32158 Camino Capistrano
Suite A   PMB 373
San Juan Capistrano, CA 92675
Sales@CaptureComponents.com
www.CaptureComponents.com

Scanning to Dropbox 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 Dropbox in an automated, time-saving manner.

Dropbox is a free service that lets you bring your photos, docs, and videos anywhere and share them easily. Dropbox was founded in 2007 by two MIT students tired of emailing files to themselves to work from more than one computer. Today, more than 45 million people across every continent use Dropbox to always have their files at hand, share with family and friends, synchronize between computers, phones, and tablets, and work on team projects. Since there is a lot more to Dropbox than this short introduction can cover please visit the Dropbox website at http://www.dropbox.com or the Dropbox Blog at http://blog.dropbox.com/ for in-depth information.

Although Dropbox 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 Dropbox, and uploading the document to Dropbox. 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 or running counters. In addition to scanning paper documents, ccScan has bulk import capabilities for already digitized documents that allow entire folders to be uploaded to Dropbox in a single step using wild card file specifications and recursion.

ccScan for Dropbox

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, documents can be captured remotely and stored in Microsoft SharePoint and SharePoint Online using SharePoint's Internet-based web services interface. Additionally ccScan works with the Microsoft Windows file system utilizing its local hard drive or networked folders for storage.

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 Dropbox 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 Dropbox. 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 Dropbox since Dropbox natively provides document viewing capabilities.

ccScan uses the Dropbox REST API to upload documents to Dropbox. The Dropbox REST API is also used to create, rename, and delete folders. Although there is no need to know anything about the API to use ccScan, Dropbox users interested in learning more about this API can visit https://www.dropbox.com/developers/reference/api.
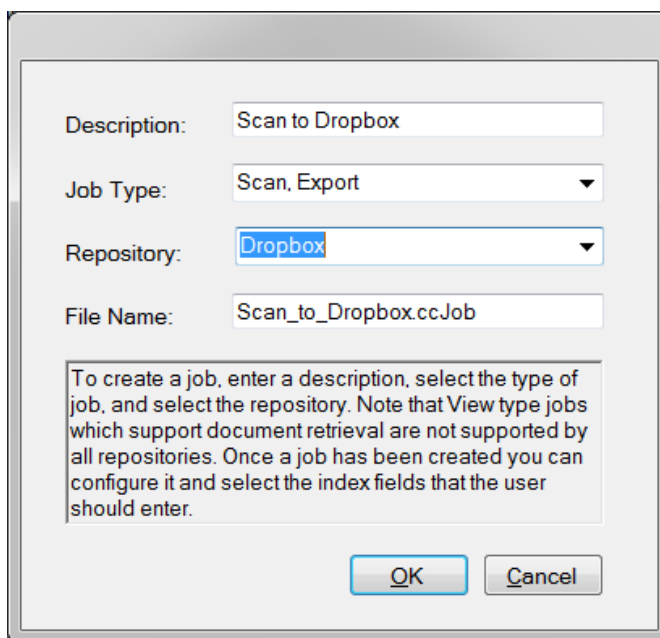
Despite its low price, ccScan has all the high-end features and no restrictions such as page click charges. For detailed information about ccScan or to download a free trial version, please visit Capture Component's website at www.capturecomponents.com.

Typical Scenario

To give the reader a flavor of capturing and storing documents in Dropbox 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 Dropbox, 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 Dropbox 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.

