



32158 Camino Capistrano  
Suite A PMB 373  
San Juan Capistrano, CA 92675  
Sales@CaptureComponents.com  
[www.CaptureComponents.com](http://www.CaptureComponents.com)

# Scanning to SharePoint® and Office 365 with ccScan®

## *Document Capture to the Cloud*

### Overview

*This white paper shows how ccScan®, a production-quality document capture application, can be used to capture documents to Microsoft SharePoint® in an automated, time-saving manner.*

*Microsoft SharePoint is a highly popular business collaboration platform for the enterprise. Typically it is the single point for all information aggregation, search and collaboration in an organization. Since SharePoint is multi-functional it can be used to build and manage internal and external websites, to build communities for internal networking and social collaboration, to search for organization-wide information, to explore the organization's business intelligence, to pull and push data from other systems used inside an organization, and last but not least to manage content such as documents, information and records.*

*SharePoint provides the tools to manage all of an organization's documents including controlling who can read and update them, categorize them for easier search and retrieval, and mark them as official records and lock them down from further change. Since there is more to SharePoint than this short introduction can cover please visit the SharePoint website at <http://sharepoint.microsoft.com>.*

*ccScan is designed to complement the comprehensive content management capabilities of SharePoint by capturing documents directly to SharePoint document libraries and lists. Although SharePoint customers can easily upload their documents if they are already in digital format, paper documents must be scanned first and then manually uploaded. ccScan eliminates the time typically wasted with such a multi-step, manual operation that consists of scanning a document locally to the PC, naming this document, creating a folder in SharePoint, uploading the document to SharePoint, and finally entering values in the input fields or "columns" attached to the SharePoint Document Library or List. Instead ccScan does all of the above in a single-step operation that is fully automated. Naming of documents and folders can be manual or automated by using barcodes, OCR with Text Pattern Search, and running counters. All input fields defined for a Document Library or List can be entered to index the captured documents. In addition to scanning paper documents, ccScan has bulk import capabilities for already digitized documents that allow entire folders to be uploaded to SharePoint in a single step using wild card file specifications and recursion.*

## ccScan for SharePoint®

Capture Components, LLC markets different versions of ccScan targeted for different document repositories. The ccScan versions for “Internet Cloud” based storage support Google Drive and Docs, Box, Salesforce, Dropbox, Microsoft SkyDrive, Amazon S3, Windows Azure, etc, but other types of document repositories are supported as well. For example ccScan works with the Microsoft Windows file system utilizing its local hard drive or networked folders for storage. And, as discussed in this white paper, documents can be captured remotely and stored in Microsoft SharePoint and SharePoint Online using SharePoint’s Internet-based web services interface.

ccScan supports SharePoint Online Office 365, SharePoint 2010, and SharePoint 2007. For online access a username/password login is supported (both Browser-based and Windows Forms-based) and in traditional Intranet environments with a domain server seamless Single Sign-On is supported through Windows login.

## Two Editions: ccScan Advanced and ccScan Standard

**ccScan Advanced for SharePoint** is best explained with some examples: run unattended jobs to import electronic faxes, upload them to automatically selected folders of a SharePoint Document Library or List, and automatically name the document, all based on information extracted from barcodes and text pattern searches in text obtained through OCR. Or scan documents and automatically populate custom Columns in a SharePoint Document Library or List. Since ccScan is highly configurable it can be applied to many scenarios where opportunities for automation and large time savings are present.

ccScan Advanced features are based upon sophisticated technologies such as barcode detection, OCR, and Text Pattern search with Regular Expressions. These capabilities are applied to both the scanning of paper documents in the paper-based office and the processing of electronic documents in the paperless office.

**ccScan Standard for SharePoint** is similar to ccScan Advanced without the automation capabilities. ccScan Standard is typically used in a paper-based office to efficiently scan paper documents to SharePoint. Because ccScan Standard scans directly to SharePoint you can eliminate manual steps like scanning locally, naming the document, creating a folder in a SharePoint Document Library or List, uploading the document to SharePoint, and finally entering values in the input fields (“Columns”) defined for the Document Library or List. Instead, ccScan Standard does all of this in a pre-configured and efficient single-step operation.

## ccScan Overview

Although ccScan costs significantly less than competing products, it is a full-featured, production-level, document capture system. The software takes full advantage of today's multi-core processors. On a quad-core computer the scanning, barcode detection, text recognition, and export to SharePoint each run on a separate core to maximize throughput. Consequently most TWAIN-compatible production scanners are operated at rated speed. Since ccScan has been designed to be enterprise-friendly, it contains a Setup mode and a User mode. The Setup mode is password protected. Jobs are created in Setup mode and can be used over and over again in User mode. To make sure ccScan can be used immediately after installation, without having to create a job first, a set of default jobs is provided to cover common document capture scenarios.

