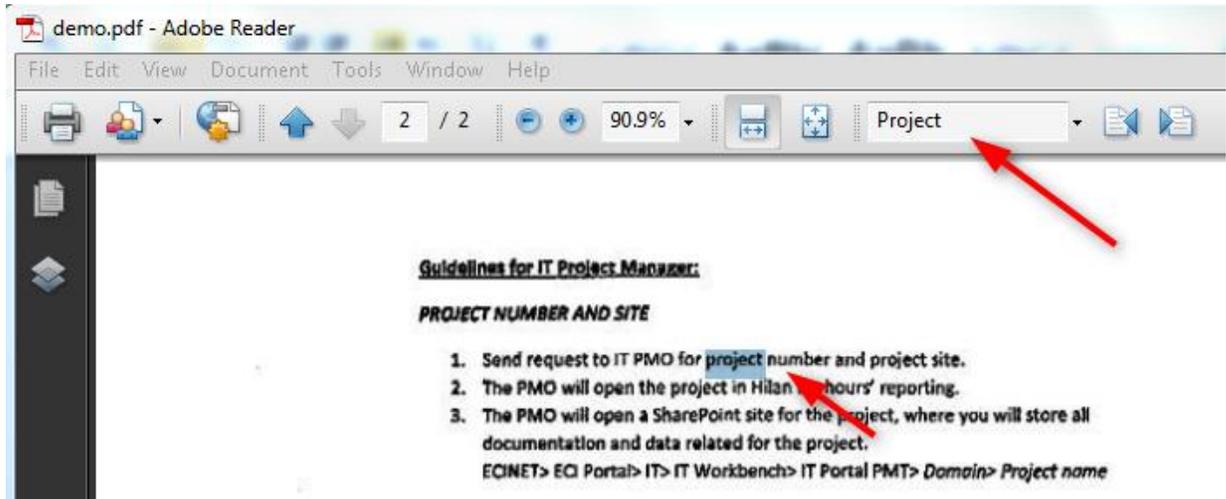
    c.   Copy text from the document:
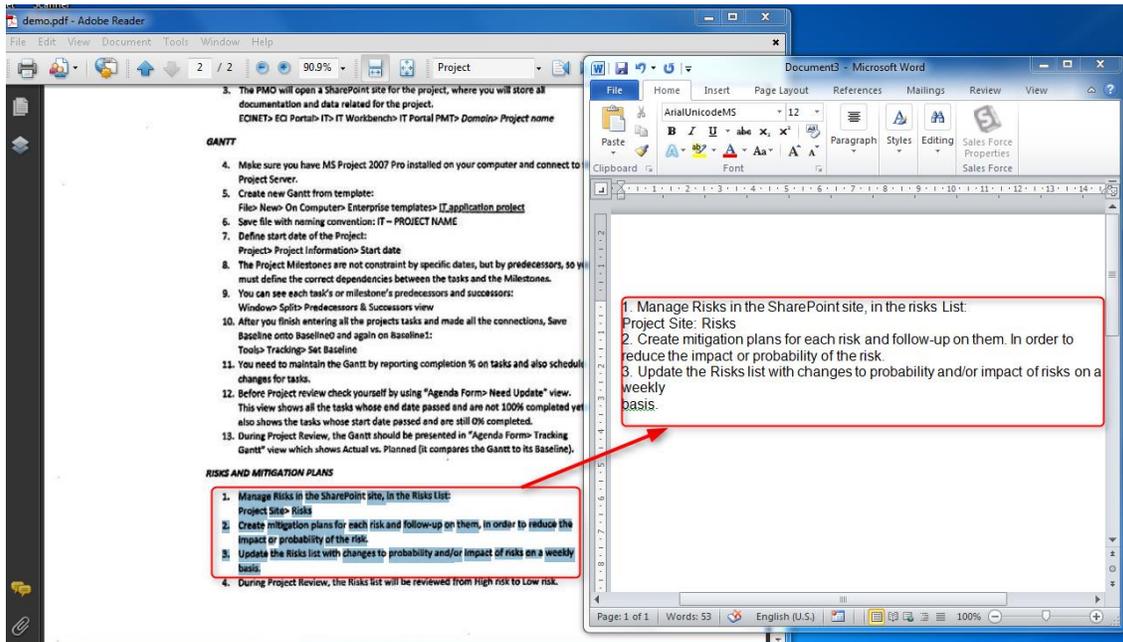
The best OCR (text recognition) quality can be reached for images scanned in grayscale with 300 dpi resolution.

Important! If you want to OCR color images, select option "I scan color images" in the eDocStation settings
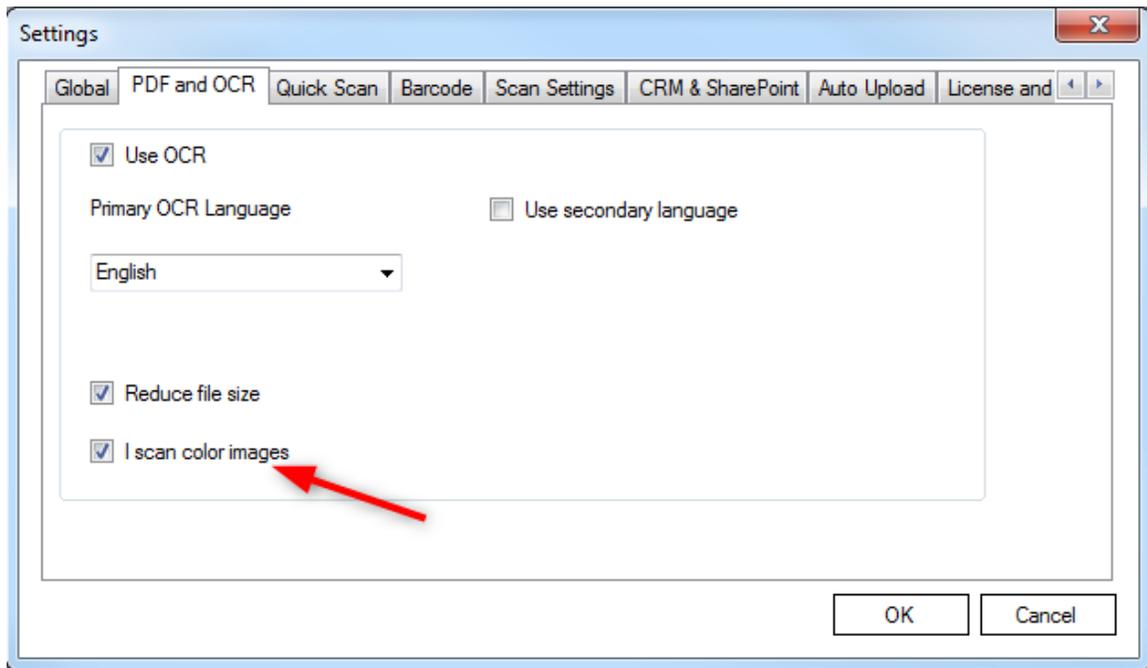
1. Select checkbox "OCR" and document languages to create searchable document.
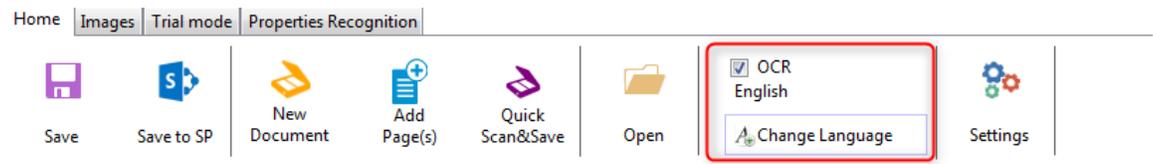
Important! Use secondary language when it is necessary - your documents have bi-lingual content. Best OCR results may be achieved by using single language.
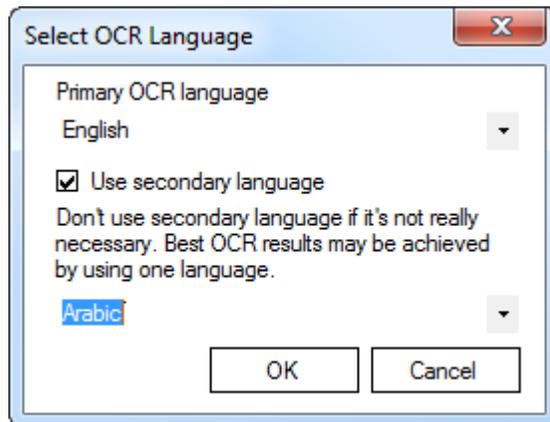
Set default OCR parameters using eDocStation settings:
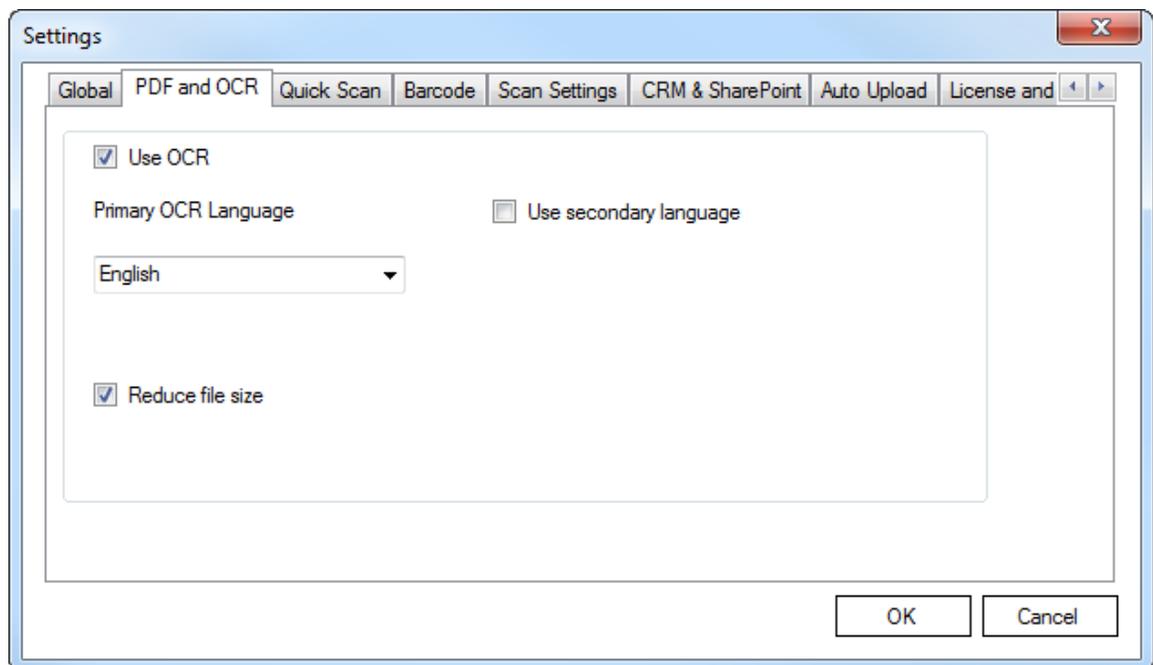
Important! SharePoint Scan, PDF and OCR Add-in *out of the box* contains English OCR language pack only. Download and install additional OCR language packs from our site: http://www.websio.com/ocrlangpacks.aspx

2. OCR process runs during document save process. The process may take some time, depending on the pages count and original images quality.

3. Our products use one of the best Optical Character Recognition (OCR) engines "Tesseract". Tesseract development has been sponsored by Google since 2006. We developed set of optimizing image procedures for best OCR recognition. These functions provide cardinal improving of the OCR results. At the same time Tesserract engine may not recognize some specific fonts. The best results may be achieved for standard Microsoft Office fonts with size from 9 to 13 px. Please pay attention, we use Tesseract OCR as-is and we cannot add support for unrecognized symbols/fonts/languages

Properties Recognition - Zonal OCR ₽

License for Zonal OCR Extension is required

This feature enables automatic properties recognition and assignment.
We highly recommend for you to watch demo video available on the product page:
http://www.websio.com/spscanocr.aspx

Examples.

Task: scan order, recognize Order ID, Customer and Amount, save document to
SharePoint and assign recognized values to document properties.

Destination document library properties:

Columns	
A column stores information about each document in the document library. The followir	
Column (click to edit)	Type
Title	Single line of text
Order ID	Single line of text
Customer	Single line of text
Amount	Currency
Created	Date and Time
Modified	Date and Time
Created By	Person or Group
Modified By	Person or Group
Checked Out To	Person or Group

Important! Zonal OCR supports text, number and currency field types only.

