

Using PDF Files in CONTENTdm

CONTENTdm® uses the Adobe® PDF Library™ to provide features for efficient processing of born-digital documents in Portable Document Format (PDF). PDF files and PDF compound objects can be displayed on the CONTENTdm Website in the Item Viewer and Compound Object Viewer.

The PDF features include:

- Creation of thumbnail images from PDF files
- Full text extraction (converted to UTF-8 format)
- Automatic conversion of multiple-page PDF files into monograph compound objects
- Pages of a compound object automatically generated from a PDF file do not count toward the total number of items on the server
- Search term highlighting within the PDF file
- Configurable options for PDF file display in your website, using the Website Configuration Tool (6.x Website only)

This tutorial provides an overview for using PDF files in CONTENTdm and steps through using the Project Client to import a multiple-page PDF file as a single item and then as a compound object. (If you use Connexion digital import to add items to your collection, see the **Using OCLC Connexion Digital Import** tutorial for more information. If you use the Web interface to add items, see **Adding Items** in the Help files.) This tutorial also steps through editing PDF items using the Project Client and the Web interface.

This tutorial assumes that you are running CONTENTdm version 6 or later. It also assumes that you have installed Adobe Reader version 8 or later on the computer running the Project Client.

Learn About

Section I: Overview

About PDF Files

Section II: Before You Begin

Choosing the Format

Configuring the Website

Confirming Collection Configuration

Section III: Importing PDF Files Using the Project Client

Importing a Single-Item PDF File

Importing a Multiple-Page PDF File as a Compound Object

Section IV: Editing PDF Files

Using the Project Client

Using the Web Interface

Section V: Troubleshooting

Section I Overview

About PDF Files

A single PDF file can contain many pages. Regardless of the number of pages, it is a single file and is uploaded as a single file. You can import multiple PDF files using the Add Multiple Items wizard in the Project Client.

Depending on how your collection is configured, multiple-page PDF files can be added to your collection to be viewed as single items or, if PDF conversion is enabled, they can be automatically converted to PDF compound objects (of the type monograph). When multiple-page PDF files are automatically converted to compound objects, each page of the PDF becomes a page with its own metadata record and each page is navigable using the features of the Compound Object Viewer.

Note: To ensure an optimal end-user experience, PDF files (or pages of a compound object) larger than 20 MB are not loaded inline in the CONTENTdm 6.x Website. These larger files can be saved to the desktop or opened outside of the browser.

To support full text searching, text from PDF files is extracted and placed in a full text search field when PDF files are approved and added to a collection. Automatic text extraction occurs when:

- The file extension of the item is .pdf.
- The text is embedded in the PDF file. (To check whether your PDF file has embedded text, save it as a .txt file. If the text file contains the text, then the PDF has embedded text.)

Note: If your PDF file was created from a born-digital document, such as a Microsoft Word file, it will almost always have embedded text. If your PDF file was created from scanned TIFF images, it does not have embedded text unless you have taken the additional step to OCR the image (or PDF file) and add that text to the PDF.

CONTENTdm supports integrated OCR functionality through the **OCR Extension**. Using the OCR Extension, full text can be generated from JPEG2000, JPEG, GIF and TIFF files. OCR is not supported for PDF files. (The automatic text extraction for PDF files mentioned above is separate functionality and does not require the OCR Extension.)

- The collection that the PDF is being added to has a full text search field.
- The full text search field is empty when the item is added to the collection.

Thumbnail images can be automatically generated for PDF files, or you can specify a custom thumbnail.

Single-Item PDF Files

Single-item PDF files are created for PDF files that only contain one page. Single-item PDF files are also created by default for multiple-page PDF files, unless your CONTENTdm administrator has configured the collection for PDF conversion. (This setting can be turned on and off for each collection. Or you can set the PDF conversion option for your project in the Project Settings Manager of the Project Client.)

In the CONTENTdm 6.x Website, PDF files are displayed in the website in the CONTENTdm viewers by default using Adobe Reader, but you can use the Website Configuration Tool **Page Types > PDF Files** options to set the website display to link to the PDF file so that it opens in Adobe Reader in a new window (your website remains open in a separate window). This option is not configurable in the CONTENTdm Responsive Website.

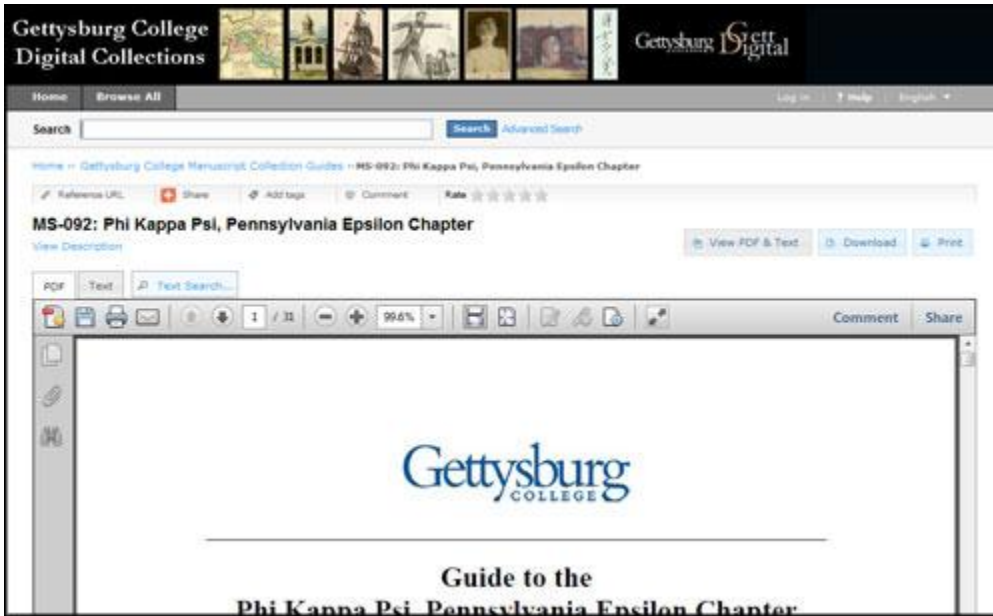


Figure 1: Default single-item PDF display in the CONTENTdm 6.x Website

You also can use the Website Configuration Tool **Page Types > PDF Files** options to configure whether portions of the Adobe Reader interface display (embedded version) and whether search terms are passed through to the Adobe Reader search function (embedded version).

Open source **Expand** Vi

PDF Compound Objects

PDF compound objects are automatically created when multiple-page PDF files are added and approved to a collection, if that collection has been configured to enable PDF conversion (or if you have enabled it for your project in the Project Settings Manager of the Project Client).

The page order of the PDF compound object matches the page order of the original, multiple-page PDF file, and the end-user can navigate and view the document using the features of the Compound Object Viewer.