Jobs can be associated with a scanner button, which can make running a job in User mode as simple as pressing a button on the scanner. ccScan also contains an Index/QA mode that allows the operator to check the document or enter indexing data before the document is sent to SharePoint. Documents are automatically separated if barcode sheets are used.

On platforms that contain no native document viewing options such as Amazon S3 and Windows Azure, ccScan provides the capability to create View jobs that will list all documents entries or a subset in either a sequential list or hierarchical format. View jobs, however, are not enabled for SharePoint since SharePoint natively provides document viewing capabilities.

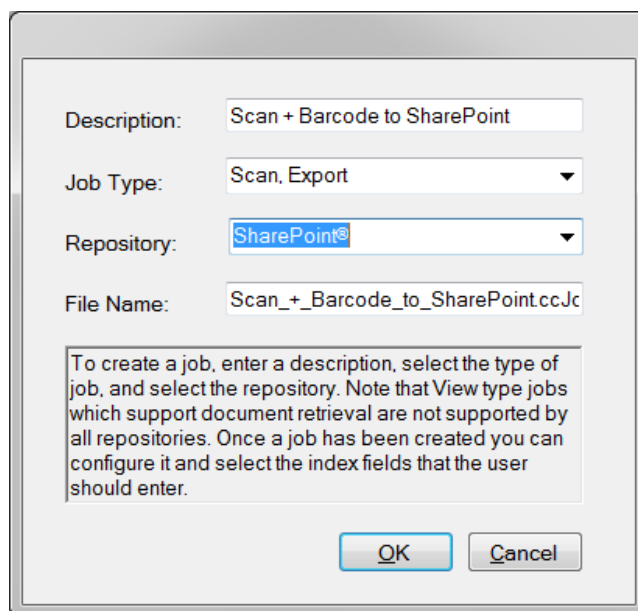
Despite its low price, ccScan has all the high-end features and no restrictions such as page click charges. For detailed information about ccScan or to download a free trial version, please visit Capture Component's website at [www.capturecomponents.com](http://www.capturecomponents.com).

## Typical Scenario

To give the reader a flavor of capturing and storing documents in SharePoint with ccScan the following sections contain walkthroughs annotated with screenshots of a typical scanning and bulk import scenario. Although the default set of jobs installed with ccScan cover capturing and storing documents in SharePoint, the example scenarios assume that no ccScan Jobs have been created previously so both Setup and User modes can be shown. For ease of illustration certain details such as the creation of an account with SharePoint or how to enter Setup mode have been left out of these descriptive scenarios.

### Create Scanning Job

To start a new scanning job must be created. In ccScan jobs are created in Setup mode. Once a job has been created, it can be used over and over again in regular User mode. The Administrator selects default values for the various job settings and decides which job settings can be entered or modified by the operator. The job settings are organized by category such as Scanner settings or Export settings.