Description: Scan to Dropbox

Job Type: Scan, Export ▼

Repository: Dropbox ▼

File Name: Scan_to_Dropbox.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.
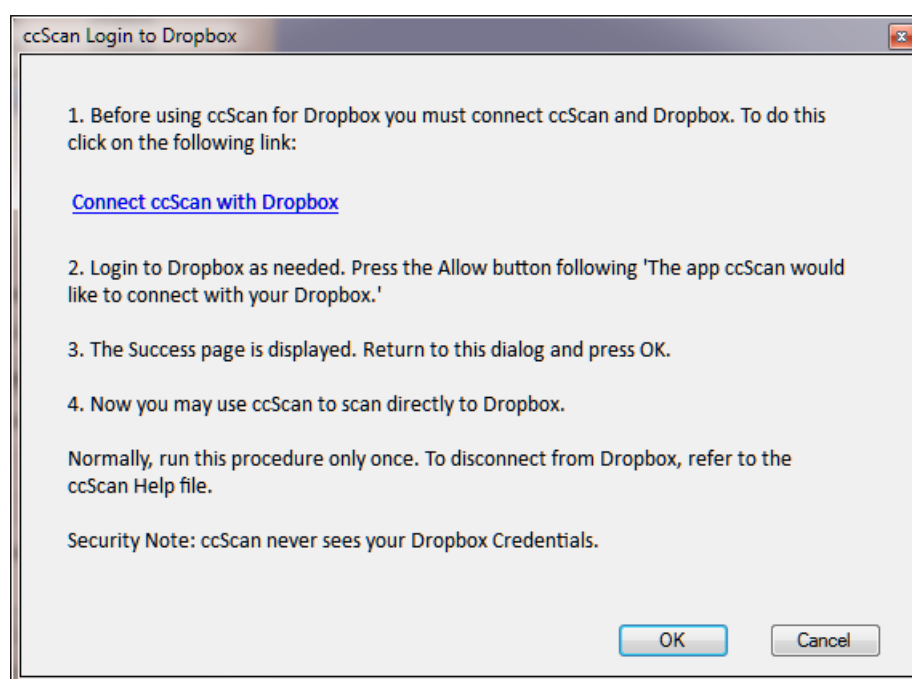
OK   Cancel

Figure 1 - Create New Scanning Job

In the dialog box above Dropbox 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.
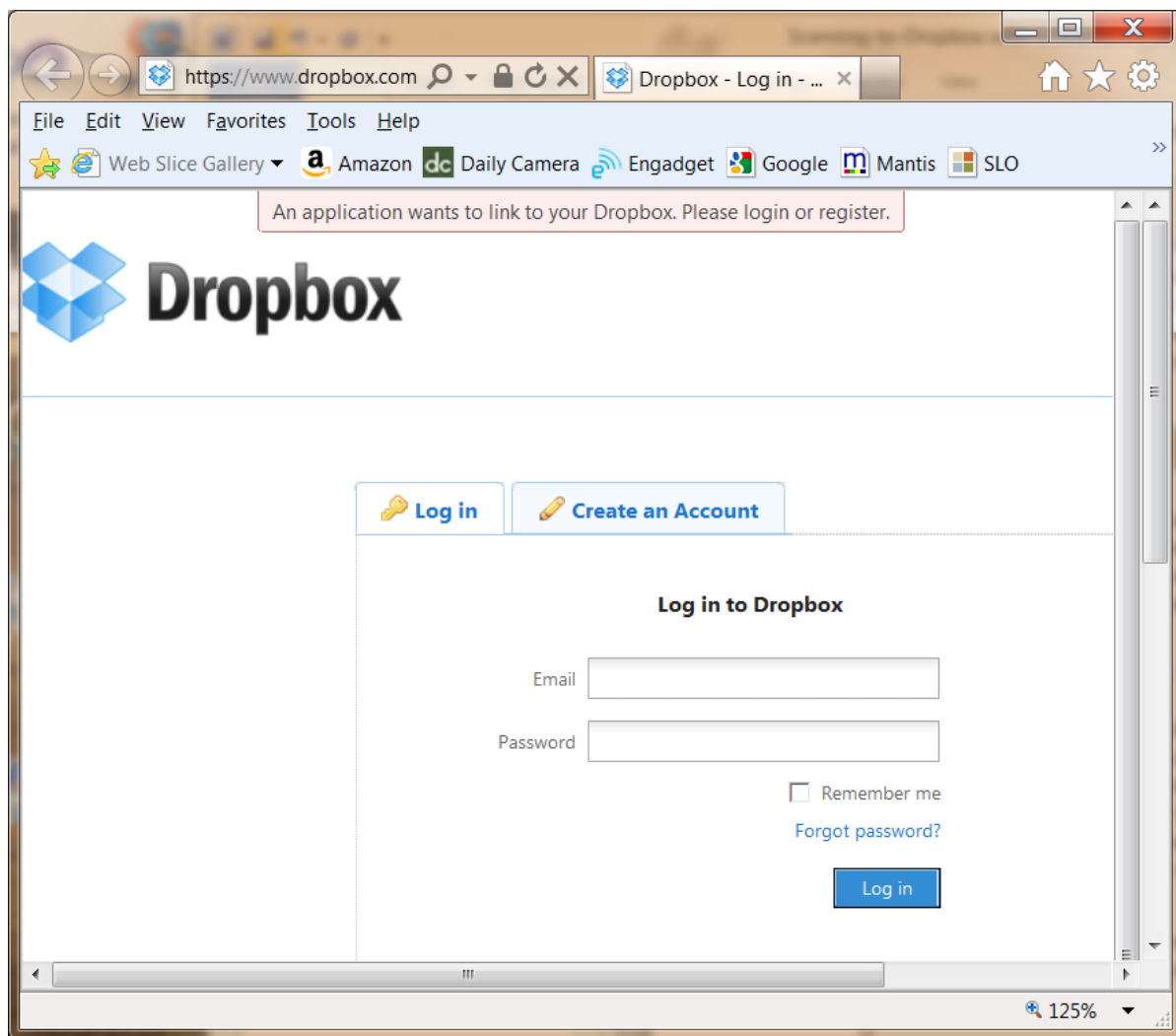
Enter Dropbox Credentials