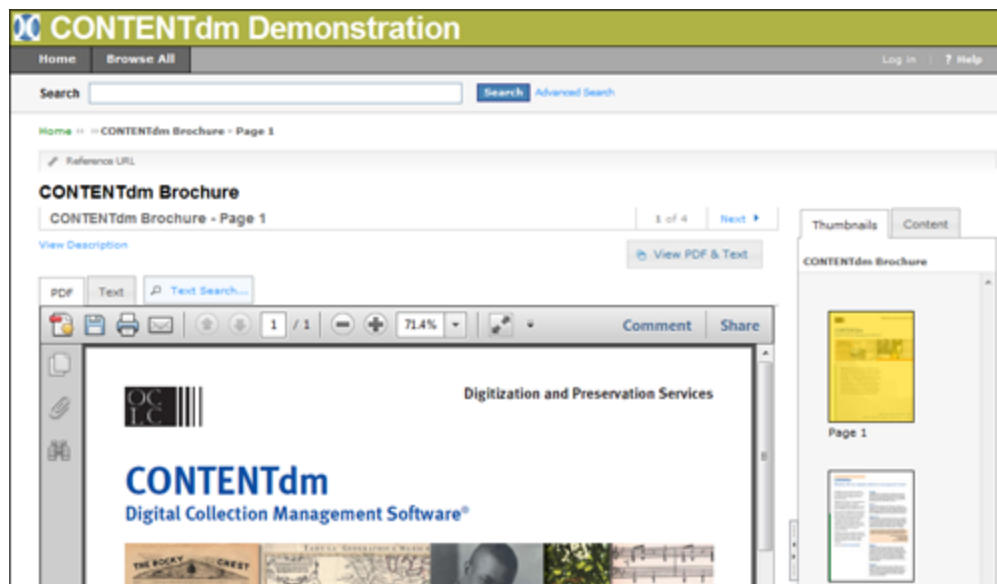


Figure 2: Multiple-page PDF file displayed in the Compound Object Viewer

Additionally, search term matches are highlighted in the PDF file in Adobe Reader and passed through to the Reader search function. (Search terms are also highlighted in the Page & Text Viewer and in the PDF inline display in the Item Viewer.)

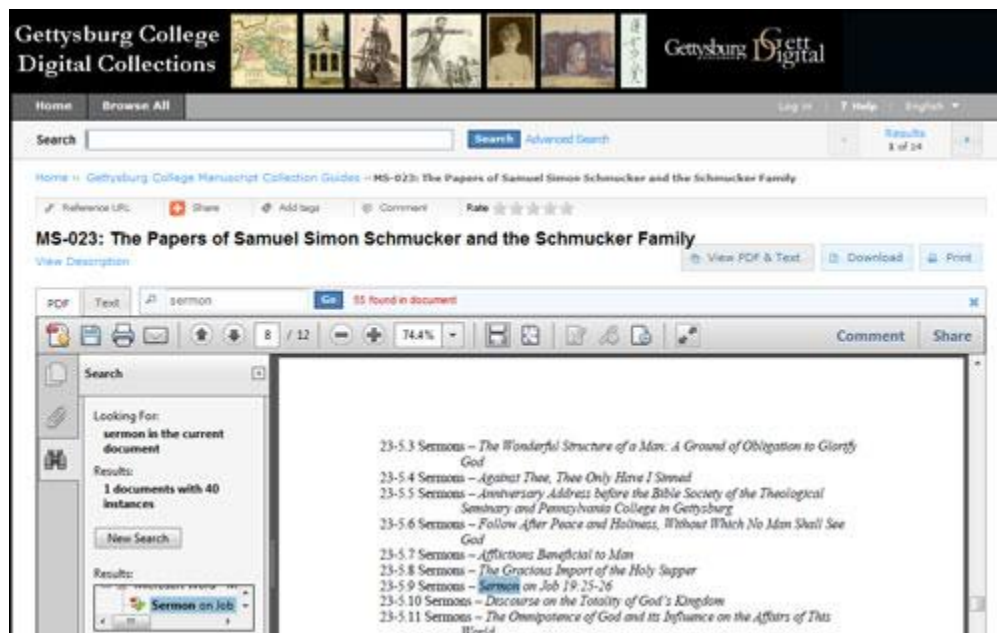


Figure 3: Search term matches are highlighted in the PDF file and passed through to Adobe Reader search function

Each page of the PDF file has its own metadata record after it is converted to a compound object, but the digital item associated with it in CONTENTdm is virtual (i.e., a link to the related page in the PDF file). The individual pages of PDF files do not exist separately on the server; they are extracted and displayed only when the user requests them. This improves the end-user's access speed because the entire PDF file does not have to download to display the requested page. You cannot set permissions on individual pages or delete individual pages. When you set permissions for the PDF compound object they apply to all pages; and to delete pages, you must delete the entire PDF compound object.

You also cannot edit or restructure the individual pages of PDF compound objects unless you remove the PDF file from the collection, edit the original PDF, and then add it to the collection again.

When a multiple-page PDF is added to a project in the Project Client, you can create compound object-level metadata by editing the record in the Project Spreadsheet or the Item Editing tab. When the multiple-page PDF file is added to the collection, text is extracted from each page and added to the full text search field in the associated page-level metadata records.

Thumbnail images are generated automatically for each page the PDF compound object.. The thumbnail that represents the PDF compound object itself is based on the first page of the PDF file. (If you choose to use a custom thumbnail for a PDF compound object, the custom thumbnail is used for the compound object, as well as for each page of the object.)

Note: Multiple-page PDF files that are imported as a part of a compound object are not converted to compound objects, even if PDF conversion is enabled for the collection. However, text is extracted and thumbnail images are generated.

Section II Before You Begin

Before you import PDF files to your collection, first determine the best format for your collection, and then confirm your style file settings for display and your collection configuration.

Choosing the Format

Before you decide to use PDF over another format, consider whether your source materials are well-suited to this format, and whether your end-user experience would be optimized by using PDF. For example, PDF files are ideal for documents that were initially created as digital documentation, such as theses and city council minutes originally saved as .doc files. PDF files may not provide an optimal end-user experience for scanned images, books, maps or newspapers.

Additionally, PDF is not ideal for scanned images because an item that has been scanned does not automatically contain embedded text. For scanned images, you can use the CONTENTdm OCR Extension for generating full text. PDF files created from images can be very large and slow to download for online viewing. For a better end-user experience, you can use CONTENTdm to create JPEG2000 or JPEG derivatives from scanned TIFF files, rather than converting the TIFF files to PDF files.

If you determine that PDF is the right format, you then need to decide whether to display the PDF file as a single-item (and if it should use the default display or display inline) or, if it consists of multiple pages, whether it should be automatically converted to a PDF compound object.

Note: For maintenance and for consistency in your collection display, you may also want to consider converting any existing PDF files in your collection so that they display and function similarly to those you add with the features introduced in CONTENTdm version 4.3. For more information, see the introduction to this tutorial.

Configuring the Website

After you have reviewed the content of your collection and determined the best overall format, you may need to edit the default Configurations in the Website Configuration Tool. For more information, see **Page Types**.

Confirming Collection Configuration

Before you import PDF files, you need to confirm your collection settings because they determine how your PDF files are processed when they are added to the collection. The CONTENTdm administrator can confirm and change collection configuration and metadata field settings by using CONTENTdm Administration.

Settings to confirm:

- **PDF conversion:** If you want to automatically convert multiple-page PDF files to compound objects when they are added to the collection, confirm that the PDF Conversion setting is set to **Yes** on the Collection Configuration page in CONTENTdm Administration. For more information, see **Collection Configuration** in the Help files.
- **Metadata fields:** If you want text extracted for searching, you must set one metadata field to the data type Full Text Search. This field holds the PDF text extraction that happens for each page of the PDF compound object or the first 128,000 characters of a single-item PDF file. For more information, see **Editing Field Properties** in the Help files.

Section III Importing PDF Files Using the Project Client

PDF files can be imported using the single or batch operations in the Project Client. The following scenarios step you through importing single PDF files to existing collections. For more information about batch operations, see **Add Multiple Items**.

When importing a file into your project, the original file remains in its existing location. A copy is added to the project.

Note: To view PDF files in the Project Client, you must have Adobe Reader installed.

Importing a Single-Item PDF File

You can import PDF files to display in your collections as single items, whether they contain one or more pages. In this example we will import a multiple-page PDF file.