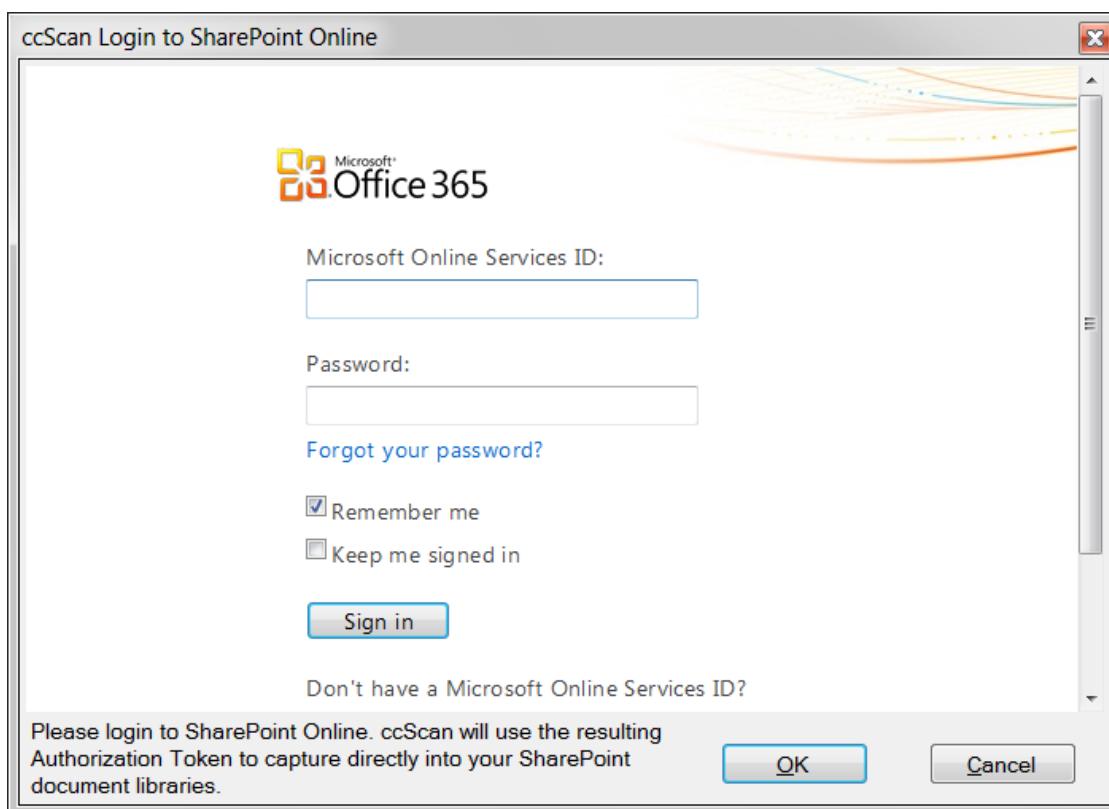
**Figure 1 - Create New Scanning Job**

In the dialog box above SharePoint has been selected as the repository for export of the captured documents. If ccScan were licensed for multiple repositories, the target repository for the job would be selected here from the list of available repositories. Note that the Job Type is Scan, Export.

## Enter SharePoint Credentials

Immediately after the new job has been created, depending on the selected login method, the operator may need to enter credentials. In the case of SharePoint Online Office 365 the dialog below comes up to enter the user name and password for a SharePoint account.

Both in User mode and Setup mode the operator must enter an email/username and password of an active SharePoint account. When the Sign In button is pressed ccScan obtains an authorization token from SharePoint that is subsequently used to request information from SharePoint such as a collection of existing document libraries and lists.

The image shows a Windows-style dialog box titled "ccScan Login to SharePoint Online". Inside the dialog, the Microsoft Office 365 logo is displayed at the top. Below the logo, there are two text input fields: "Microsoft Online Services ID:" and "Password:". A blue link "Forgot your password?" is positioned below the password field. There are two checkboxes: "Remember me" (which is checked) and "Keep me signed in" (which is unchecked). A "Sign in" button is located below the checkboxes. At the bottom of the dialog, there is a line of text: "Don't have a Microsoft Online Services ID?". Below this, a larger text block states: "Please login to SharePoint Online. ccScan will use the resulting Authorization Token to capture directly into your SharePoint document libraries." At the very bottom right, there are "OK" and "Cancel" buttons.

**Figure 2 - Enter SharePoint Online Office 365 Credentials**

To prevent users from having to login each time ccScan is started, the obtained authorization token is stored in between ccScan sessions in an encrypted file that can only be read by ccScan and only on the PC where it was created. Note, however, that the SharePoint Online Office 365 authorization tokens are valid for less than 24 hours. ccScan can be configured to disable the default auto-logon which requires that users re-enter their SharePoint credentials each time they use ccScan.

## Configure New Scanning Job

In the form below the Administrator can configure the new Scanning job. The same form is used to make modifications to the job later. The type of Login Method to select depends on the SharePoint environment. For SharePoint Online Office 365, select Browser. Browser login will request a user name and password using a browser page (see the screenshot in Figure 2).

The screenshot shows the 'ccScan' application window with the 'Setup Job: Scan + Barcode to SharePoint' configuration form. The form is organized into a table with columns: Field Name, Default Value, Show, Order, Input Control, and Edit Field Name. The fields are as follows:

Field Name	Default Value	Show	Order	Input Control	Edit Field Name
<b>Job</b>					
Description:	Scan + Barcode to Share	<input type="checkbox"/>	0	TextBox	Description
<b>Export</b>					
Login Method:	Browser (Office 365)	<input type="checkbox"/>	0	ComboList	Login Method
Host URL:	http://capturecomponent:	<input type="checkbox"/>	0	TextBox	Host URL
Logged-In User:		<input type="checkbox"/>	0	TextBox	Logged-In User
Named User:		<input type="checkbox"/>	0	TextBox	Named User
Document Library:	Publications	<input checked="" type="checkbox"/>	1	ComboList	Document Library
View:	All Documents	<input type="checkbox"/>	0	ComboList	View
SharePoint Folder:		<input checked="" type="checkbox"/>	2	FolderBox	SharePoint Folder
Document Name:	Barcode_%c4:00%_%b1	<input checked="" type="checkbox"/>	3	TextBox	Document Name
Author:		<input checked="" type="checkbox"/>	4	TextBox	Author
Publisher:		<input checked="" type="checkbox"/>	5	TextBox	Publisher
When Published:		<input type="checkbox"/>	0	TextBox	When Published

At the bottom of the form are buttons for 'Validate', 'Save Job', and 'Close'. The status bar at the bottom shows 'OK' and a set of icons for 'Scn', 'Imp', 'Ocr', and 'Exp'.