Example 1. Manual recognition

1. Scan document, open tab "Properties Recognition" (1) and click on the "New" (2). Document properties pane (3) will be displayed. The pane contains supported properties from the current library.
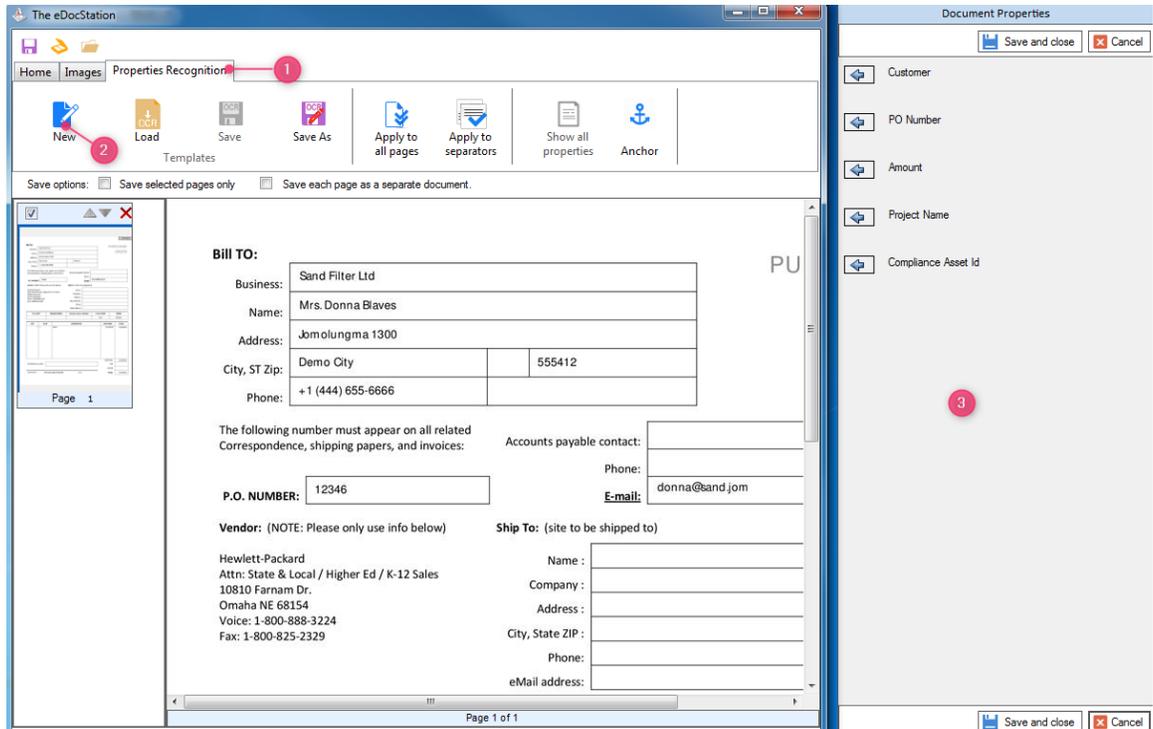
2. Click on the arrow or drag property to the scanned image surface.
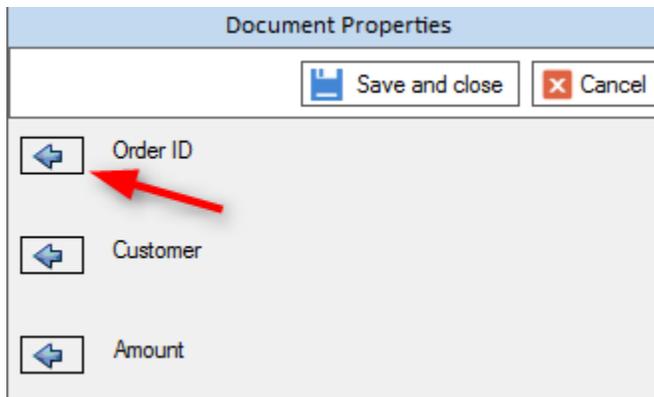
3. Drag and resize property placeholder to fit actual text on the picture

4. Recognized text appears on the Properties Pane under property name. Edit value if necessary.
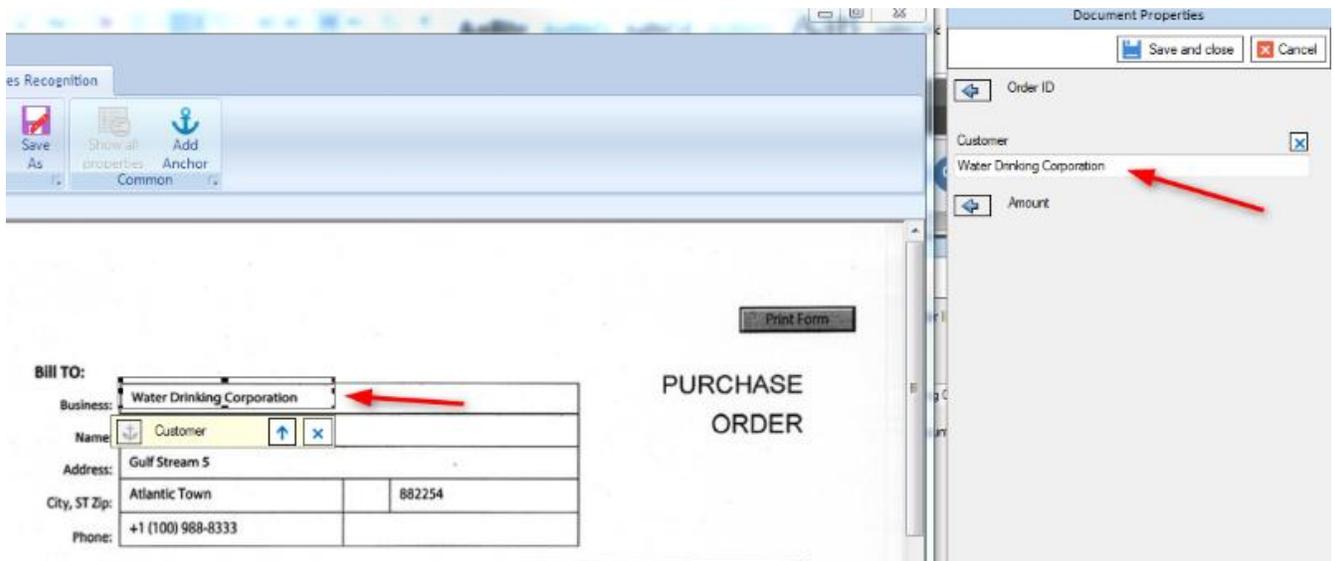
5. Position other properties, click "Save and close"
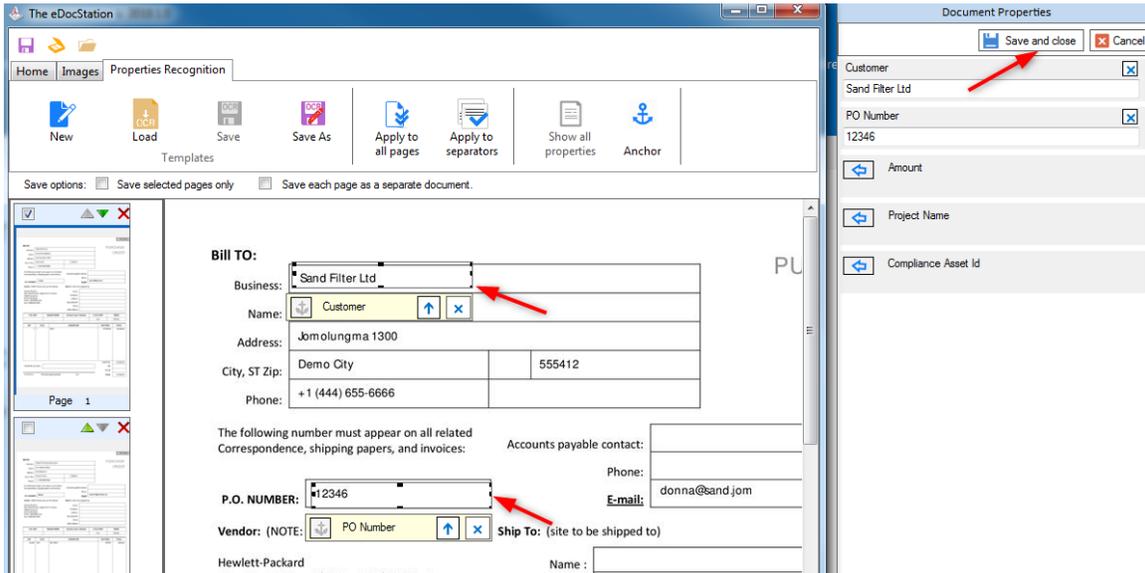
6. You'll see a new button on the page thumbnail:
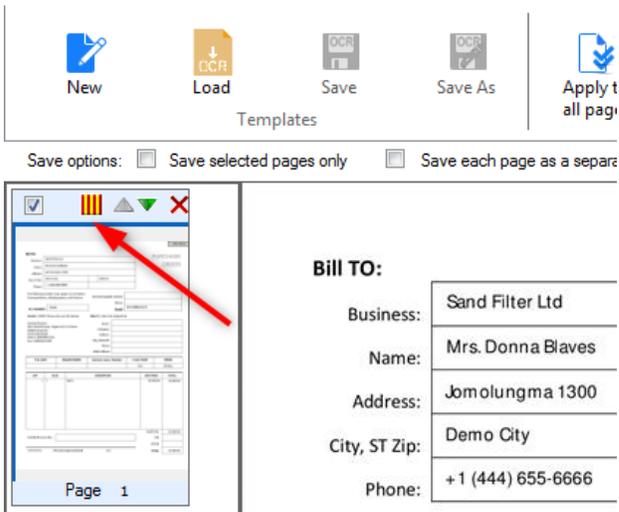
7. Use this button to review and edit page properties values
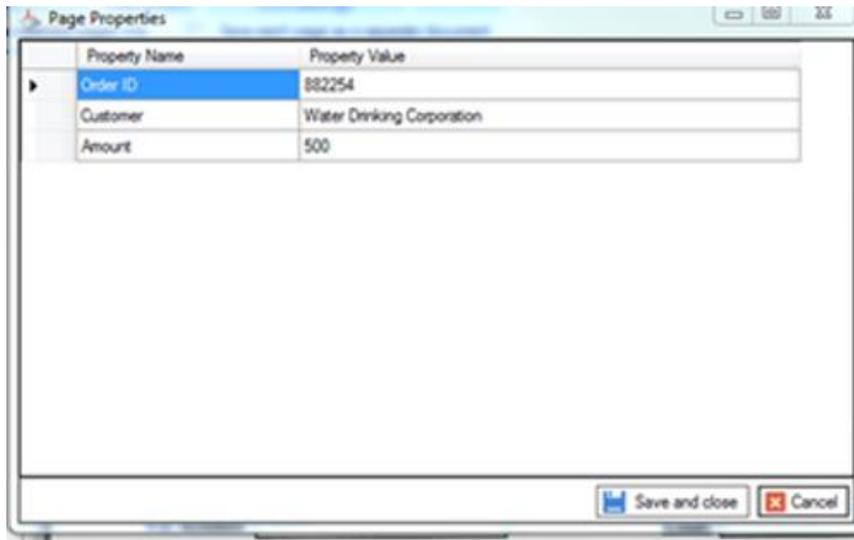
8. Save document. Result:
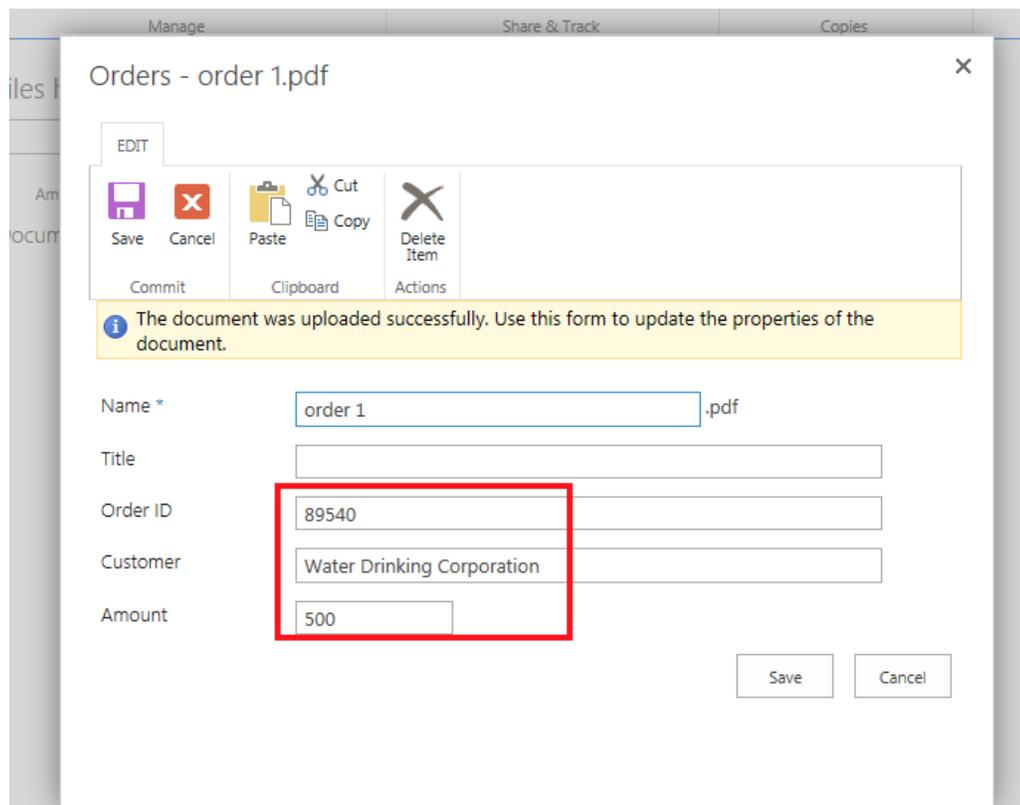
Example 2. Templates