Step 1: Confirm that Adobe Reader is installed

Because we want to be able to view the PDF in the Project Client, we first confirm that Adobe Reader is installed on the Project Client computer. (To download and install the latest version of Adobe Reader, see <http://get.adobe.com/reader/>.)

Step 2: Confirm source file type and security settings

Confirm that the file name for the item you want to add has the extension .pdf. In this example, the source document is named Census92.pdf. (If the file extension is .doc or another format, you will need to first convert to PDF by using software such as Adobe Acrobat®.)

Next confirm that the PDF file is not password-protected, encrypted or locked. (If it is encrypted or locked, a thumbnail may not be generated and text is not extracted in CONTENTdm.) One way to test for security settings is to try opening the document on your computer using Adobe Reader. You may see a message similar to the following.



Figure 4: Adobe Reader security warning

If you need to modify the security settings, you may need permission to change the settings from the document owner, and then use Adobe Acrobat to edit and save the PDF file.

Step 3: Confirm collection settings

Because we have decided to add this multiple-page PDF file as a single-item, we confirm with the CONTENTdm administrator that the collection has PDF conversion disabled for multiple-page PDF files in CONTENTdm Administration. (This setting is displayed in CONTENTdm Administration under the Collections tab on the Collection Configuration page.)

Note: See also the following Confirm Project Setting step because some settings can be set locally in the Project Client to override the server settings.

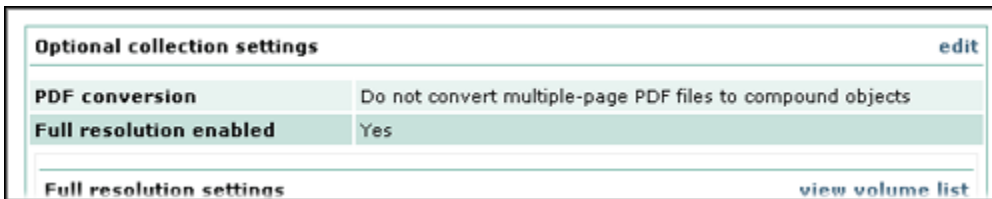


Figure 5: PDF conversion setting disabled for the collection in CONTENTdm Administration

We want the first 128,000 characters extracted for full text search, so we also need to confirm that one metadata field has been set to the type Full Text Search.

For this example, we have set the Description field for the collection to the data type Full Text Search in CONTENTdm Administration.

CONTENTdm Administration

admin home | server | collections | items

profile | fields | website | reports | export | view collection | help

Current collection: Medieval Architecture [change](#)

Metadata fields

View and configure collection and administrative fields.

Collection field properties

View, add, edit and delete fields. Enable full text searching and controlled vocabulary. After you have added, changed, or deleted fields, index the collection to update changes.

	Field name	DC map	Data type	Large	Search	Hide	Required	Vocab	add field
1	Title	Title	Text	No	Yes	No	Yes	No	move to ▼ edit delete
2	Subject	Subject	Text	No	Yes	No	No	No	move to ▼ edit delete
3	Description	Description	Full Text Search	Yes	Yes	No	No	No	move to ▼ edit delete
4	Creator	Creator	Text	No	No	No	No	No	move to ▼ edit delete
5	Publisher	Publisher	Text	No	No	No	No	No	move to ▼ edit delete
6	Contributors	Contributors	Text	No	No	No	No	No	move to ▼ edit delete

Figure 6: Metadata Fields page in CONTENTdm Administration

Step 4: Confirm project settings

Although we confirmed settings on the CONTENTdm Server, settings can be set for the project in the Project Client. These settings can override the server settings.

1. **Access the Project Settings Manager**, and click the **Processing** tab. The Processing Settings page displays.

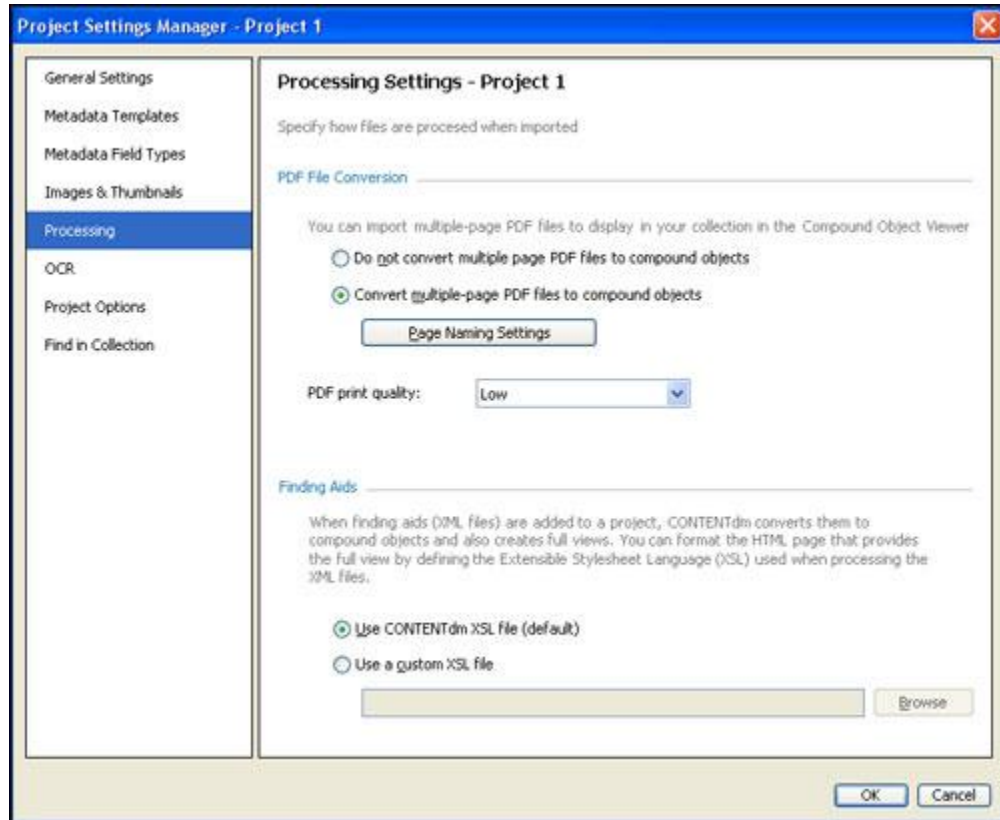


Figure 7: Processing Settings

2. Confirm that PDF File Conversion is set to **Do not convert multiple-page PDF files to compound objects**.
3. Check the PDF Print Quality setting.

This setting enables you to specify the quality of the optional print versions (PDF) of the compound object. The quality is determined by this setting in relation to the quality of the original source material. Select a higher-quality setting for items where the PDF quality is important to your end users, or select lower-quality settings to improve processing and rendering times of the user experience. (For more information, see **Processing**.)

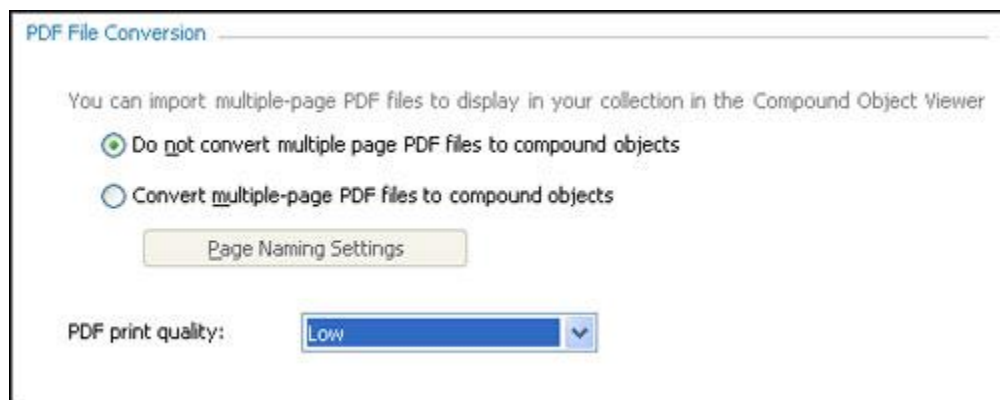


Figure 8: PDF file conversion settings

4. Click **OK** to save any changes.

Step 5: Add the PDF file to your project

The general steps for adding a single-item PDF file are the same as adding other individual files.