Immediately after the new job has been created, the login dialog below comes up to ask you to bring up the default Internet browser on your system so you can enter the email address and password for a Dropbox account.

Both in User mode and Setup mode the operator must enter an email address and password of an active Dropbox account. When the login or register in the Internet browser page has been completed (see Figure 3) and the OK button in the dialog is pressed, ccScan obtains an authorization token from Dropbox that is subsequently used to request information from Dropbox such as a list of existing folders on Dropbox.

ccScan Login to Dropbox

1. Before using ccScan for Dropbox you must connect ccScan and Dropbox. To do this click on the following link:

Connect ccScan with Dropbox

2. Login to Dropbox as needed. Press the Allow button following 'The app ccScan would like to connect with your Dropbox.'

3. The Success page is displayed. Return to this dialog and press OK.

4. Now you may use ccScan to scan directly to Dropbox.

Normally, run this procedure only once. To disconnect from Dropbox, refer to the ccScan Help file.

Security Note: ccScan never sees your Dropbox Credentials.

OK   Cancel

Figure 2 – Login Dialog and Instructions

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. At any time the user has the option to revoke the authorization token on the Dropbox website. Note that ccScan can be configured to disable the default auto-logon which requires that users re-enter their Dropbox credentials each time they use ccScan.

Figure 3 – Enter Dropbox Credentials with the Browser

To allow the administrator flexibility jobs can be configured such that only a specific Dropbox user can run a particular job. Normally jobs can be run by any logged-in Dropbox user. In addition to this "named user" feature, ccScan allows the administrator to set a time-out for a particular job; if the job has not been run for e.g. 60 seconds, the currently logged-in user is automatically logged out (and the login dialog in Figure 2 will come up when the job is started again).

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 name of the newly scanned document is entered in Document Name. Optionally, the name of a folder located on Dropbox can be entered or selected in Dropbox Folder (or selected from the list of existing folders).

Figure 4 - Configure New Scanning Job

To allow for automated naming, ccScan supports the use of "macros" in document and folder names that automatically insert counter and barcode values. Automated naming is

used with high-capacity scanners whose hoppers can hold large stacks of documents separated by barcode sheets.

The Upload Document settings determine what happens if a document already exists on Dropbox: Overwrite Existing, Rename New Upload, or Conflict If Error (which means stop the job in progress and report an error). Named User can be set to the username of an existing Dropbox user, which ensures that only this particular user can run the job. Job Expiration can be set to the number of seconds of idle time after which the logged-in Dropbox user is automatically logged out. The default setting of zero disables this automatic logout.

When Shareable Link is checked, ccScan will request a shareable URL from Dropbox for the uploaded document. The URL and its expiration date are returned in read-only fields that could be copied and pasted. For example, the URL could be copied in an email to share the document with anyone, even non-Dropbox users.

These return values are also recorded in the daily log file that ccScan produces to allow automated post-processing. This JSON format log file contains information about all jobs that were run on a particular day and the Dropbox return values are recorded for each document in the job. Typically this capability is used in environments where Dropbox is used for document storage but links to the documents are stored in a database.