How to create template

1. Scan document and position placeholders as described in the previous example.

2. Add anchor. Anchor stores position of the selected text and text itself. When you apply template for the document, the eDocStation searches for anchor's text and corrects position of the fields. Try to find unique single word or short phrase for anchors. Add two anchors for the best positioning.
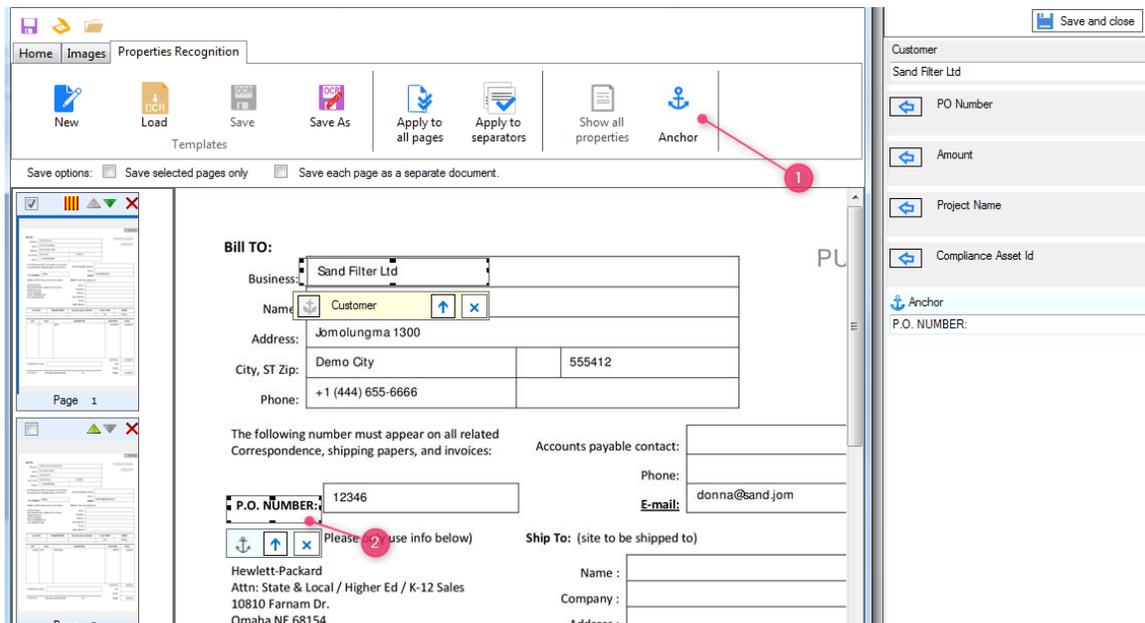
3. Save template.
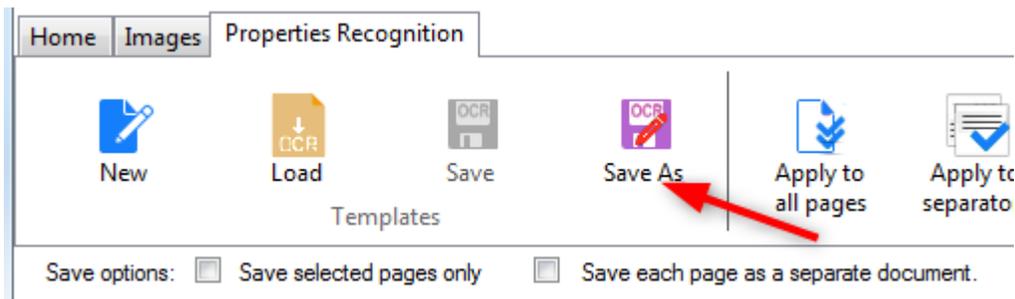
How to use previously saved template

1. Scan or load document
2. Press on the button "Load"
3. Select template from the list
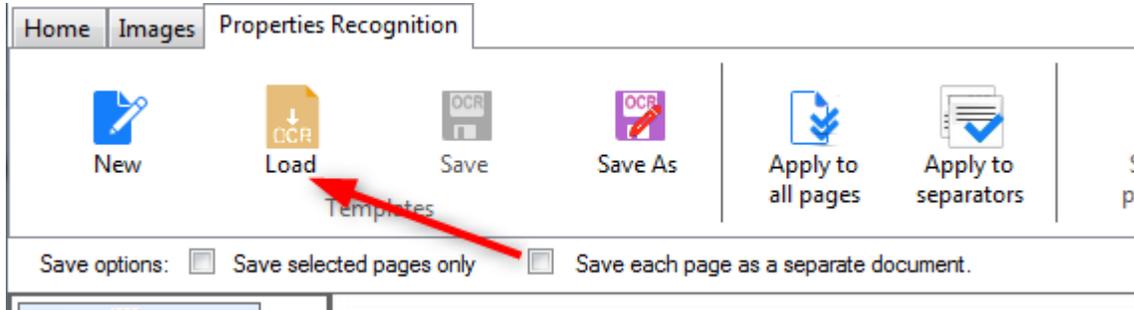
4. The template will be loaded and placeholders will be positioned automatically
   Adjust placeholders positions and/or edit properties values if necessary.

5. Please pay attention, the eDocStation displays only properties saved in the template. If you want to add properties from the current library but not stored in the template, use button "Show all properties"
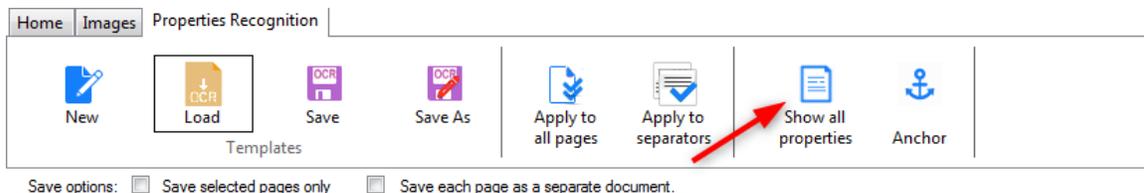
Barcodes ℙ

Use Barcode tab of the Settings window to define barcode recognition settings.
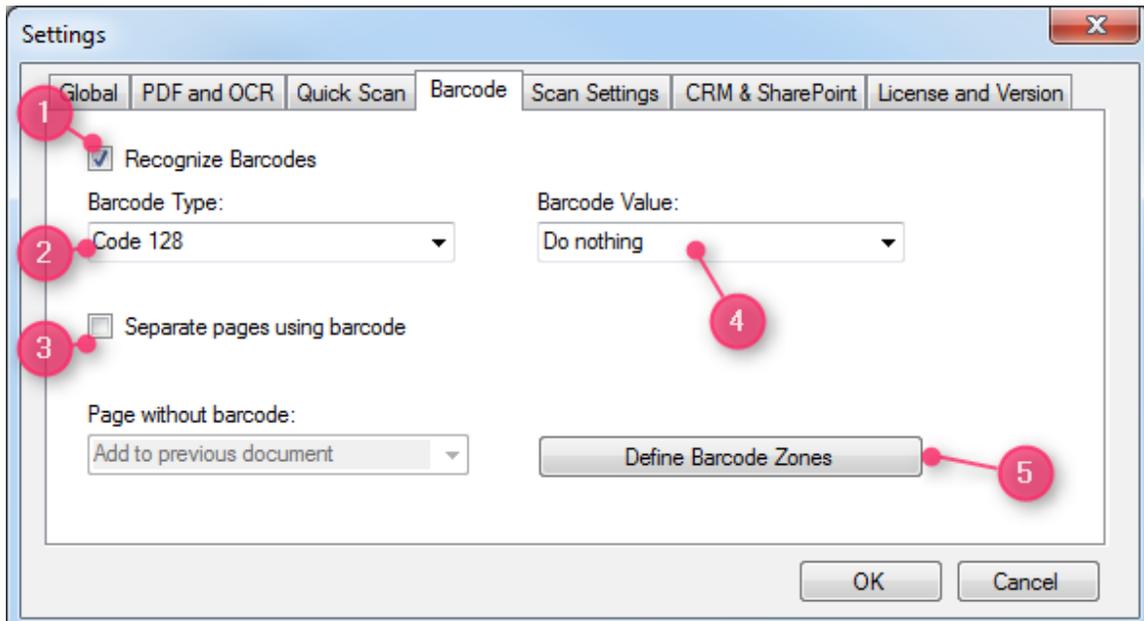
Turn barcode recognition on

Select checkbox "Recognize barcodes" (1) and select barcode type (2)

Separate pages using barcodes

Barcode-based pages separation allows the scan application to detect when a new document should be started. It helps the scan operators to increase productivity, allowing them to organize a stack of paper into logical documents. That eliminates the need to scan each document separately.

Select checkbox (3) to activate this option.

You can select behavior for pages without barcodes using option "Page without barcode" (5):

1. Add to previous document
2. Save as a separate document

Barcode Value:

You can save recognized barcode value to the SharePoint document property.

1. Define property for barcode value in the destination SharePoint library. For example, "Document Barcode":

2. Open the eDocStation settings and set field "Barcode Value"(3) to "Save to SharePoint Property". Field "SharePoint Property" (6) will be displayed.

3. Set "SharePoint Property" (6) value to the name of the previously created SharePoint property.

4. Load page with barcode

5. Save document. Barcode value is saved as a SharePoint property:

Barcode Zones

Barcode recognition on the entire page surface is a time-consuming process. The best way to reduce barcode recognition time is to define barcode zones.
Barcode zone is a rectangle where the eDocStation looks for the barcode. You can define as many barcode zones as you wish. Of course, reducing count of barcode zones will reduce barcode recognition time.

Important rules:
1. When barcode zones are not defined, the eDocStation looks for barcode on the entire page
2. When one or more barcode zones are defined, the eDocStation looks for barcode in the barcode zones only.