1. Navigate to the Home tab or Project tab.
2. On the left task pane under **Common Tasks**, click **Add Item**. The Add Item wizard displays.

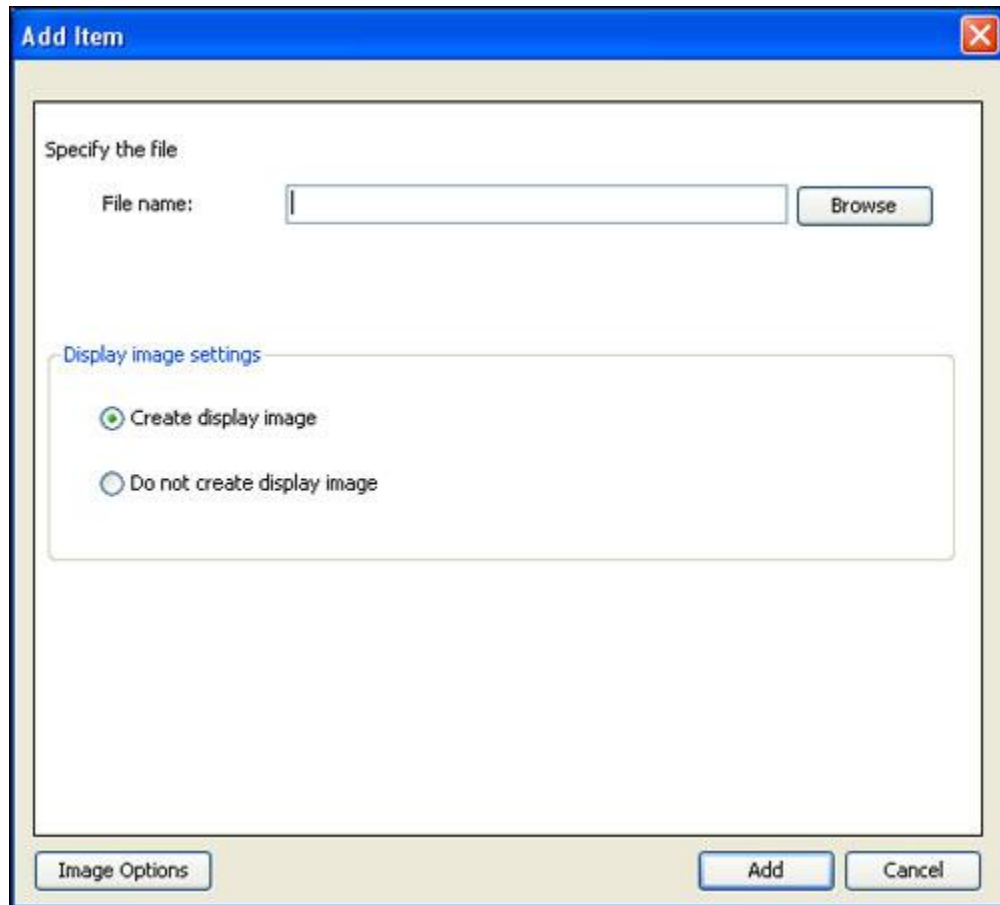


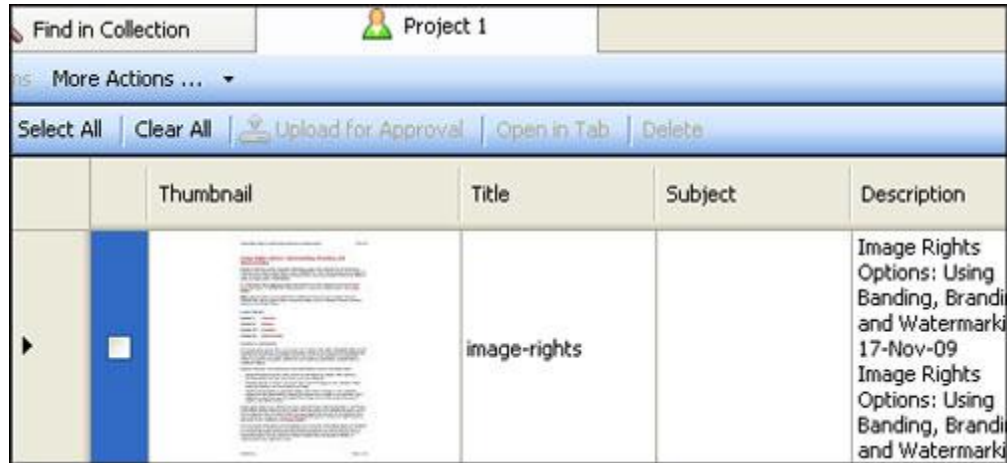
Figure 9: Adding an item

3. Specify or browse to the file you want to add.

Tip: If you want to add more than one item, press Ctrl and click on the individual files to add.

4. Select whether to create a display image.
5. Click **Add**. A progress screen displays as the item is processed. When complete, a summary screen displays the summary and any errors or warnings.
6. Click **Close**. The item is displayed in the project spreadsheet.

Note: The original file remains in its current location. A copy is added to the project spreadsheet.




	Thumbnail	Title	Subject	Description
		image-rights		Image Rights Options: Using Banding, Brandi and Watermark 17-Nov-09 Image Rights Options: Using Banding, Brandi and Watermark

Figure 10: Project spreadsheet

- In the Project Spreadsheet, edit the title in the required **Title** field and enter other metadata. (Or open the item in the Item Editing tab for access to more editing features.)

For more information, see **Entering Metadata** and **Using the Item Editing Tab**.

Note: We do not enter metadata in the **Description** field because it has been designated as the full text search field and contains the extracted text.

Step 6: Upload the PDF file

After you have added the PDF file to your project, you can upload it to the pending queue on the server, where it can be reviewed and approved by an administrator before being added to your collection.

For more information about how to upload a file, see **Uploading Items from the Project Tab**.

Step 7: Approve the PDF file

After your file is uploaded to the pending queue, it must be approved by your CONTENTdm administrator in order to be added and viewable in the collection.

When it is approved, the PDF file is added to the collection on the CONTENTdm Server. The first 128,000 characters are extracted from the PDF and put into the Description field, which we designated as the full text search field in step 3. (If the text in the PDF is longer than that, the text is truncated. If the full text search field already contained data, the text would not be extracted.)

Step 8: Index the collection

Finally, after the file is approved, your CONTENTdm administrator must index the collection. This builds the text index that makes the PDF file discoverable via search.

The PDF file is now searchable and viewable inline in the CONTENTdm Item Viewer.