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.
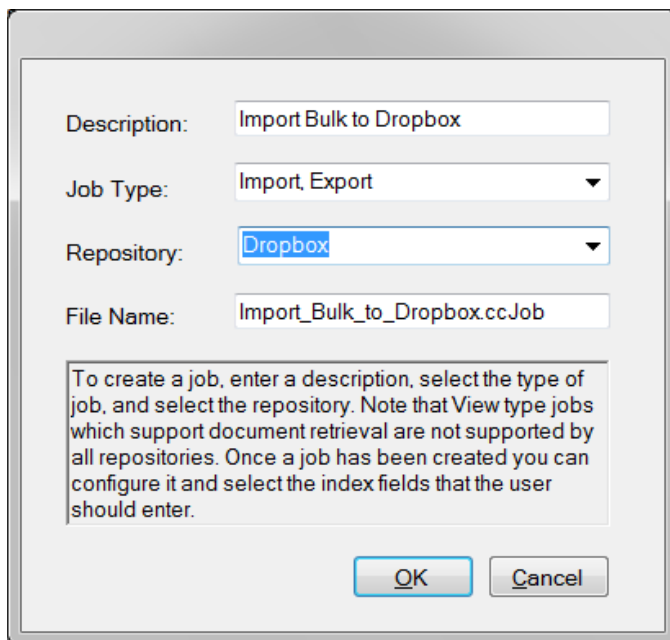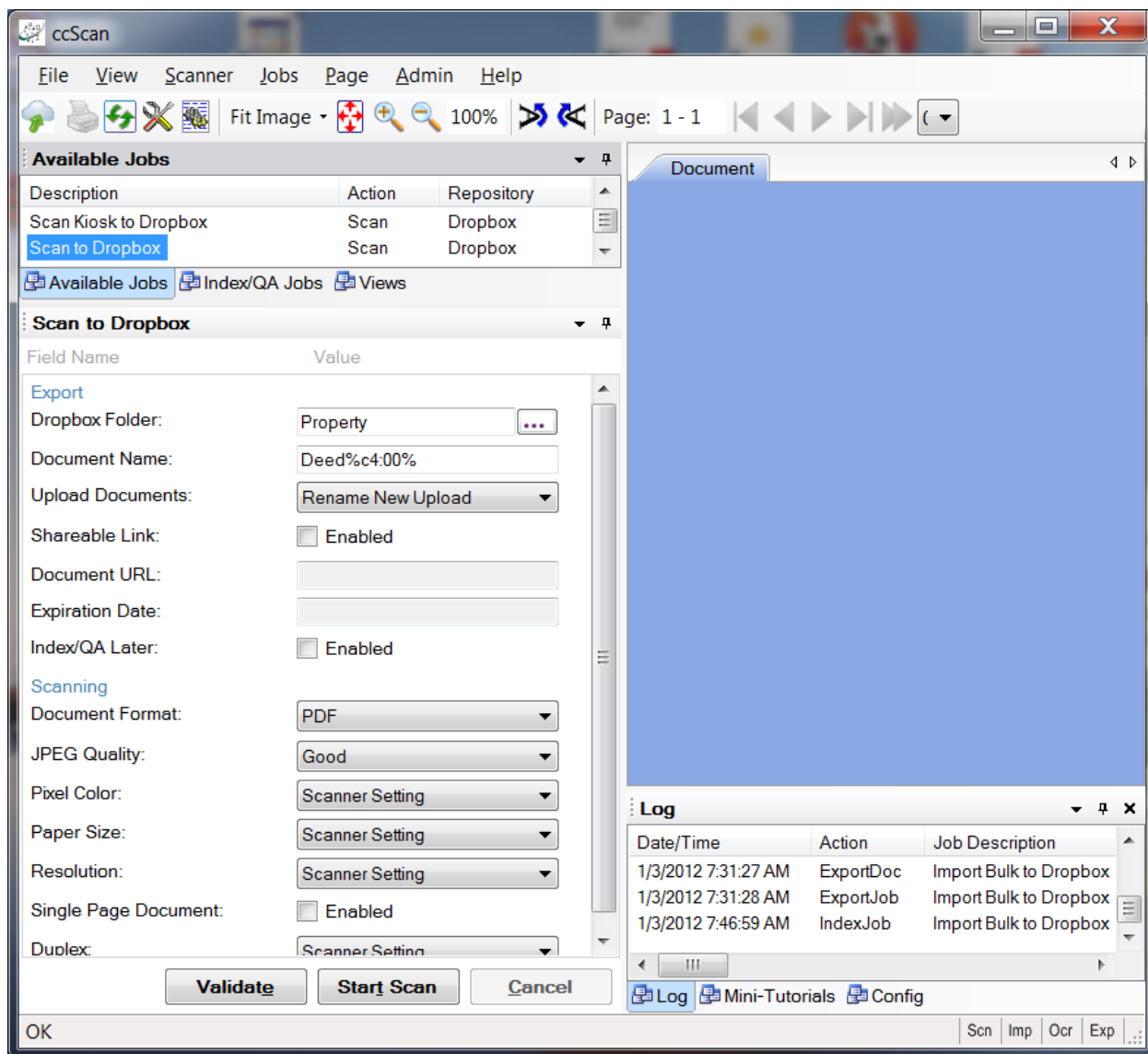

Figure 5 - Run this Scanning Job

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

Figure 6 - Create New Import Job

In the dialog box above Dropbox 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.