How to define barcode zones:

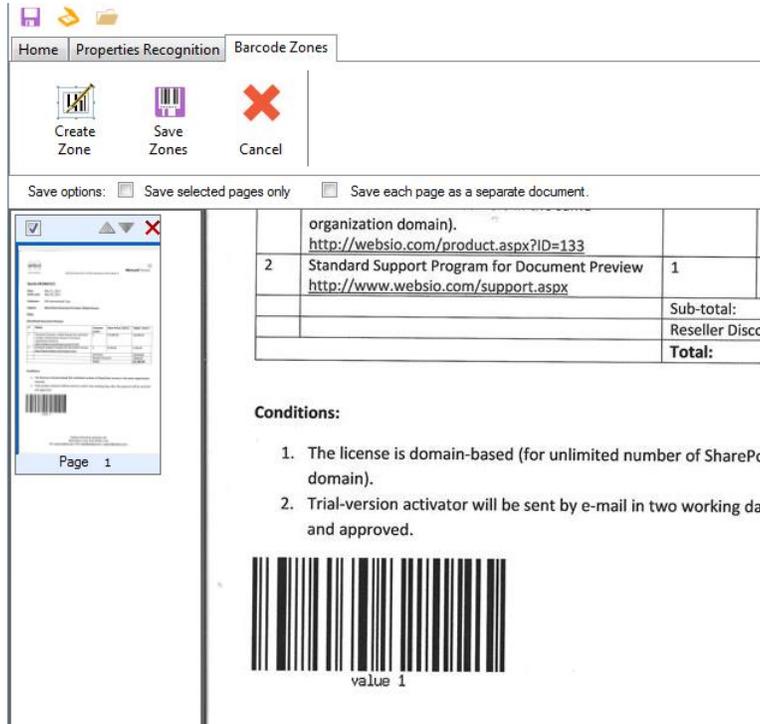1. Load page containing barcode

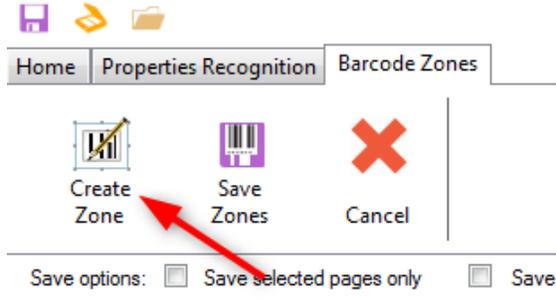2. Go to Settings -> Barcode and click on the "Define barcode zones"
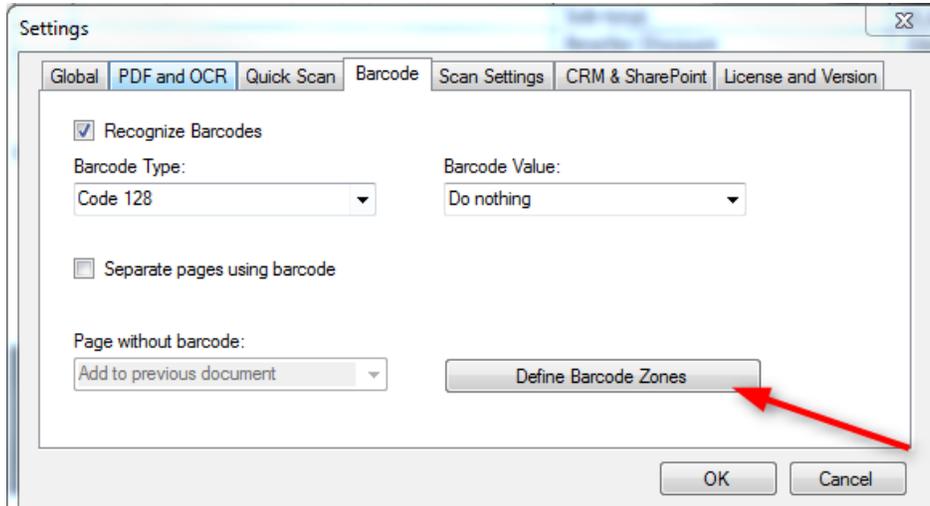
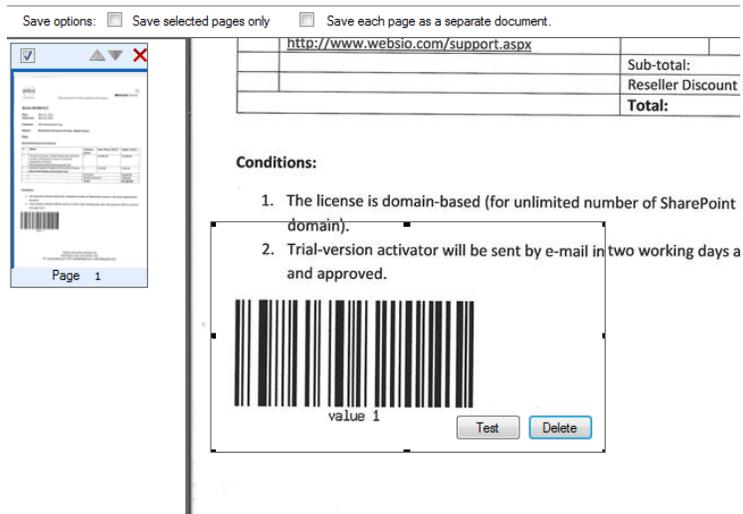3. You'll see tab "Barcode Zones". Previously defined zones will be shown on the image surface.
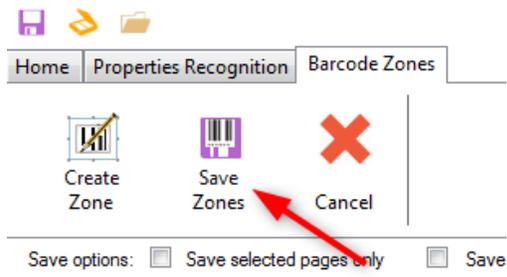
4. Create new zone:
   - Click on the button "Create Zone"

   - Position zone selector

   - Add additional zone selectors (if required) and save zones

Auto-Upload Extension

Auto-upload extension for SharePoint Scan, PDF and OCR Add-in listens to the predefined directories and automatically uploads new or updated files to the SharePoint folders.

Auto-upload processes documents according to file type and content:

| #	Source	Result
1	PDF file containing image only, non-searchable (for example, scanned document).	Searchable PDF containing text after OCR process
2	PDF file already containing text	Uploaded as-is
3	Image files (tif, jpg, png)	Searchable PDF containing text after OCR process
4	Any other file type	Uploaded as-is

Important!

The Listener is not covered by SharePoint Scan, PDF and OCR Add-in license and requires special license. Write to sales@websio.com for price proposal.