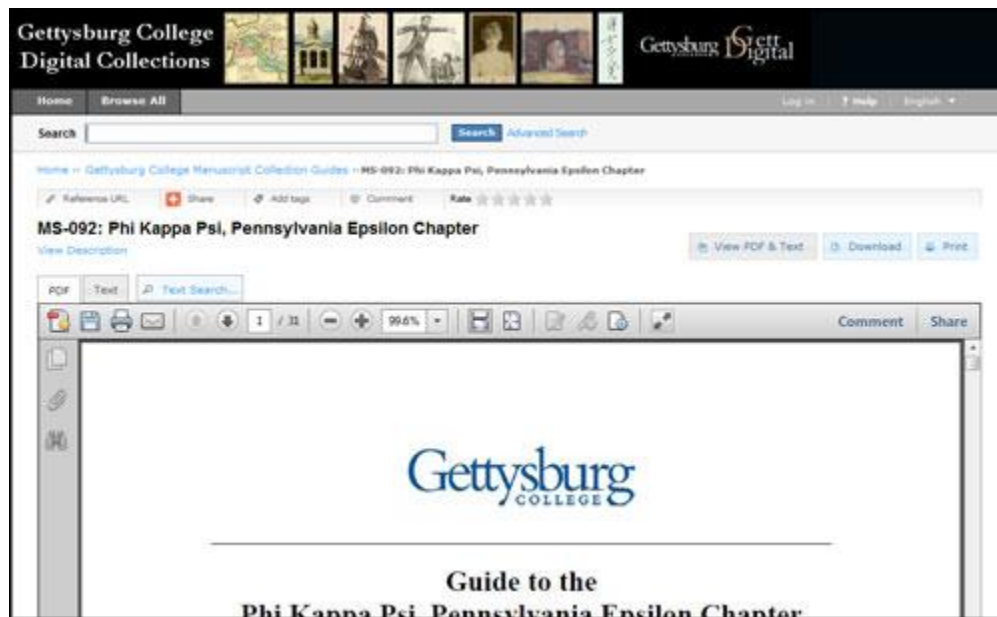


Figure 11: Multiple-page PDF file displayed as a single-item PDF inline in the Item Viewer

Importing a Multiple-Page PDF File as a Compound Object

You can import multiple-page PDF files to display in your collections in the Compound Object Viewer. In this example we will import the same multiple-page PDF file as a PDF compound object.

Step 1: Confirm that Adobe Reader is installed

Because we want to be able to view the PDF in the Project Client, we first confirm that Adobe Reader is installed on the Project Client computer. (To download and install the latest version of Adobe Reader, see <http://get.adobe.com/reader/>.)

Step 2: Confirm source file type and security settings

Confirm that the file name for the item you want to add has the extension .pdf. In this case, the source document is named Census92.pdf. (If the file extension is .doc or another format, you will need to first convert to PDF by using software such as Adobe Acrobat®.)

Next confirm that the PDF file is not password-protected, encrypted or locked. (If it is encrypted or locked, a thumbnail may not be generated, it may not be converted to a compound object, and text is not extracted in CONTENTdm.) One way to test for security settings is to try opening the document on your computer using Adobe Reader. You may see a message similar to the following. For more information, see the **Troubleshooting** section of this tutorial.



Figure 12: Adobe Reader security warning

If you need to modify the security settings, you may need permission to change the settings from the document owner, and then use Adobe Acrobat to edit and save the PDF file.

Step 3: Confirm collection settings

Because we have decided to add this multiple-page PDF file as a compound object, we confirm with the CONTENTdm administrator that the collection has PDF conversion enabled for multiple-page PDF files in CONTENTdm Administration. (This setting is displayed in CONTENTdm Administration under the Collections tab on the Collection Configuration page.)

Note: See also the following Confirm Project Setting step because some settings can be set locally in the Project Client to override the server settings.

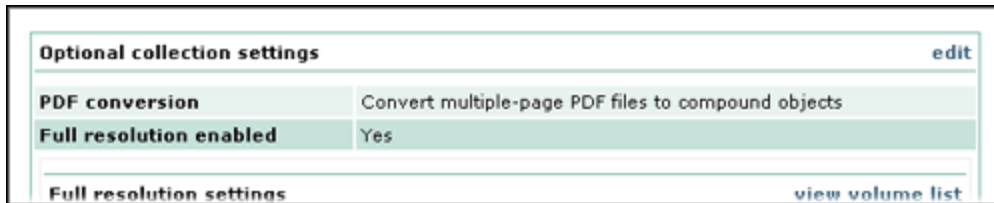


Figure 13: PDF conversion setting enabled for the collection in CONTENTdm Administration

Each page of this multiple-page PDF file will have a metadata record. When the PDF compound object is added to the Project Spreadsheet, the text from each page (up to 128,000 characters) is extracted from the PDF and put into the full text search field for the metadata record for each page.

Therefore, as with the single-item PDF file, we need to confirm that one metadata field has been set to the type Full Text Search.

For this example, we have set the Description field for the collection to the data type Full Text Search in CONTENTdm Administration.

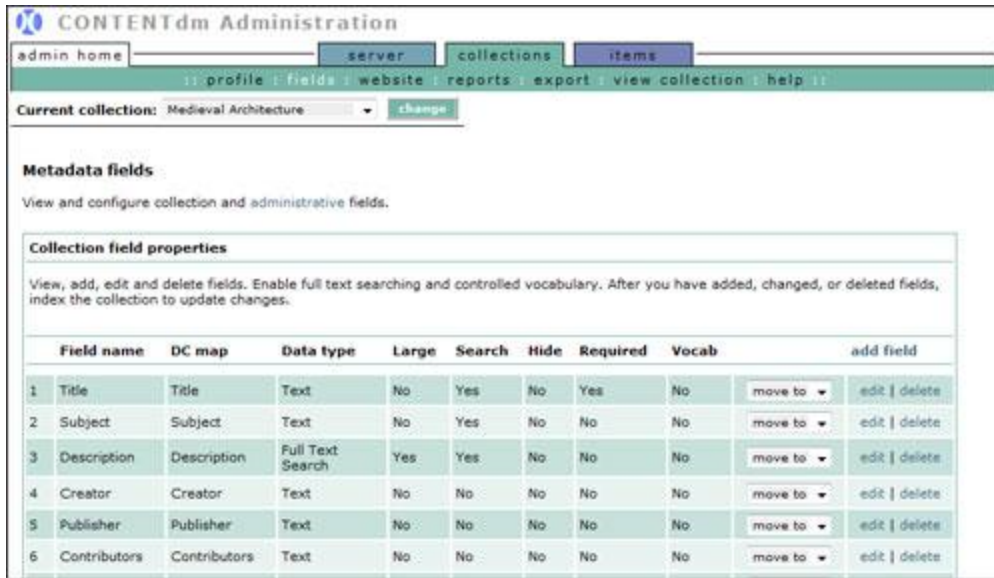


Figure 14: Metadata Fields page in CONTENTdm Administration

Step 4: Confirm project settings

Although we confirmed settings on the CONTENTdm Server, settings can be set for the project in the Project Client. These settings can override the server settings.

1. **Access the Project Settings Manager**, and click the **Processing** tab. The Processing page displays.