**Figure 3 - Configure New Scanning Job**

For Internet hosted SharePoint environments other than Microsoft Office 365, select Forms. Forms login will request a username and password using a Windows dialog instead of a browser page. For Intranet environments where *single sign-on* is desired, select Windows login. The obtained user credentials are cached in memory and re-used over and over again (for each job) while ccScan is running. The dialog comes up automatically when ccScan does not have valid credentials in its memory. After selecting the Login

Method, enter the complete URL to the SharePoint site on the SharePoint server in Host URL.

To allow the administrator flexibility jobs can be configured with Named User such that only a specific SharePoint user can run a particular job. Normally jobs can be run by any logged-in SharePoint user. Note that Named User and Logged-In User requires Forms login. In addition to this *named user* feature, ccScan allows the administrator to set a time-out for a particular job with Job Expiration; if the job has not been run for e.g. 60 seconds, the currently logged-in user is automatically logged out (and the user name and password dialog will come up when the job is started again). The default setting of zero disables this automatic logout. Note that Job Expiration requires either Brower or Forms login. Typically Job Expiration is used when ccScan is operating in “kiosk mode” where ccScan runs on a computer with a scanner that is shared among a work group.

The name of the newly scanned document is entered in Document Name. Optionally, the name of a folder located on SharePoint can be entered or selected in Document Folder (or selected from the list of existing folders). To allow for automated naming, ccScan supports the use of “macros” in document and folder names that automatically insert counter, barcode values, and OCR text pattern search results. Automated naming is used with high-capacity scanners whose hoppers can hold large stacks of documents separated by barcode sheets.

Next, select the desired SharePoint Document Library or List from the dropdown list in Document Library/List. The captured documents will be stored in this Document library or List. For each Document Library or List SharePoint allows the creation of input fields which typically are displayed as columns in the user interface. ccScan makes all defined input fields available to the job so they can be used to index the document (either manually or automatically using e.g. barcodes).

For those document libraries and lists that have a large number of input fields defined, ccScan allows SharePoint views to be used as “filters” in the sense that only the input fields defined in those views are made available. To take advantage of this capability make a selection from the provided dropdown lists in View.

When Prevent Overwrite is enabled, ccScan first checks whether the document or list entry already exists on SharePoint before storing the document (and report an error instead in case the document is already present). Also provided are a series of standard PDF settings, such as user and owner password, which will be applied to the captured documents before they are exported to SharePoint.

If your current Document Library is setup to support Versioning, ccScan provides two fields: Check-in Type and Check-in Comment. See Figure 4. New documents are automatically checked-in after uploading with proper input fields. Check-in Types available are Minor Version (default), Major Version or Overwrite. If versioning is not enabled, the type will be Major version. You may enter comments for each Check-in. Note that check-in/check-out does not apply to SharePoint Lists.

For each setting, the Administrator selects or enters a value. Default values are provided where possible. The Administrator can also make settings available for the operator to enter, except for the Administrator-only settings. For the settings to be entered by the operator, the Administrator determines the order in which these settings are shown,