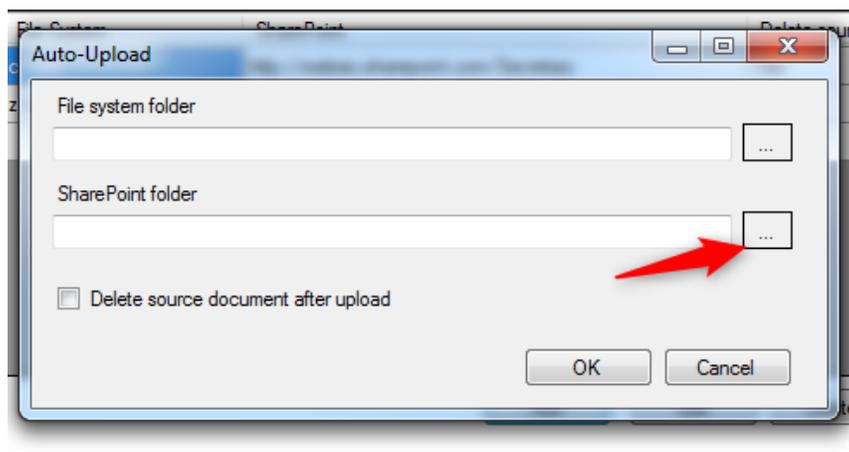1. Open the eDocStation, go to Settings->Auto-Upload
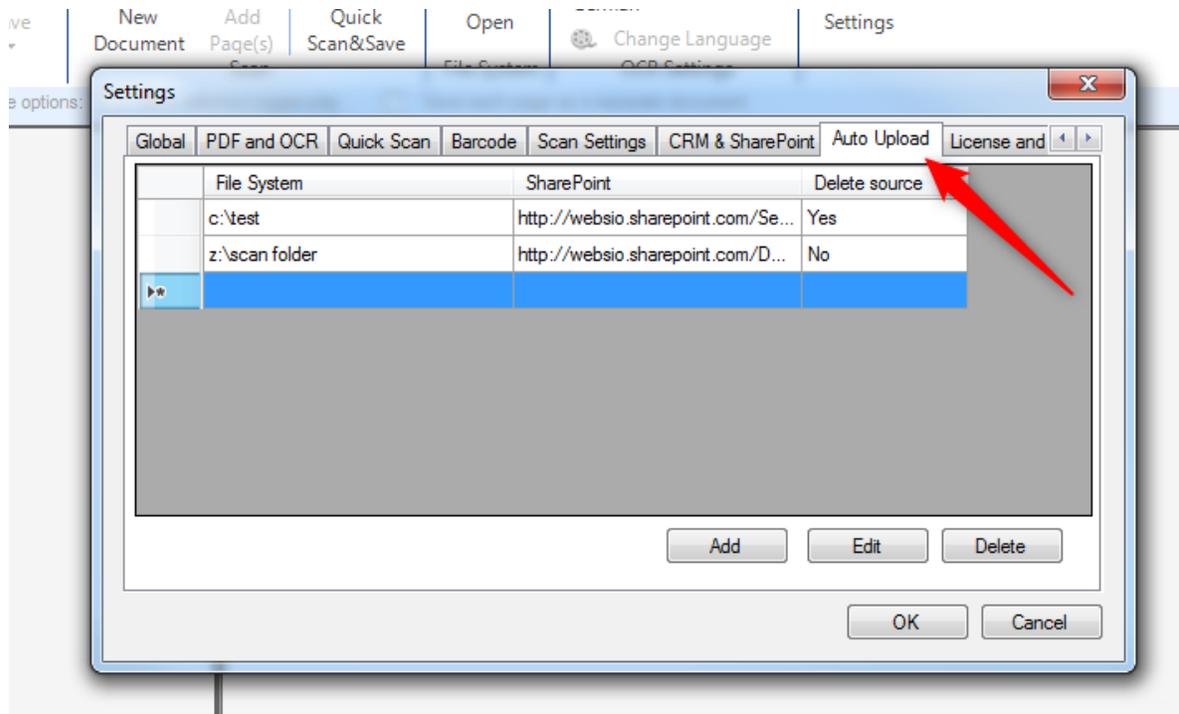
2. Click on the "Add" to define new pair "File System Folder –> SharePoint Folder"
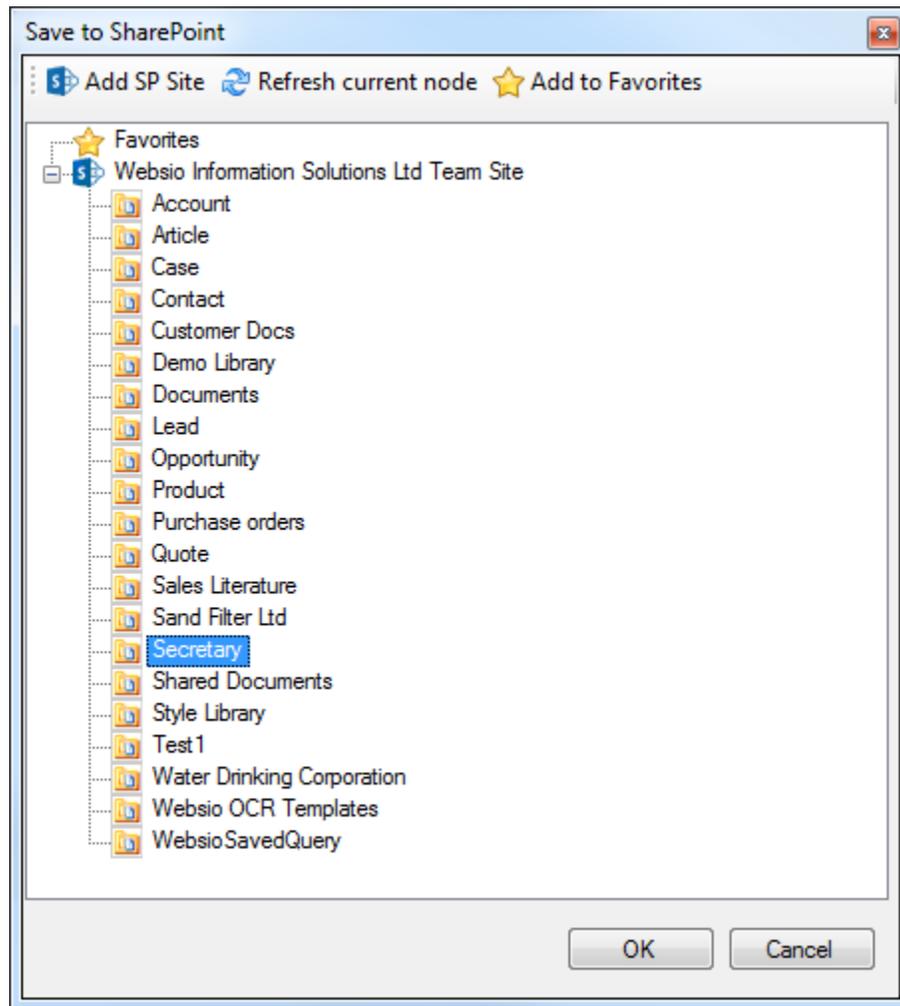
3. Select source directory
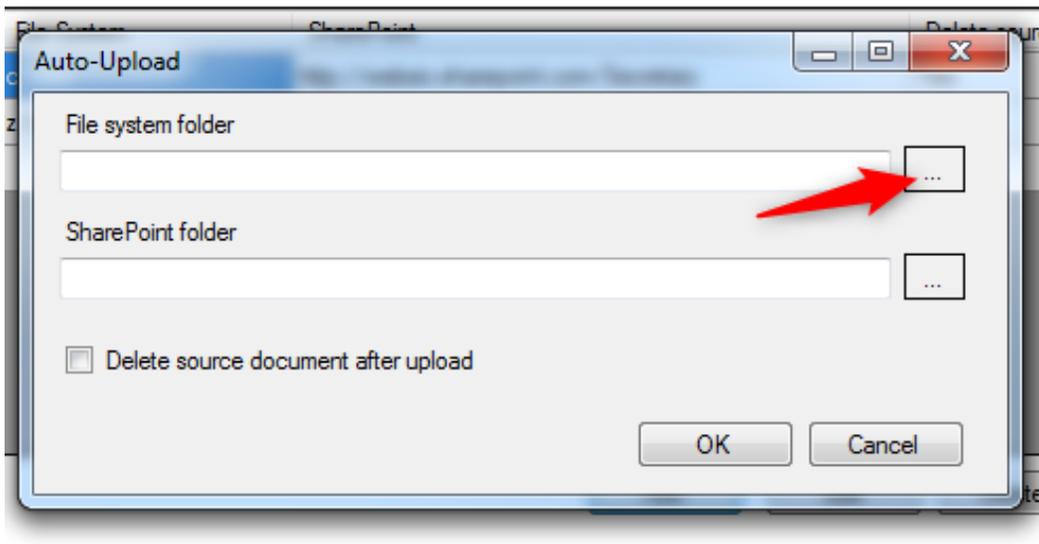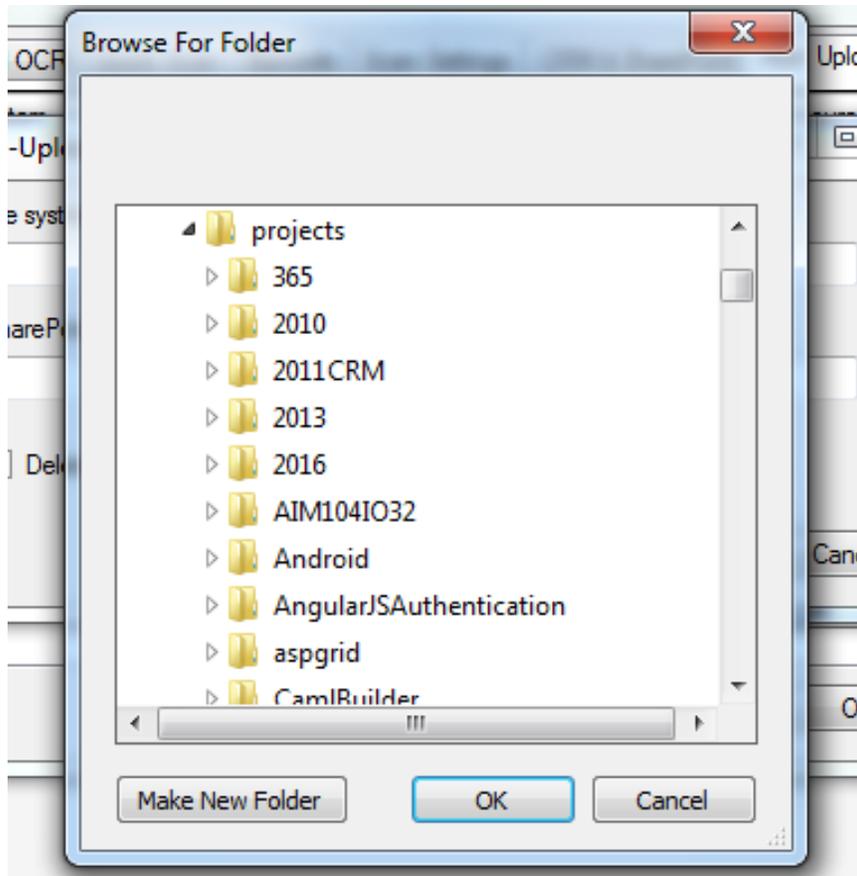
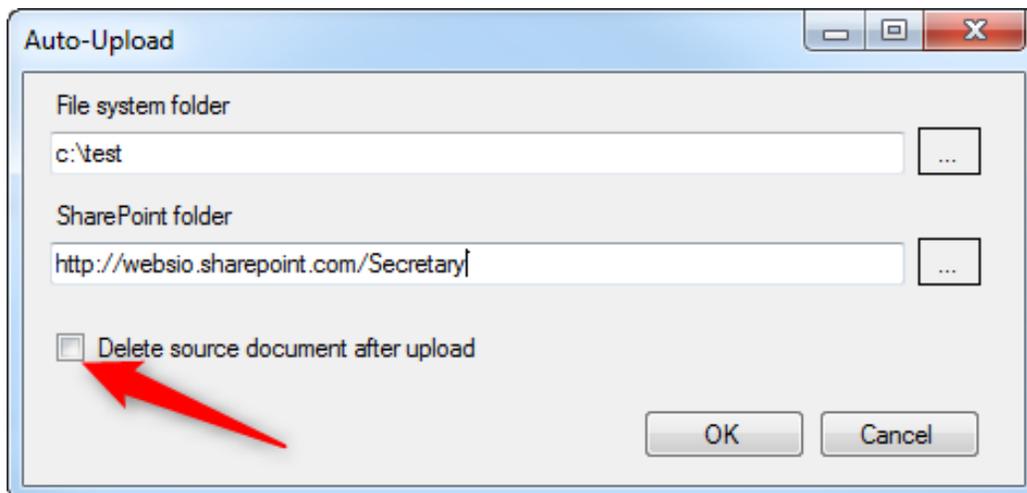4. Select corresponding SharePoint folder
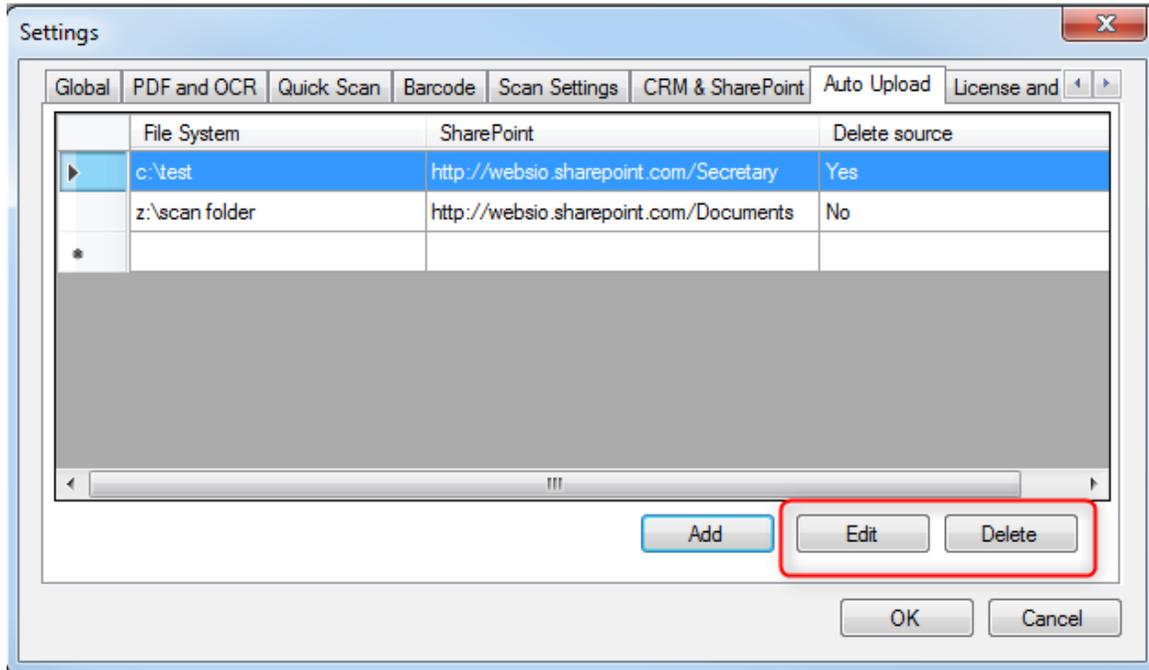
5. Decide if you want to delete source document after upload

6. Press "OK"

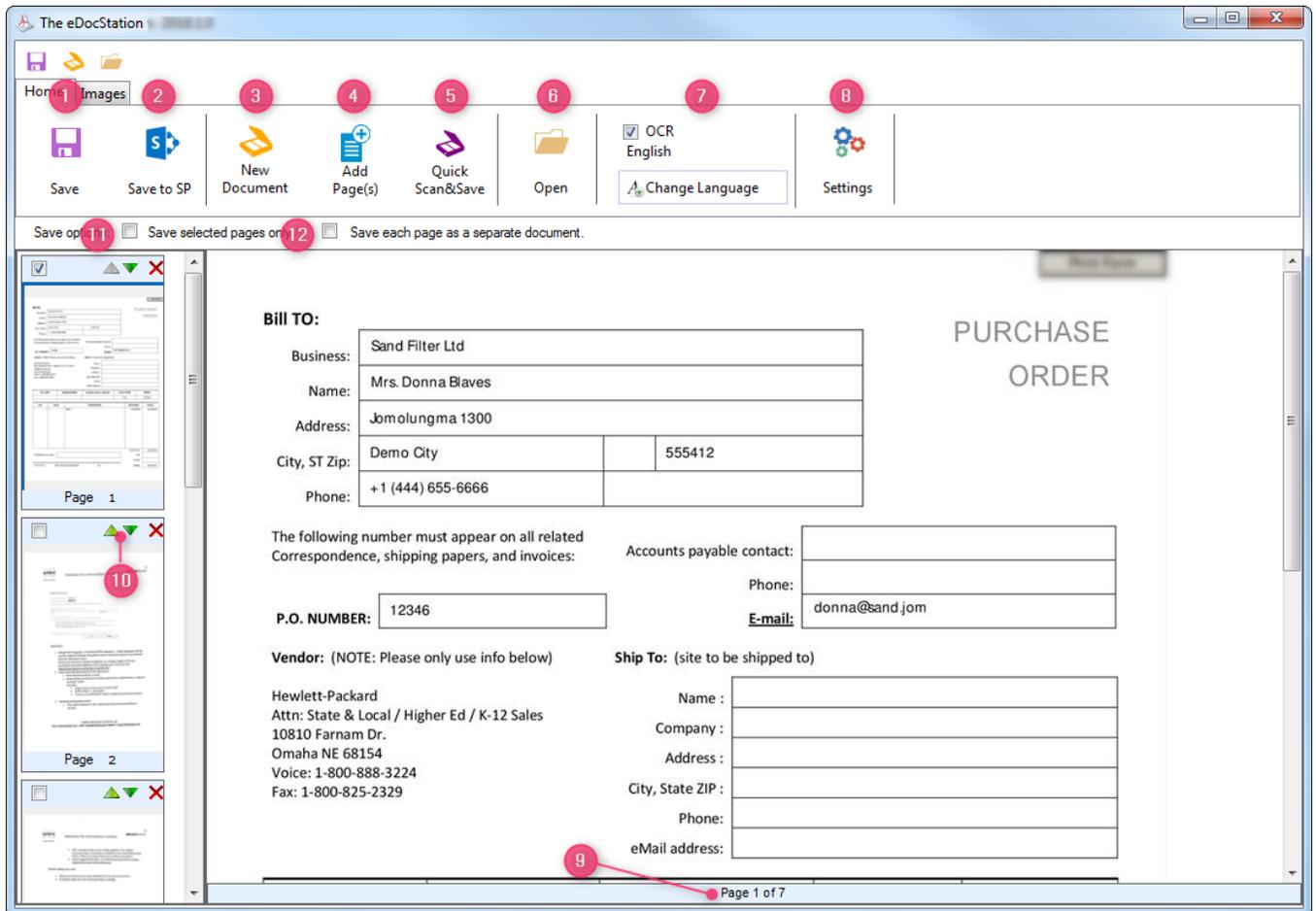7. New pair appears in the grid. Use buttons "Edit" or "Delete" to edit or delete previously defined pairs.

8. Close the eDocStation. Documents will be uploaded by background service "eDocClient".

9. Copy document(s) into the previously selected file system folder, it will be automatically processed and uploaded.

User Interface

Tab "Home". Main view.