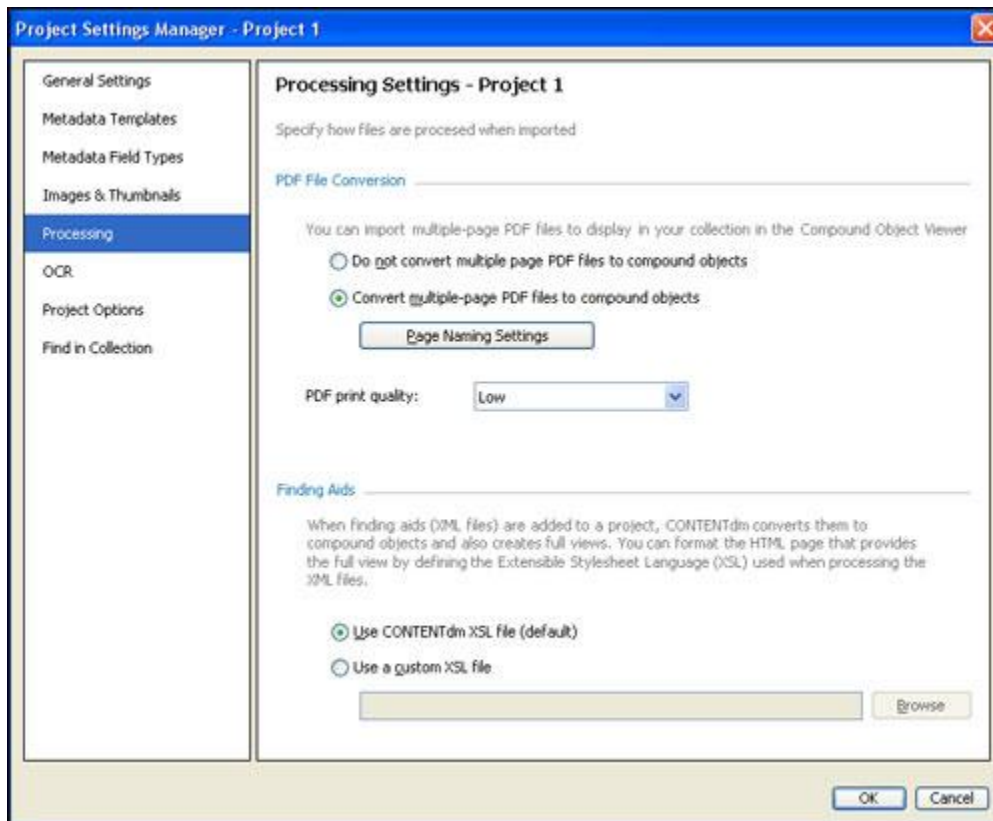


Figure 15: Project Settings Manager, Processing page

2. Confirm that PDF File Conversion is set to **Convert multiple-page PDF files to compound objects**.
3. Check the PDF Print Quality setting.

This setting enables you to specify the quality of the optional print versions (PDF) of the compound object. The quality is determined by this setting in relation to the quality of the original source material. Select a higher-quality setting for items where the PDF quality is important to your end users, or select lower-quality settings to improve processing and rendering times of the user experience. (For more information, see **Processing**.)

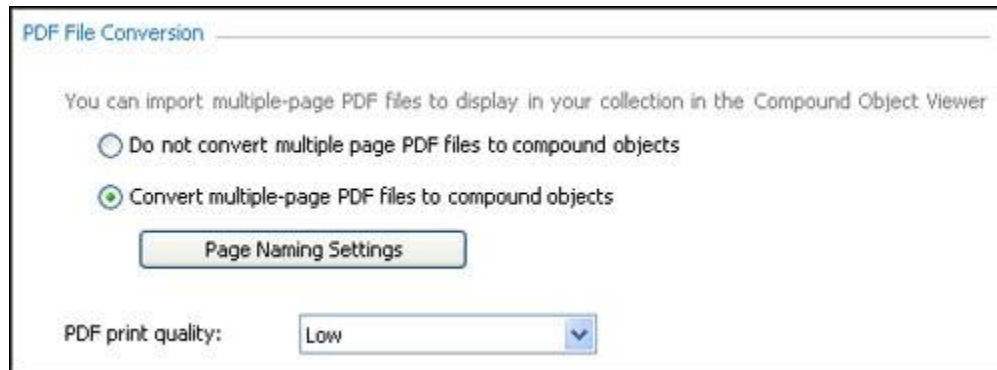


Figure 16: Convert multiple-page PDF files to compound objects

4. Click **OK** to save any changes.

Step 5: Add the PDF file as a compound object

The general steps for adding a multiple-page PDF file to your project are the same as adding a single-item PDF file or other individual files.

For more information, review the steps outlined for **adding a single-item PDF file** in the example above.

After the item is added to the Project Spreadsheet, you can edit the title for the compound object-level record in the required **Title** field and enter other metadata. (Or open the item in the Item Editing tab for access to more compound object editing features.)

For more information, see **Entering Metadata** and **Using the Item Editing Tab**.

Note: We do not enter metadata in the **Description** field because it has been designated as the full text search field and contains the extracted text.

Step 6: Upload the PDF file

After you have added the PDF file to your project and entered metadata, you can upload it to the pending queue on the server, where it can be reviewed and approved by an administrator before being added to your collection.

For more information about how to upload a file, see **Uploading Items from the Project Tab**.

Step 7: Approve the PDF file

After your file is uploaded to the pending queue, it must be approved and indexed by your CONTENTdm administrator to be viewable in the collection.

Step 8: Index the collection

Finally, after the file is approved, your CONTENTdm administrator must index the collection. This builds the text index that makes the PDF compound object discoverable via search.

The PDF compound object and all of its pages are now searchable and viewable inline in the CONTENTdm Compound Object Viewer.

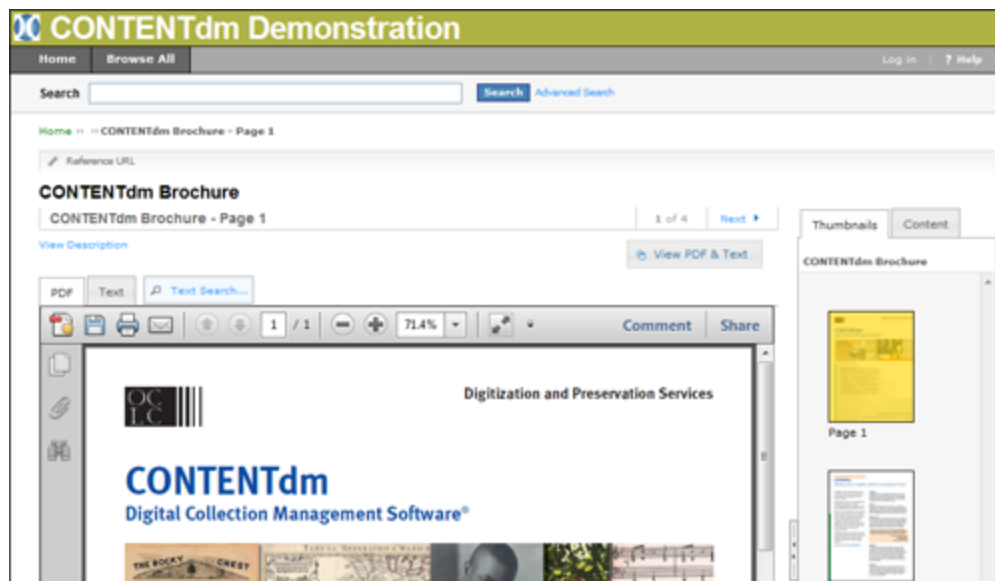


Figure 17: Multiple-page PDF file displayed as a compound object

Section IV Editing PDF Files

After your PDF files have been uploaded, approved, and indexed to a collection, the metadata for single-item PDF files and PDF compound objects can be edited by using the Web interface or the Project Client, similar to how you edit other items and objects but with some considerations for how they are structured and processed.

- PDF compound objects have editable metadata records for each page, as well as for the compound object itself.
- You cannot replace or remove individual pages of PDF compound objects using the Project Client or the Web interface. To replace individual pages of PDF compound objects, you must delete the compound object from the collection, replace the pages in the original PDF file, and import the file again.
- You cannot delete individual pages of a PDF compound object.
- Permission settings for PDF compound objects apply to all pages of the compound object. You cannot set permissions for individual pages of PDF compound objects.

Editing a PDF Compound Object Using the Project Client

You can edit the metadata, display page names, thumbnails and permissions of a PDF compound object using the Project Client. (For more information about using the Project Client for editing, see **Editing Items in a Collection** in the Help files.)