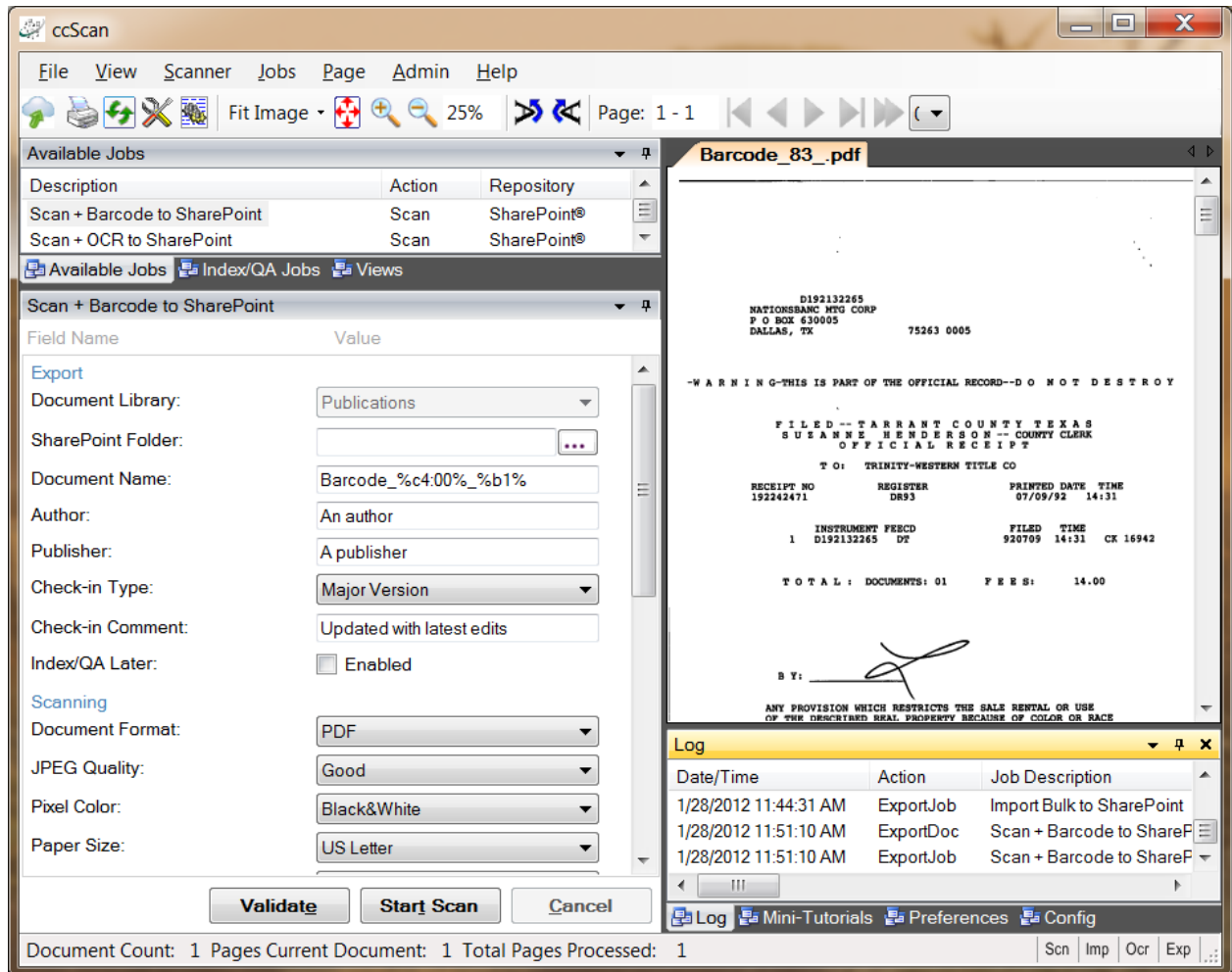
optionally modifies the default description of each setting, and selects which type of input control should be presented to the operator.

After the Export Settings, the Scanner Settings are entered. The scanning of documents in ccScan is based on the TWAIN standard. The scanner manufacturer provides TWAIN drivers. All the scanner settings you would expect in a production-level document capture system are provided: basic settings such as export format, resolution, paper size, and advanced settings such as barcode detection, which for example, can be used to automatically separate documents or automatically specify folder names. Jobs can also be associated with a button at the scanner front panel so that the job is started automatically with a press of the button.

TWAIN settings highly specific to the type of scanner can optionally be set with the TWAIN dialog box. These settings are saved locally on the computer and associated with the job. For those settings common between ccScan and the TWAIN dialog box, the ccScan settings can be given priority over the TWAIN settings. With this capability the administrator can give the operator access to the TWAIN dialog box and still, for example, enforce company standards such as 200 DPI resolution.

## Run this Scanning Job

Once the job has been created it can be run over and over again in User mode. The job is started either from the scanner front panel or with the Start Job button in ccScan.



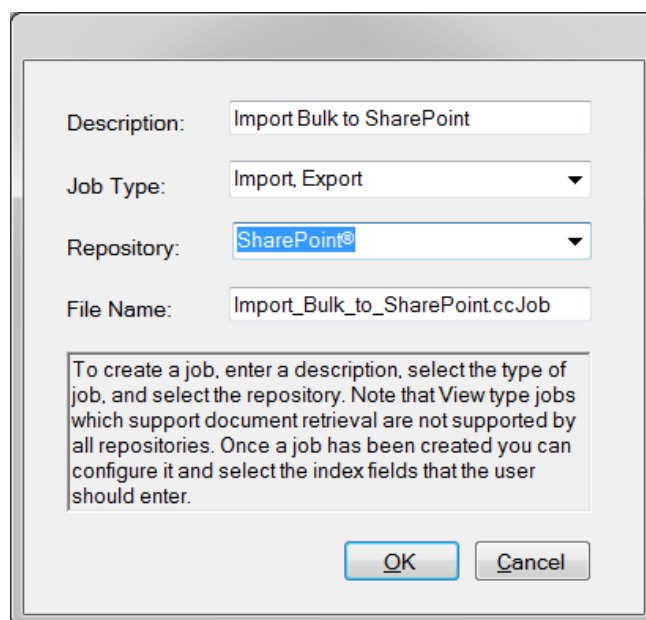
**Figure 4 - Run this Scanning Job**

Typically jobs are configured so that the operator must enter at least a document name or title, unless ccScan's automated naming capabilities are used. Optionally the operator can change default scanner settings or enter or select additional SharePoint fields.