Enter Dropbox Credentials

Immediately after the new job has been created and the Dropbox credentials have not been entered yet, the email address and password for a Dropbox account must be entered with the dialog described earlier in this document.

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.

Figure 7 - Configure New Import Job

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 Dropbox fields mentioned earlier for scanning are available for import as well.

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.

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.

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.
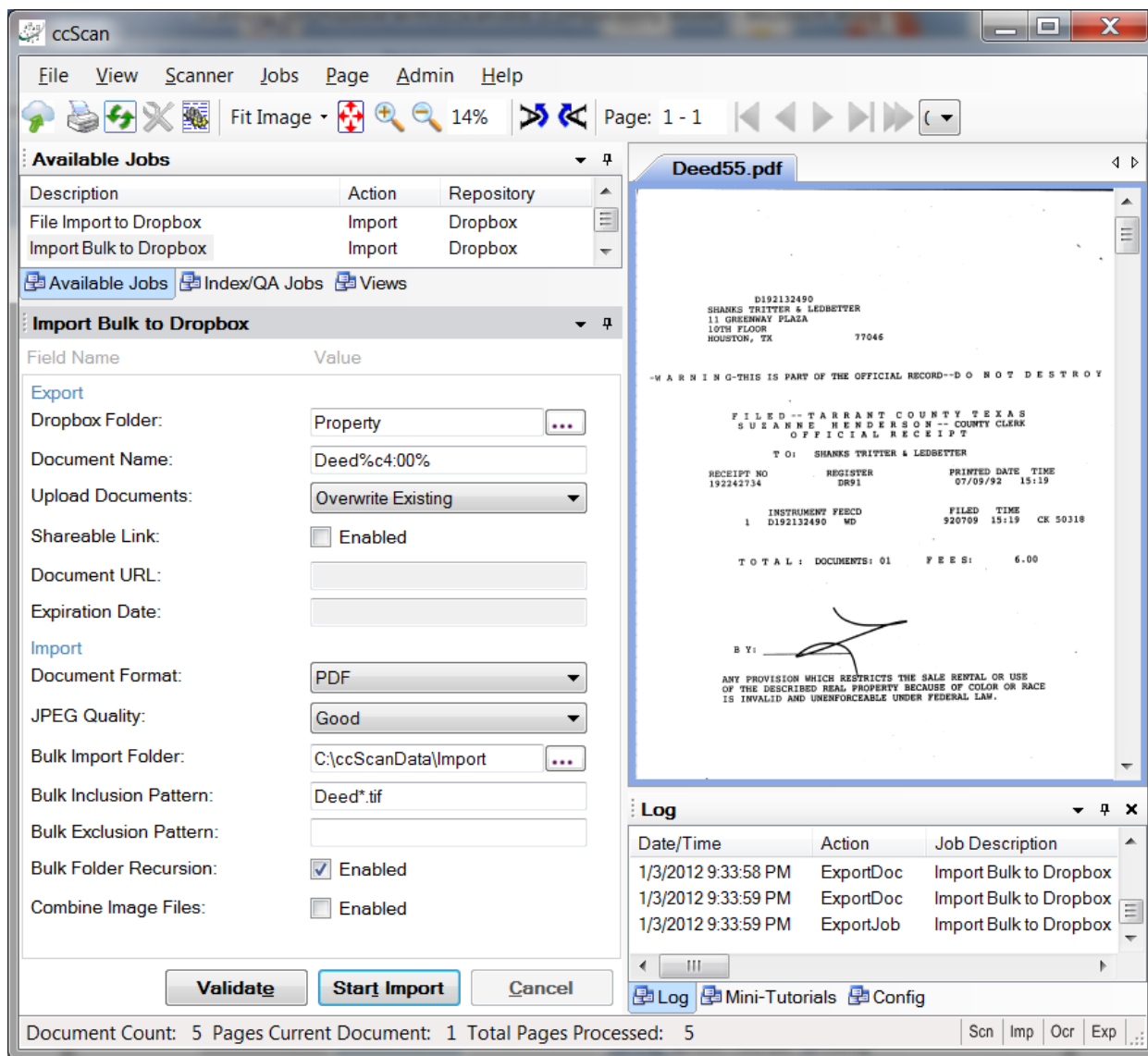
Figure 8 - Run this Import Job