Note: When you download items from a collection into your project, the items are **locked** from editing by other users but are still available for viewing in the Web interface.

Step 1: Find the compound object and page-level records & add to the Project Spreadsheet for editing

1. Open the Project Client project that is associated with the collection you wish to edit.
2. Click the **Find in Collection** tab.
3. Browse for the items in the collection by clicking **Browse Collection** or conduct a search by typing keywords in the search box and clicking **Search**. (Or construct a more detailed search by clicking **My Advanced Search**.)

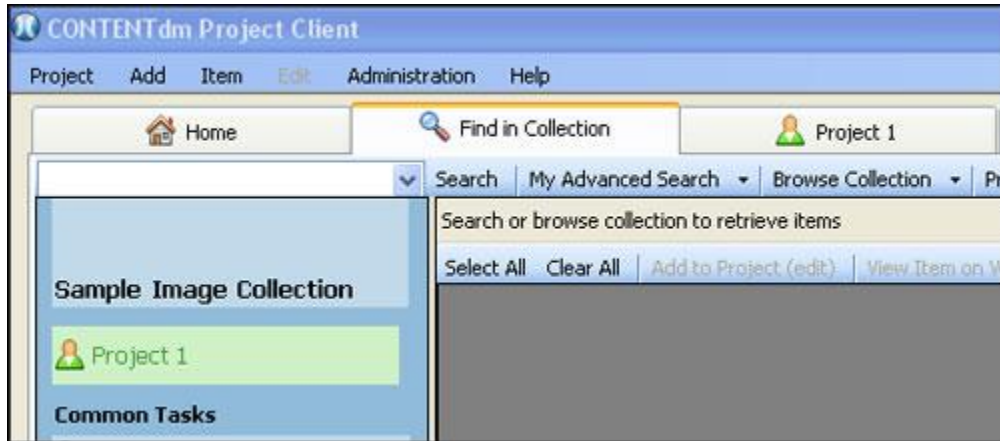


Figure 18: Finding items in the collection

For more information about using the advanced search or the other features in the Find in Collection tab, see **Using the Find in Collection Tab**.

4. Search and browse results are displayed.

Thumbnail	Title	Subject	Description	Creator	Publisher
	Allen Library.	University of Wa...	A long shot of th...	Carol Martin	University of Wa...
	Allen Library Arch.	University of Wa...	A long shot of th...	Jim Barrett	University of Wa...
	Allen Library Arc...	University of Wa...	A medium shot o...	Jim Barrett	University of Wa...
	Suzzallo and Alle...	University of Wa...	A wide shot of S...	Jim Barrett	University of Wa...

Figure 19: Search results in Find in Collection tab

5. Locate the PDF compound object to edit, and select the checkbox. Then click **Add to Project (Edit)**.

The selected items are added to the Project Spreadsheet, a locked icon displays and the item rows change color on the Find in Collection tab.

Note: Although you can add more than 10,000 items at a time, we recommend downloading less than 10,000 items at a time from the server to maintain acceptable download times.

Step 2: Edit the records in the Project Spreadsheet

Using the Project Spreadsheet and the Item Editing Tab, we can edit metadata for the compound object and the individual pages of the compound object.

Note: If you download and then delete an item from your project (whether or not you make any changes), the item is unlocked and returns to its original state within the collection.

In this example, we open a PDF compound object in the Item Editing tab and edit the contents of the Subject field for several pages. For more information about all of the compound object-editing features provided by the Item Editing Tab, see **Using the Item Editing Tab**.

1. From the Project tab Project Spreadsheet, double-click the thumbnail. The PDF compound object opens in a new tab, with the compound object-level record displayed.

Note: Use the **View Structure** drop-down menu to change the view of the compound object navigation pane. This enables you to switch between a structural, thumbnail or spreadsheet view of the PDF file. (The spreadsheet view provides **fill functions** and the other editing options that are available when using the project spreadsheet.)

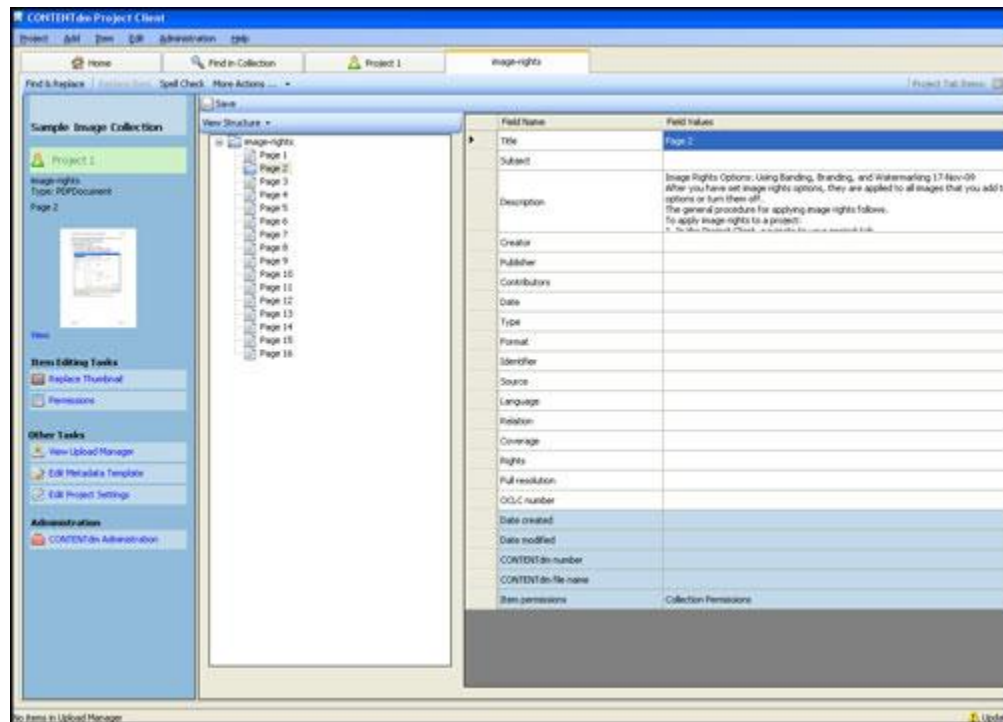


Figure 20: Item Editing Tab

- After you enter metadata in the Subject field for the compound object-level record, use the navigation pane to click through the pages of the object and enter metadata for the Subject field of those records. (Alternatively, if all Subject fields will have the same metadata, you can switch to the spreadsheet view and use the **Fill Compound Object** option. This option fills the same data in the selected field for all pages of the compound object. For more information, see **Using Fill Functions**.)

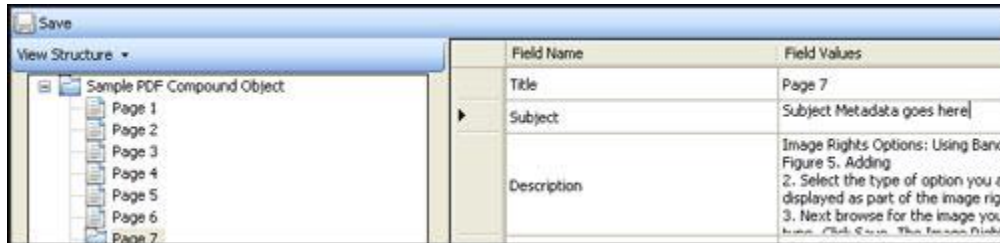


Figure 21: Editing metadata of a page of the compound object

- Click **Save** to save your changes.

Step 3: Upload the items for approval

When you are finished editing, select the item in the Project Spreadsheet and then click **Upload for Approval**. Items are sent to the pending queue for approval.