Once the job is started, ccScan disables the Start button and does not allow any changes to the form. However, the operator is allowed to press the Cancel button. To provide visual feedback the scanned pages are displayed while the job is running.

## Create Import Job

To start a new Import job it must be created. In ccScan jobs are created in Setup mode. Once a job has been created, it can be used over and over again in regular User mode. The Administrator selects default values for the various job settings and decides which job settings can be entered or modified by the operator. The job settings are organized by category such as Import settings or Export settings.



Description: Import Bulk to SharePoint

Job Type: Import, Export

Repository: SharePoint®

File Name: Import\_Bulk\_to\_SharePoint.ccJob

To create a job, enter a description, select the type of job, and select the repository. Note that View type jobs which support document retrieval are not supported by all repositories. Once a job has been created you can configure it and select the index fields that the user should enter.

OK Cancel

**Figure 5 - Create New Import Job**

In the dialog box above SharePoint has been selected as the repository for export of the imported documents. If ccScan was licensed for multiple repositories the target repository for the job would be selected here from the list of available repositories. Note that the Job Type is Import, Export.

## Configure New Import Job

In the form below the Administrator can configure the new Import job. The same form is used to make modifications to the job later. The name of the newly imported document is entered in Document Name. If Document Name is left blank, the new document will get the same name as the imported file (or the name of the first file if multiple files are imported). All the SharePoint fields mentioned earlier for scanning are available for import as well.

The screenshot shows the 'ccScan' application window with the 'Setup Job: Import Bulk to SharePoint' configuration form. The form is organized into a table-like structure with columns for Field Name, Default Value, Show, Order, Input Control, and Edit Field Name.

Field Name	Default Value	Show	Order	Input Control	Edit Field Name
<b>Job</b>					
Description:	Import Bulk to SharePoint	<input type="checkbox"/>	0	TextBox	Description
<b>Export</b>					
Login Method:	Browser (Office 365)	<input type="checkbox"/>	0	ComboBox	Login Method
Host URL:	http://capturecomponent:	<input type="checkbox"/>	0	TextBox	Host URL
Logged-In User:		<input type="checkbox"/>	0	TextBox	Logged-In User
Named User:		<input type="checkbox"/>	0	TextBox	Named User
Document Library:	Publications	<input checked="" type="checkbox"/>	1	ComboBox	Document Library
View:	All Documents	<input type="checkbox"/>	0	ComboBox	View
SharePoint Folder:	Table Of Contents	<input checked="" type="checkbox"/>	2	FolderBox	SharePoint Folder
Document Name:	Deed%c4:00%	<input checked="" type="checkbox"/>	3	TextBox	Document Name
Author:		<input checked="" type="checkbox"/>	4	TextBox	Author
Publisher:		<input checked="" type="checkbox"/>	5	TextBox	Publisher
When Published:		<input type="checkbox"/>	0	TextBox	When Published

Buttons at the bottom: Validate, Save Job, Close. Bottom status bar: OK, Scn, Imp, Ocr, Exp.

**Figure 6 - Configure New Import Job**

For each setting, the Administrator selects or enters a value. Default values are provided where possible. The Administrator can also make settings available for the operator to enter, except for the Administrator-only settings. For the settings to be entered by the operator, the Administrator determines the order in which these settings are shown,

optionally modifies the default description of each setting, and selects what type of input control should be presented to the operator.

After the Export Settings, the Import Settings are entered. First the type of import must be selected, either File Import or Bulk Import. File Import brings up the standard Windows file selection dialog and allows the user to navigate to a particular folder and select one or more files to import.