1. Save.

2. Save to SharePoint (visible in the standalone mode only)

3. Scan new document

4. Add Page(s). Scan pages and add them to the current document

5. Quick Scan & Save. Read section "Quick Scan" for details

6. Open. Load files from file system.

7. OCR. Select OCR checkbox to use text recognition process. ₽

8. Settings.

9. Current page number and total pages count

10. Reorder pages.

11. Save selected pages only.

12. Save each page as a separate document.

Thumbnails zone displays up 30 thumbnails. When pages count exceeds 30 pages, buttons "Previous pages" and "Next pages" appear.
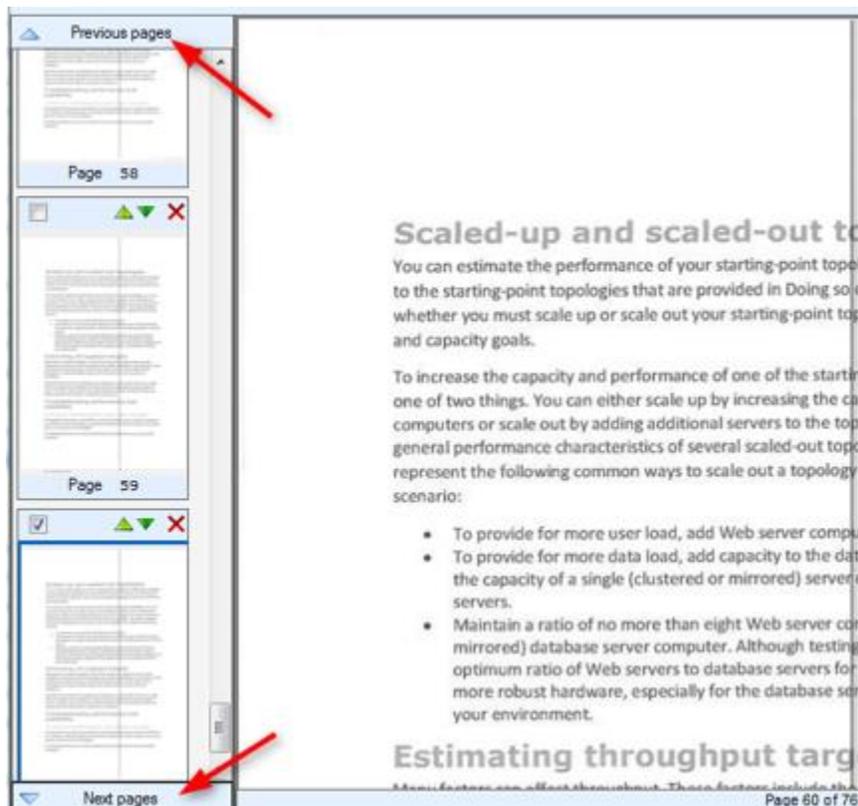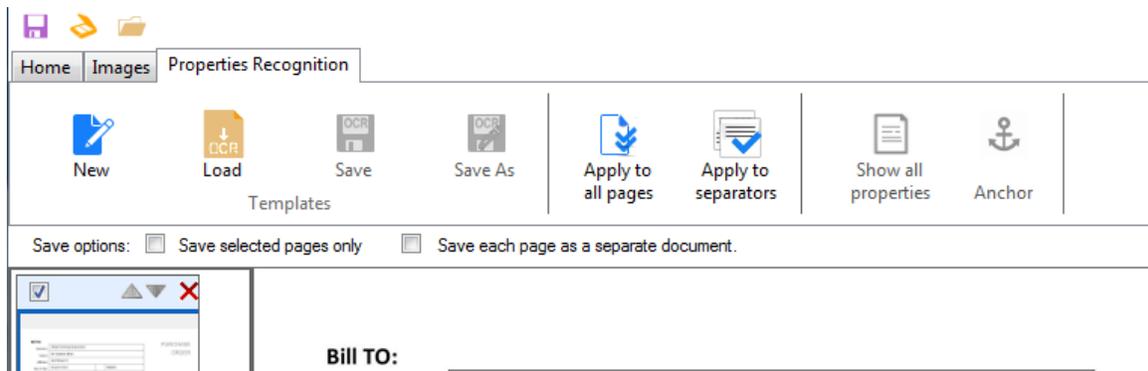
Tab "Images"
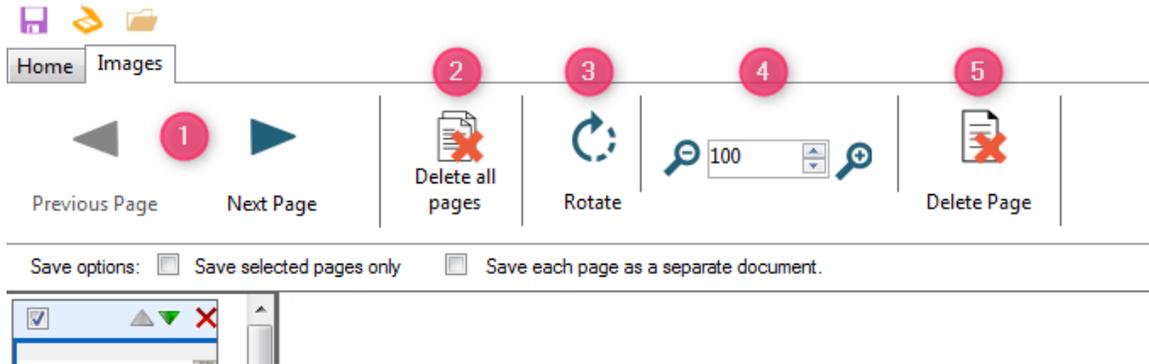
1. Navigation buttons

2. Clear. Delete all pages.
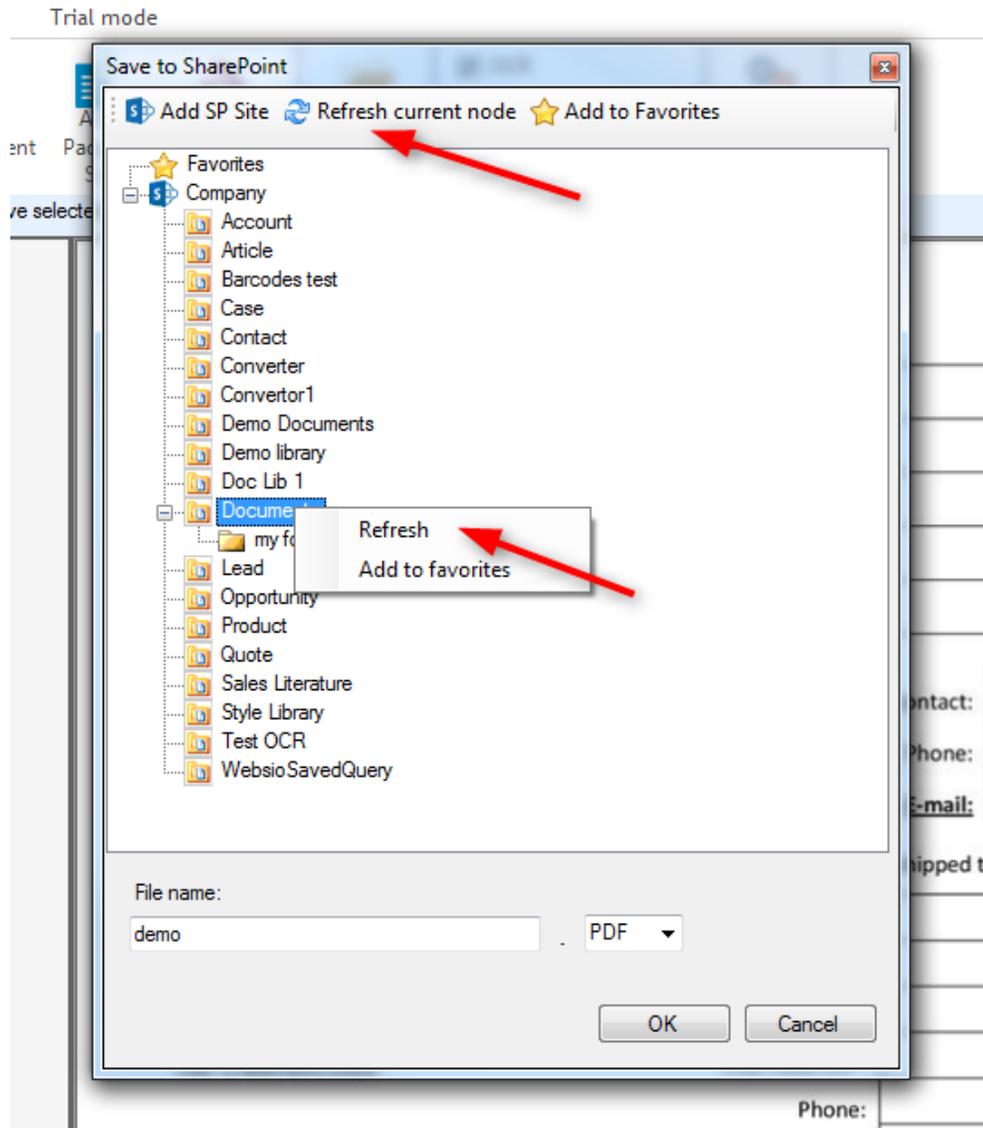
3. Rotate image.

4. Zoom

5. Delete current page

Tab "Properties Recognition"

Read section "Properties Recognition" for details.

"Save to SharePoint" dialog

1. The dialog stores site structure in the cache. If you added/removed SharePoint libraries or folder after they were cached, use option "Refresh" to load them from SharePoint
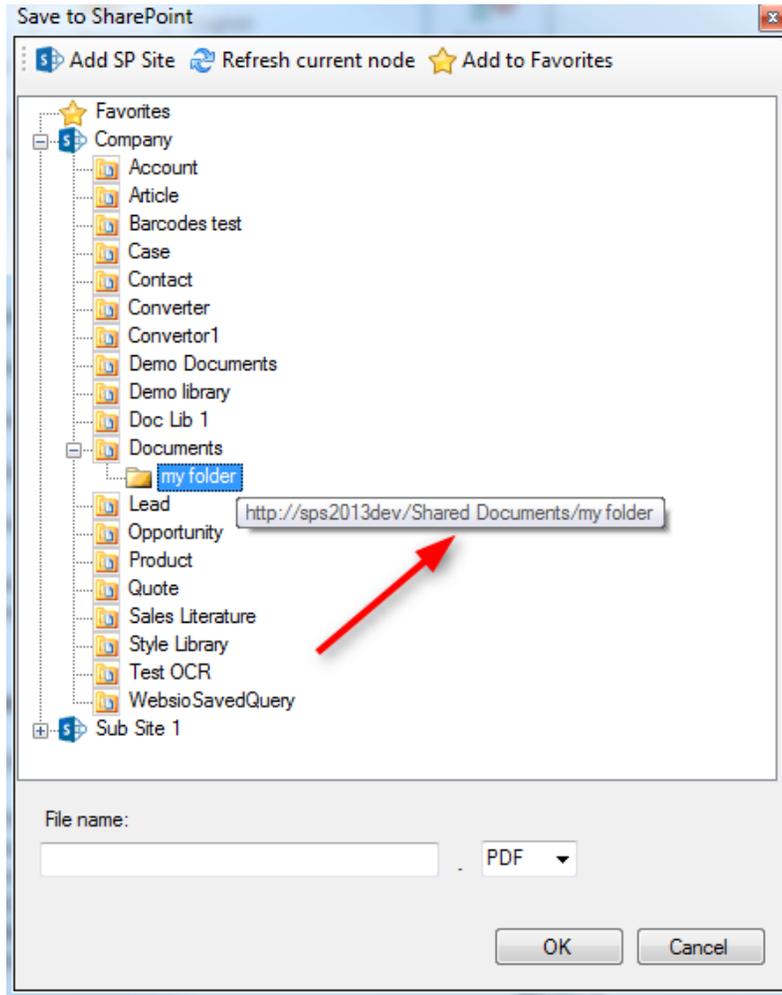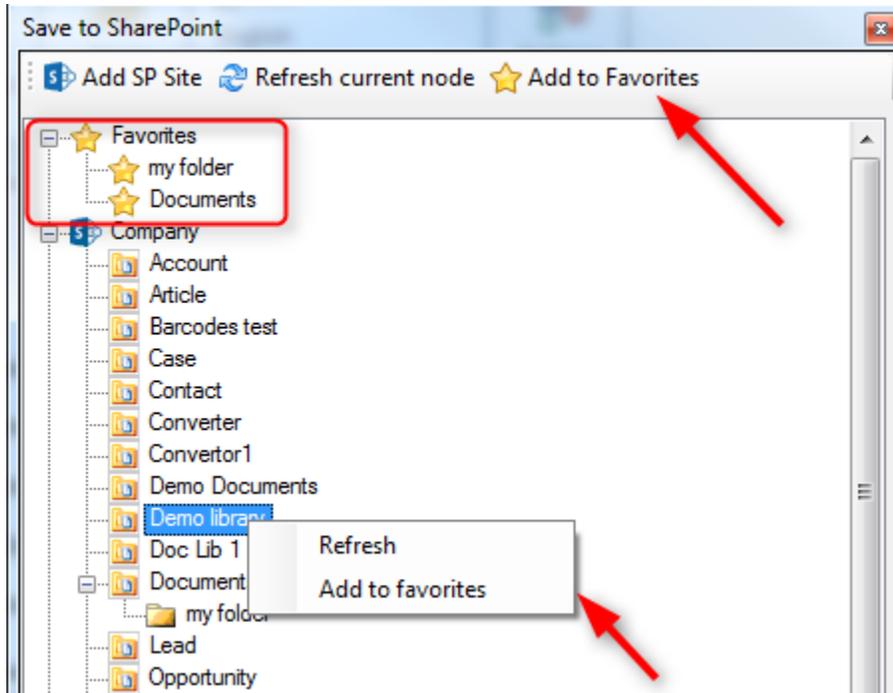
2. Site tree does not contains sub-sites. Use button "Add SP Site" to add URL of each sub-site separately. For example, sites http://spscompany and http://spscompany/subsite1 :
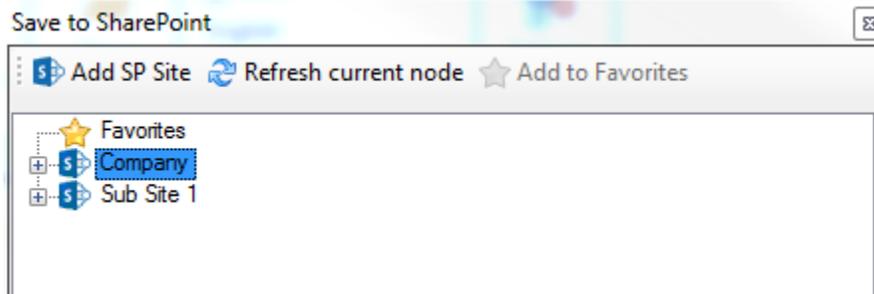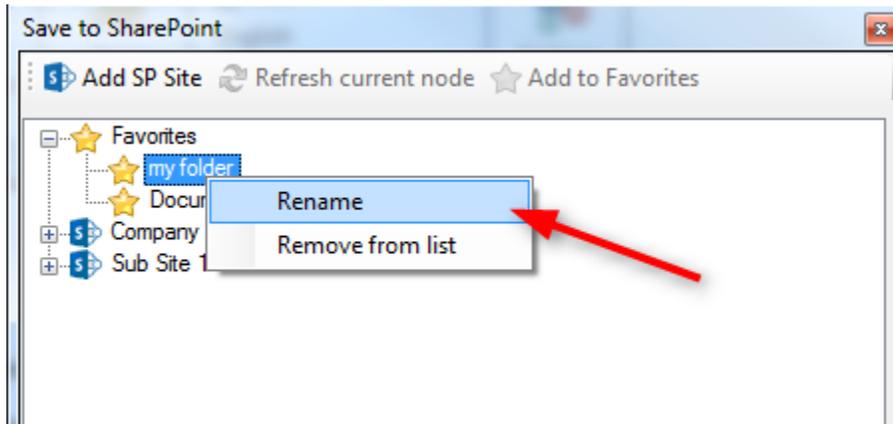