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 Dropbox 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 Dropbox, 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.
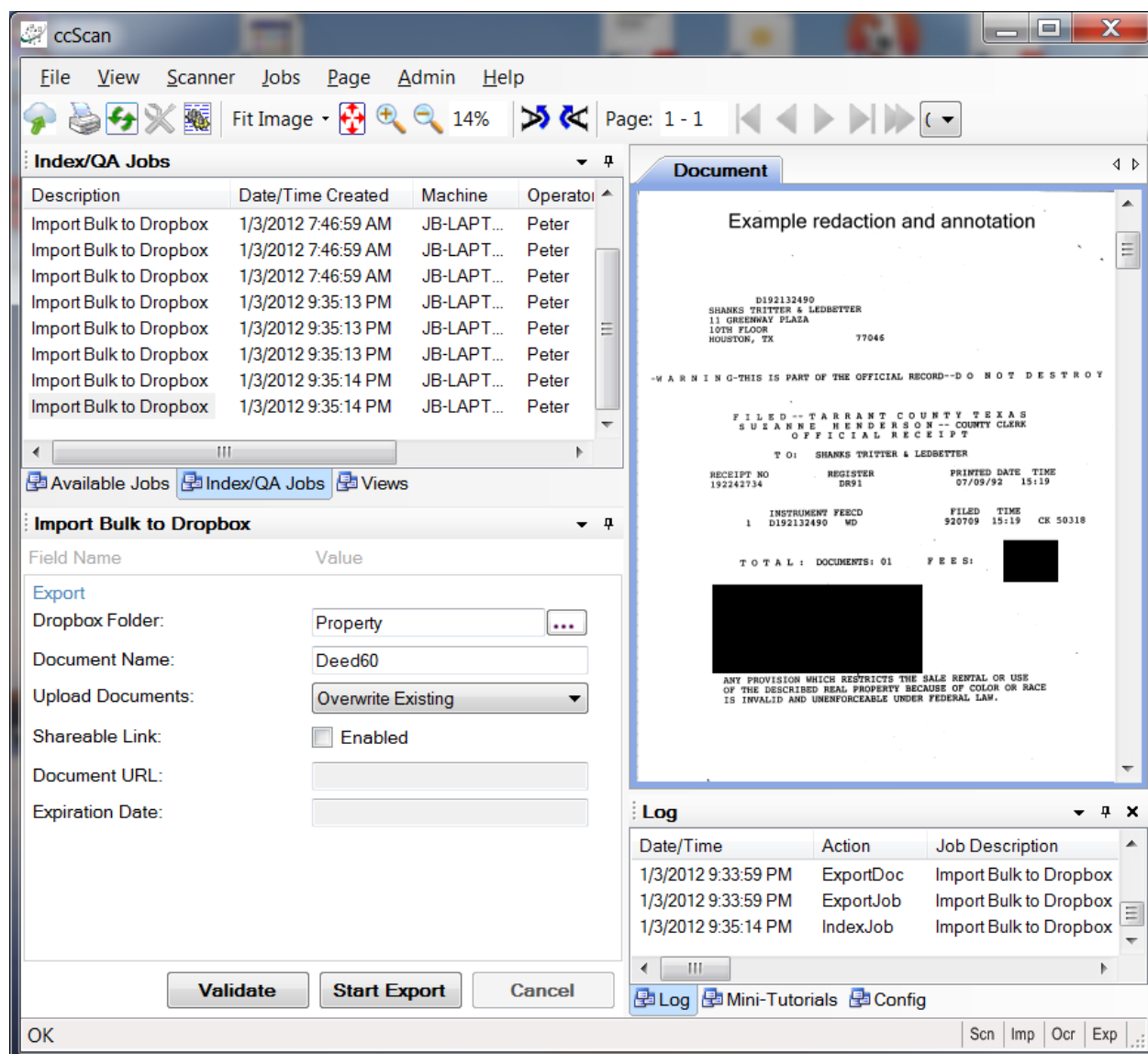

Figure 9 - Index/QA Mode

Before releasing the documents to Dropbox 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 Dropbox, 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