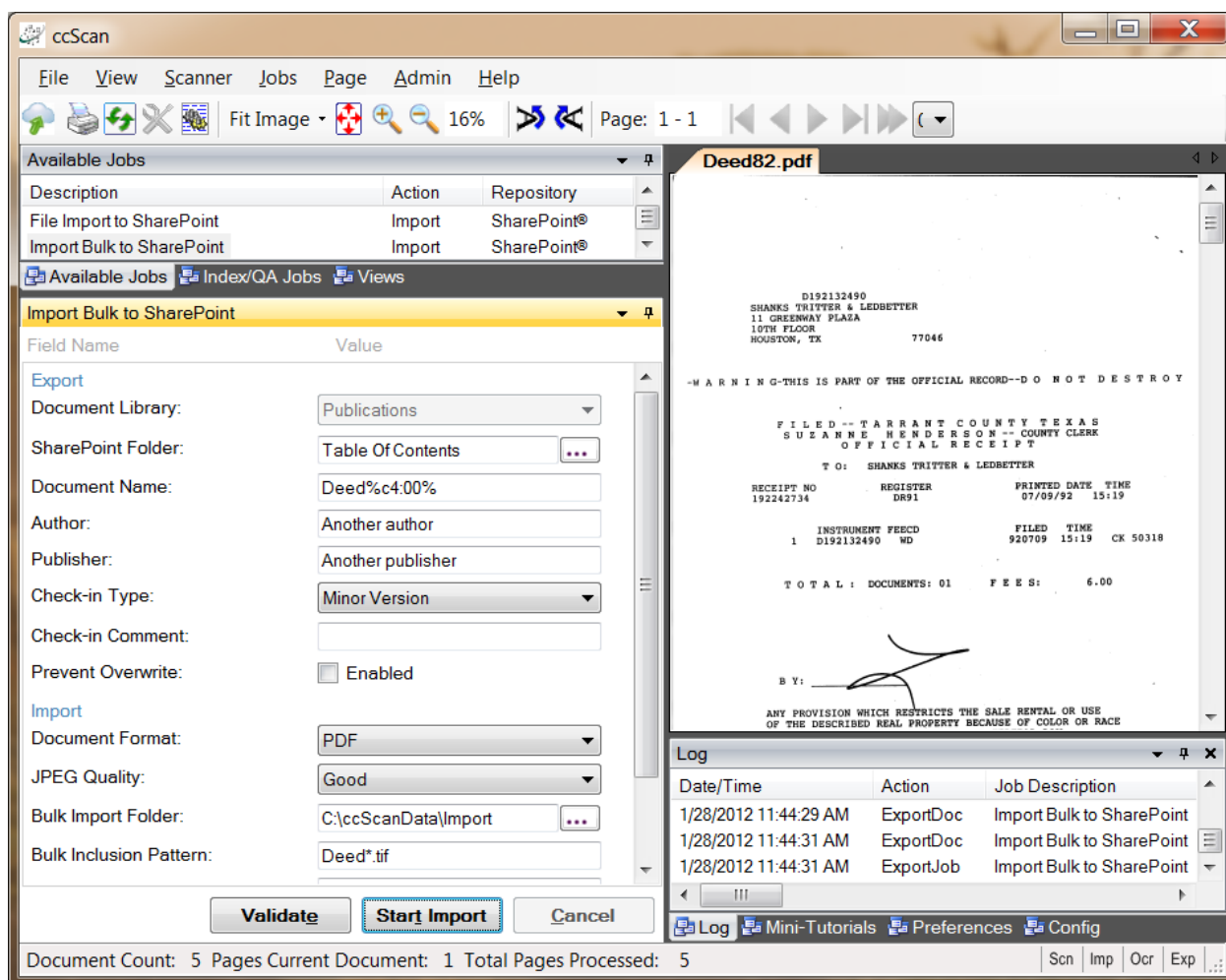
To import a large number of files in a single step select Bulk Import. Bulk Import allows the selection of a folder and for ccScan to automatically walk through this folder (optionally recursively so all subfolders and their subfolders are included) and import all files that match one more wildcards (such as \*.\* , \*.pdf, ccitt\_?.tif?, etc). Even wildcards can be specified to exclude certain types of files. Also, for files containing scanned images only, there is an option to combine all files into a single file and import only this combined file (optionally de-skewing the images first).

In Bulk Import Folder either enter the name of the folder to traverse (or select from a list of folders). The wildcard specification to include files is entered in Bulk Inclusion Pattern while the wildcard specification to exclude files is entered in Bulk Exclusion Pattern. To recursively traverse subfolders enable Bulk Folder Recursion.

To combine a number of separate image files (e.g. TIF, PNG, etc) into a single image file, check Combine Image Files.

## Run this Import Job

Once the Import job has been created it can be run over and over again in User mode. The job is started with the Start Job button.



**Figure 7 - Run this Import Job**

Typically jobs are configured so that the operator must enter at least a document name unless the new document must be named after the imported document. Optionally the operator can enter or select a SharePoint folder name where the imported document will be stored.

Once the job is started, ccScan disables the Start button and does not allow any changes to the form. However, the operator is allowed to press the Cancel button. To provide visual feedback, the imported pages are displayed while the job is running.

## Annotation and Redaction

It is not uncommon to annotate or redact selected paragraphs of documents that are shared. For example, redaction is used routinely in the legal sector. To allow review, correction, or annotation of documents before they are stored in SharePoint, ccScan provides an optional Index/QA mode. This mode can be enabled on a per-job basis by checking the Index/QA Later field. When enabled the documents produced with a Scanning or Import job are cached locally and made available in the Index/QA tab of ccScan.