3. Tooltip of each node displays URL of the current site/library/folder.

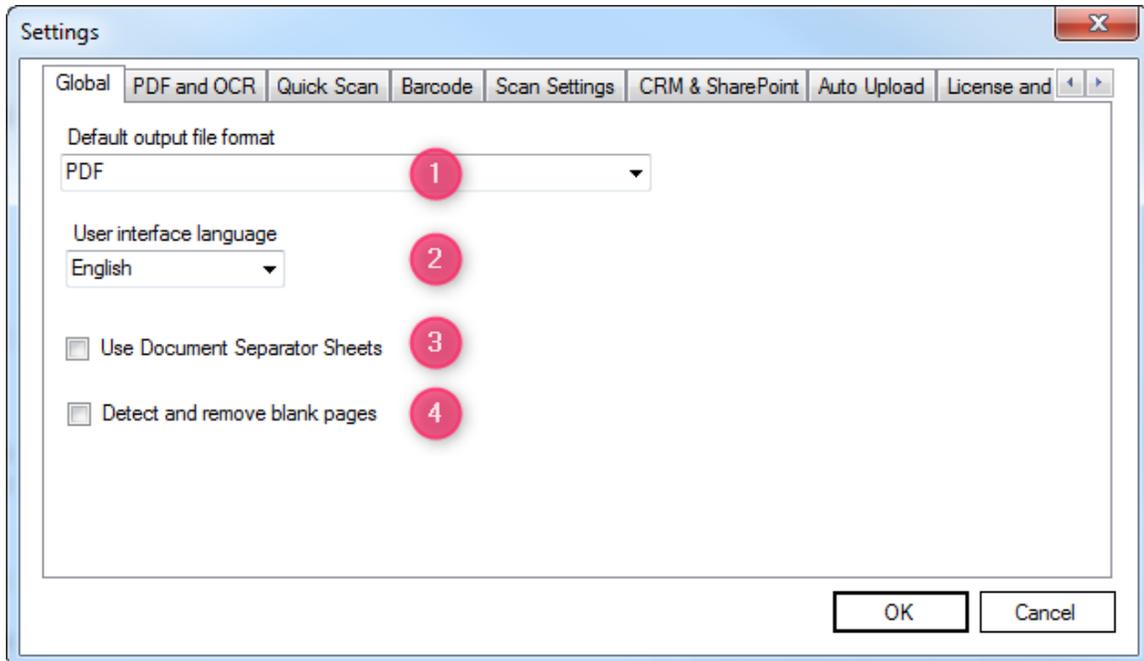4. Add frequently used libraries/folders to Favorites

5. Use menu option "Rename" to provide to Favorite item more meaningful value.
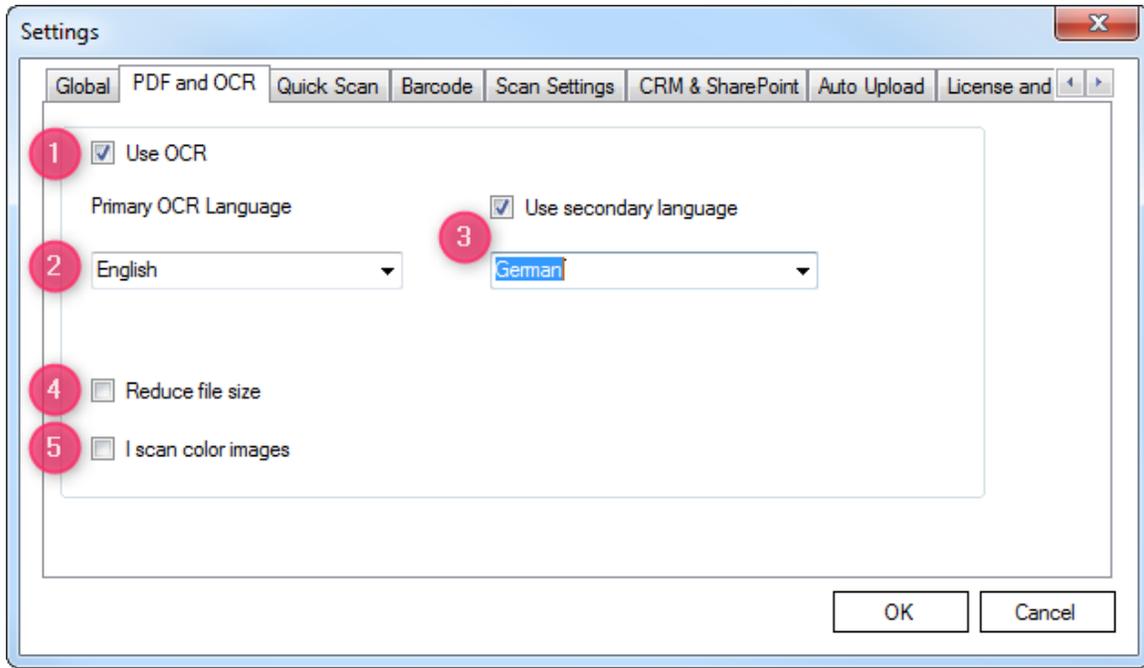
Settings

Global

1. Default output file format. The following formats are supported: PDF, TIFF, JPG, GIF, PNG or BMP

2. User interface language. English, German, Czech, Russian, Arabic and Hebrew languages are available.

3. Use Document Separator Sheets. See section "Document Separator Sheets" above. ℙ
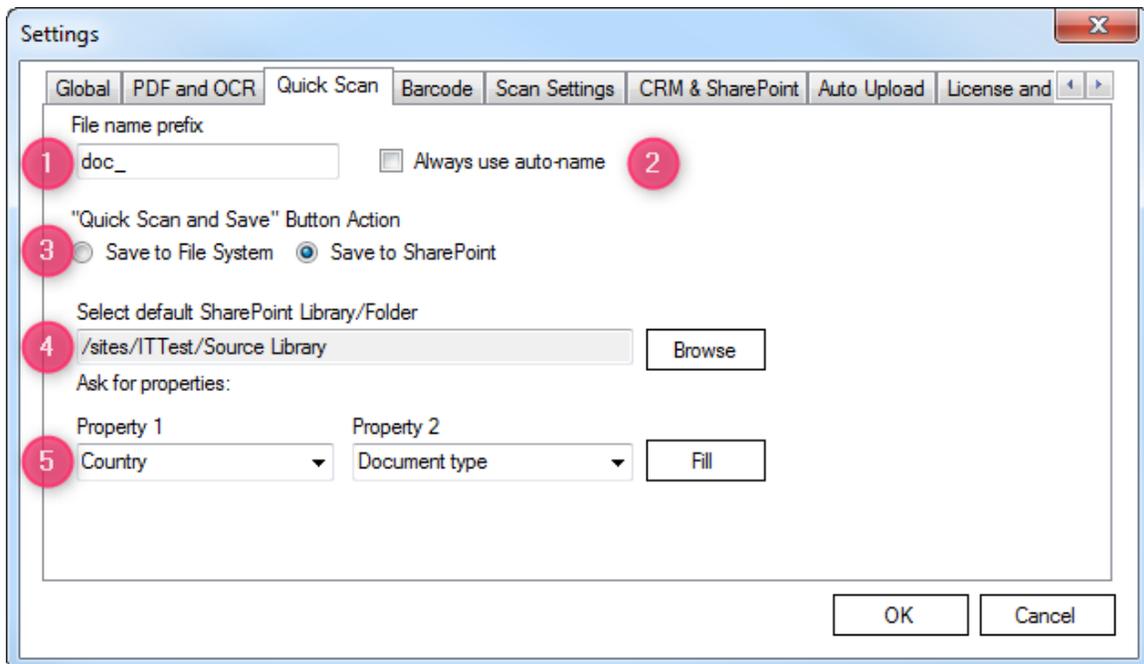
4. Detect and remove blank pages

PDF and OCR ℙ

1. Use OCR – enable/disable OCR.

2-3 OCR languages.

> **Important!** Use secondary language when it is necessary (your documents have bi-lingual content). Best OCR results may be achieved by using single language.