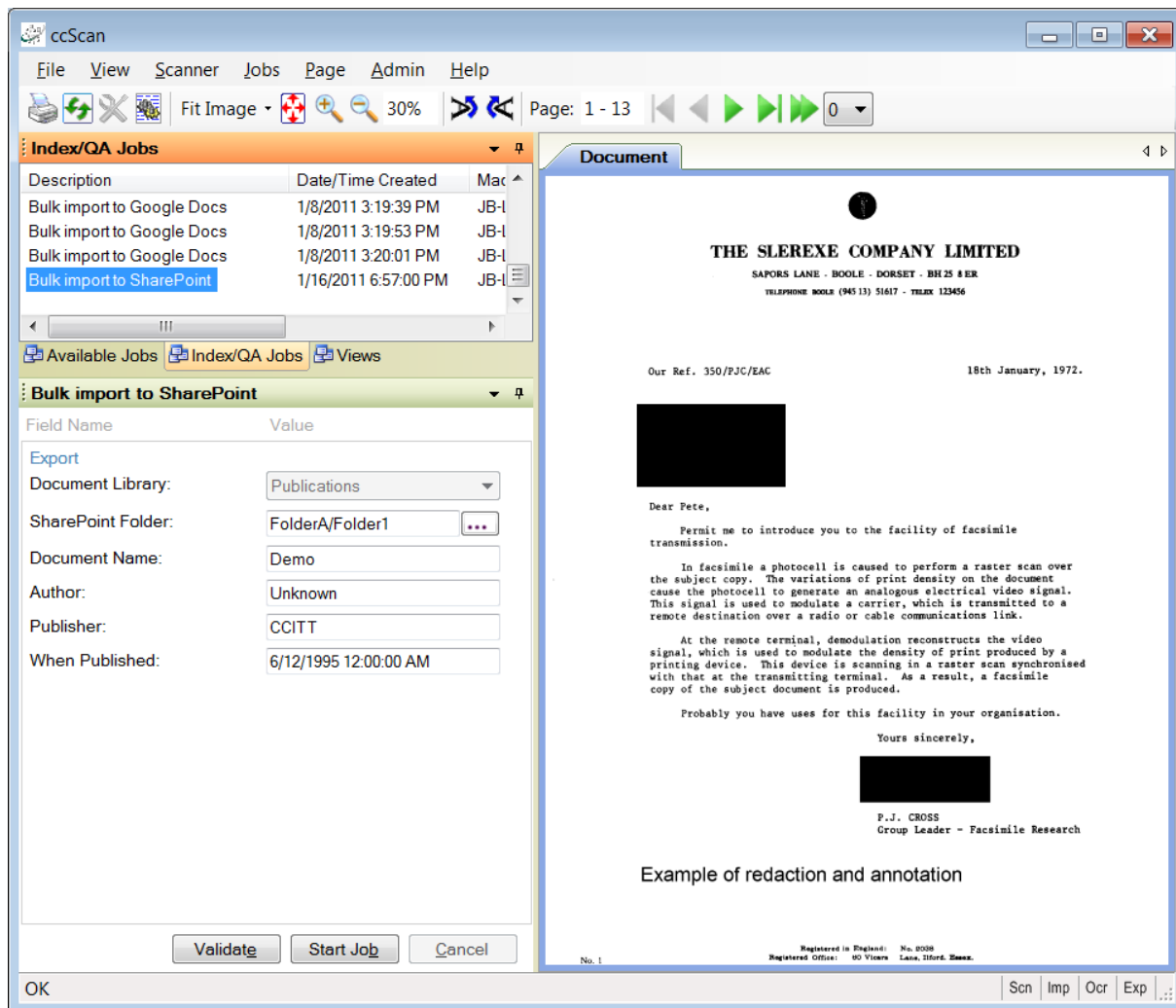


Figure 8 - Index/QA Mode

Before releasing the documents to SharePoint from Index/QA, descriptive information can be added to index fields, and pages can be deleted or inserted. Index information can be entered manually using the keyboard or extracted automatically from selected zones of the image using OCR (Optical Character Recognition).

In addition, pages can be annotated with text and selected portions of the pages could be redacted. Note that the annotations and redactions become permanent once they are saved and cannot be removed later.

## **Afterword**

*We at Capture Components LLC hope that this White Paper is of benefit to you. If it is and you are able to conveniently and efficiently capture and import your documents to SharePoint, then we have achieved a measure of success. If you choose to use ccScan as your document capture application, we will do our best to provide you with the highest quality product and support.*

Visit us at <http://www.capturecomponents.com>