> **Please pay attention:** SharePoint Scan, PDF and OCR Add-in *out of the box* arrives with English OCR language pack only. Download and install additional OCR language packs from our site: http://websio.com/ocrlangpacks.aspx

4. Select this option to reduce size of the created PDF files for 30%-50%

5. Select "I scan color images" to OCR color input files

Quick Scan

1. Prefix for auto-generated file name.

2. Always use auto-name. Select this option to use automatic name in the regular save mode too.

3. Save to file system or to SharePoint by default

4. Destination folder for Quick Scan.

5. Properties to fill during Quick Save to SharePoint. Please pay attention, we recommend to use standard SharePoint interface to fill properties instead of this option.
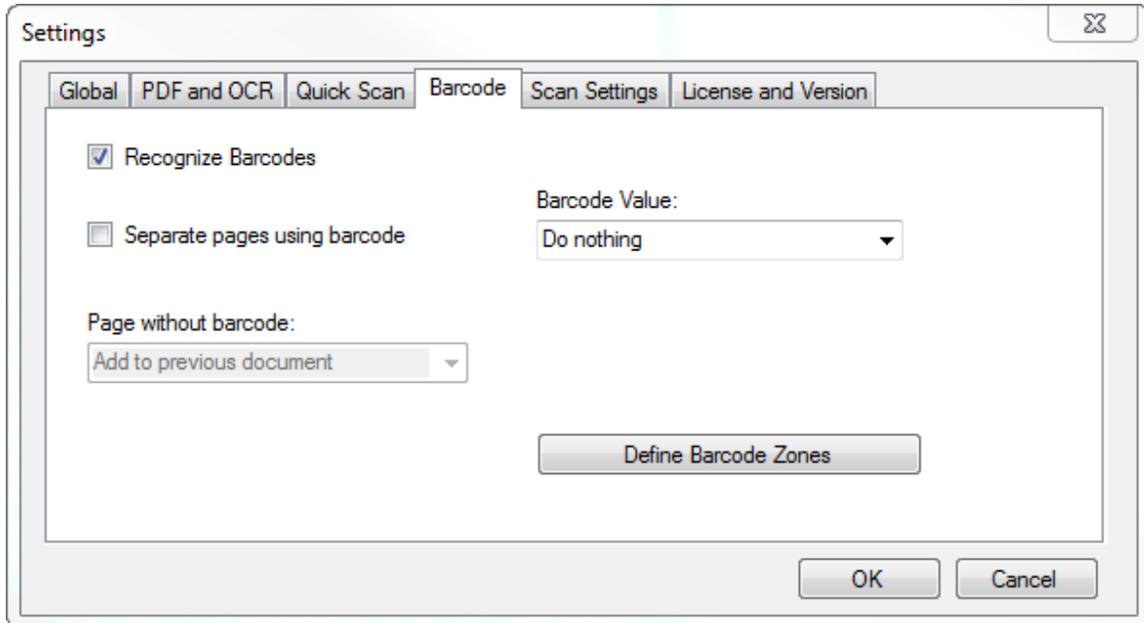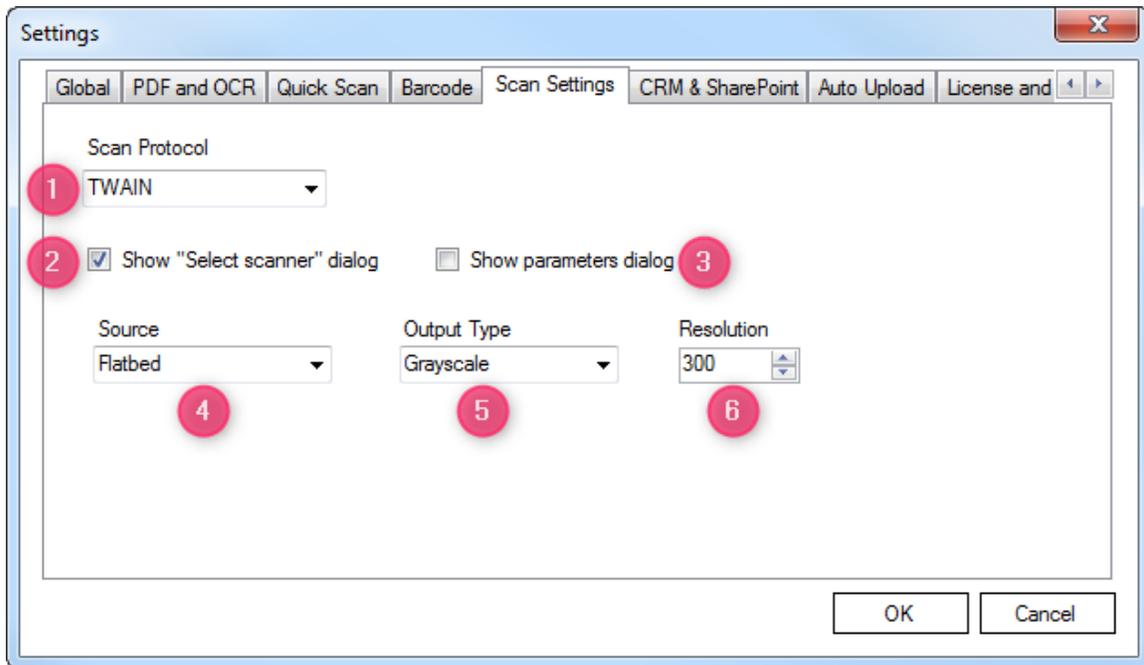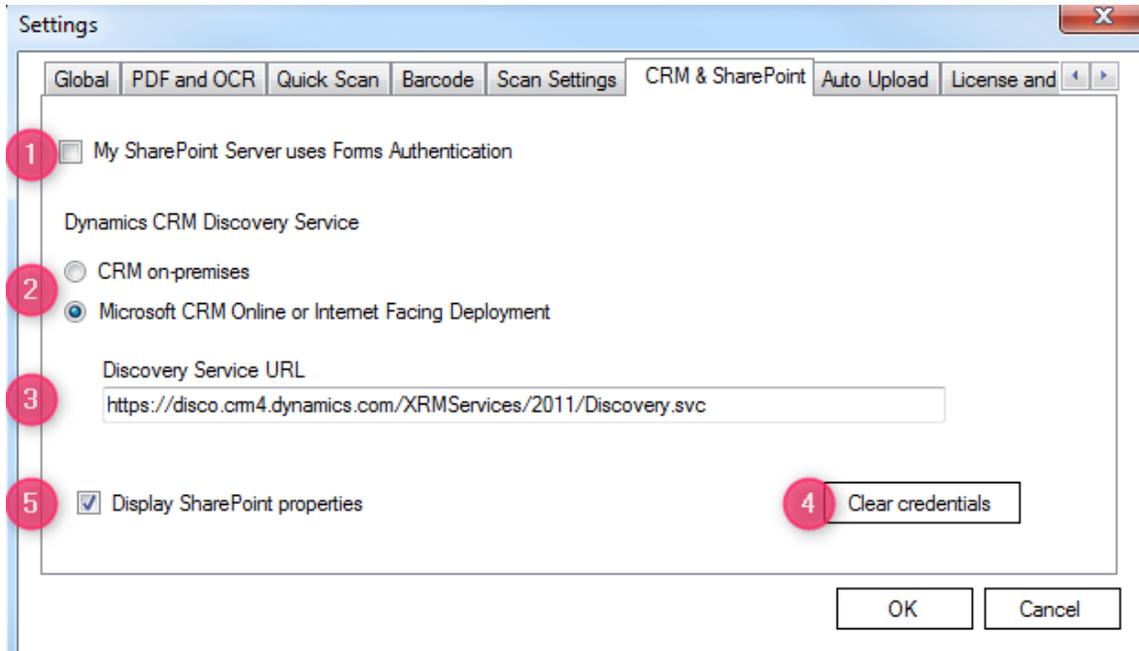
Barcode ℙ

See section "Barcodes" above

Scan Settings

1. Scan protocol. TWAIN or WIA. Default value: WIA.

2. Show "Select Scanner" dialog. Select source scanner before scanning.

3. Show parameters dialog before scan

4. Source. Scan from flatbed glass or feeder.

5. Output type. Color, Grayscale or Black & White

6. Image resolution

CRM & SharePoint

1. If your SharePoint server uses Forms Authentication, select "My SharePoint uses Forms Authentication". If you are not sure which authentication model is in use on your SharePoint, leave this checkbox unchecked.
2. Not relevant for SharePoint
3. Not relevant for SharePoint
4. Clear previously saved user credentials
5. Open SharePoint Document Properties window after upload. This parameter is relevant for the eDocStation launched as a standalone application
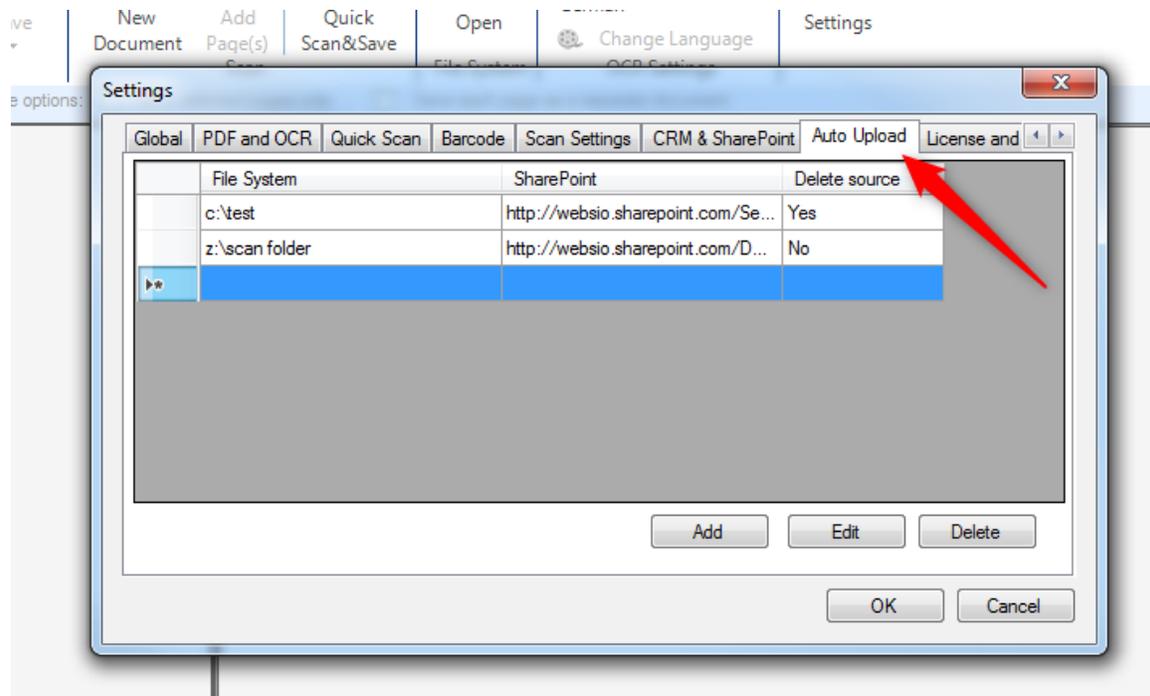
Auto-Upload

See section "Auto-Upload Extension" for details.

Troubleshooting

Error message "Problem with installation package" appears during installation process
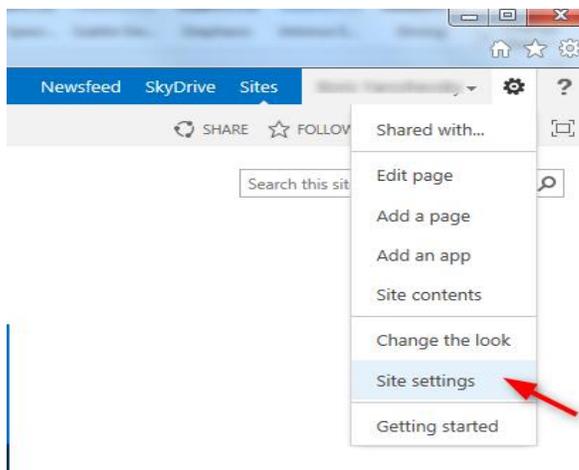

Solution:

Right-click on the "setup.exe" and select "Run as Administrator" in order to begin the installation process.

Button "Scan document" does not appear

1. Open top-level site of the site collection where you are going to use the Add-in.

2. Open "Site actions" menu and select option "Site Settings "

3. In the Site Settings window select option "Site collection features"

Site Collection Administration
Recycle bin
Search Result Sources
Search Result Types
Search Query Rules
Search Schema
Search Settings
Search Configuration Import
Search Configuration Export
Site collection features
Site hierarchy
Site collection audit settings
Audit log reports
Portal site connection

4. In the list of installed features find "Websio Scan, PDF and OCR Add-in" and press "Activate".

If you can't find this option go on to the next step.

	Websio Scan, PDF and OCR Add-in	Deactivate	Active
	Websio Send Mail	Deactivate	Active
	Workflows Aggregated set of out-of-box workflow features provided by SharePoint.	Activate	

5. Open SharePoint Central Administration, select "System Settings" and click on the "Manage farm solutions" link.

6. In the Solution Management window be sure that websioscannerplugin2016.wsp has status "Deployed".

7. In the Solution Properties window press "Deploy solution". Wait until status "Deployed" will appear. Deployment process may take few minutes.

8. Return to step 1.

After click on the "Scan Document" the eDocStation freezes or error is displayed.

1. Probably your scanner does not support selected scan protocol. Open the eDocStation settings and change scan protocol in the "Scan settings".

After scan on the Windows 8 / 10 GDI+ error is displayed.

1. Close all instances of the Internet Explorer. Launch Internet Explorer once as Administrator (right-click on the IE icon and select "Run as Administrator").
2. Open SharePoint portal, scan document and save it.
3. Next time you can launch IE in the common mode.

Another problem

Ask for our support! Write to: support@websio.com

This Software uses:
1. *iTextSharp library v.4.1.6 (http://sourceforge.net/projects/itextsharp/) distributed under Lesser General Public License (LGPL) to convert images to PDF*
2. *Tesseract OCR engine library (http://code.google.com/p/tesseract-ocr/) distributed under Apache License 2.0 for Optical Character Recognition (OCR)*
3. *AForge.NET Framework published under LGPL v3 license.*