Editing a PDF Compound Object Using the Web Interface

To edit items using the Web interface, you must have CONTENTdm administrator rights to edit and approve metadata. For more information about using the Web interface for editing, see **Editing Items** in the Help files.

Step 1: Find the compound object and page-level records

Editing a compound object begins with searching for the compound object and its pages. (You can browse the entire collection, however, to edit pages within a compound object, you must search for the pages, because subpages are suppressed while you are browsing.)

- In CONTENTdm Administration, go to the **Items** tab and then click **Edit**. In this example, we then search for the term "arizona."

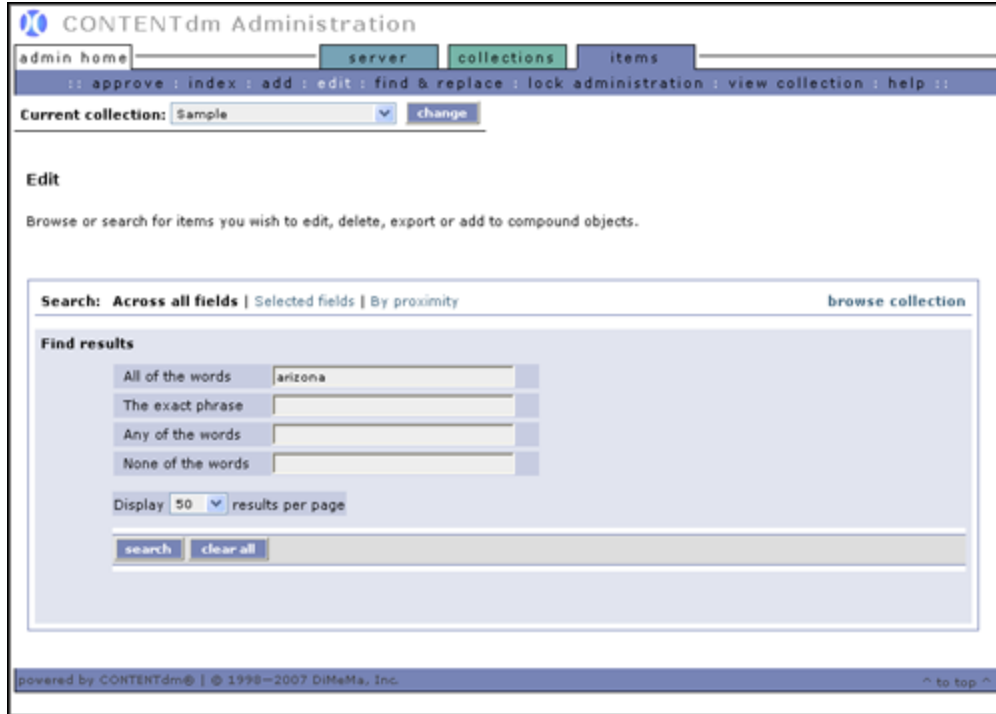


Figure 22: Searching for items to edit using CONTENTdm Administration

2. In this example, search results display 15 items (the PDF compound object and the individual pages that make up that compound object).

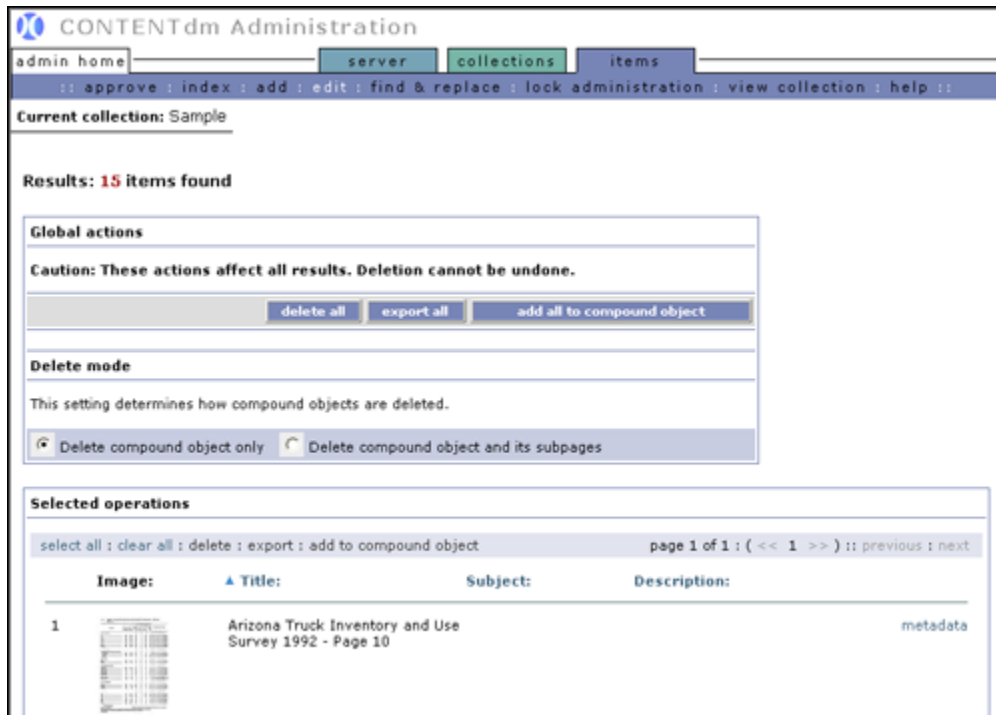


Figure 23: Search results

Note: PDF compound objects do not display a structure option. You cannot edit or delete individual pages of PDF compound objects by using CONTENTdm Administration. To edit or delete individual pages of PDF compound objects, you must

delete the compound object from the collection, update the structure, edit, or remove the pages in the original PDF file, and then import the file again.

Step 2: Edit the metadata

From the Results page, click the Metadata link associated with each page or the compound object record to edit, add, and delete metadata. You also can apply permissions.

Note: Permission settings for PDF compound objects apply to all pages of the compound object. You cannot set permissions for individual pages of PDF compound objects.

The screenshot shows the 'CONTENTdm Administration' interface. At the top, there are navigation tabs: 'admin home', 'server', 'collections', and 'items'. Below these are links: 'approve', 'index', 'add', 'edit', 'find & replace', 'lock administration', 'view collection', and 'help'. The 'Current collection' is 'Sample'. The main section is titled 'Edit item'. Below this, there is a message: 'Edit metadata then click **save changes**. To remove this item from the collection, click **delete**. After making changes, you must rebuild the text index.' Below the message is a form with the following fields:

- Title: Arizona Truck Inventory and Use Survey 1992
- Subject: Department of Commerce Report
- Description: JOBNAME: No Job Name PAGE: 1 SESS: 42 OUTPUT: Thu Apr 28 15: 41: 59 1994 / pssw02/ disk2/ economic/ tc92t/ 3/ 07.txtint
Census of transportation
- Creator: Department of Commerce
- Publisher: (empty)
- Media: (empty)
- Material: (empty)
- Measurements: (empty)
- Collection: (empty)
- Full text: (empty)

Figure 24: Editing item metadata in CONTENTdm Administration

Click **Save Changes** to save edits or deletions. A confirmation page displays. Items do not have to be approved, but the collection must be indexed to update changes.

The screenshot shows the 'CONTENTdm Administration' interface. At the top, there are navigation tabs: 'admin home', 'server', 'collections', and 'items'. Below these are links: 'approve', 'index', 'add', 'edit', 'find & replace', 'lock administration', 'view collection', and 'help'. The 'Current collection' is 'Sample'. The main section is titled 'Edit item'. Below this, there is a message box containing the text: 'The item has been updated. Index the collection to activate the changes.' At the bottom of the page, there is a footer: 'powered by CONTENTdm® | © 1998–2007 DiMeMa, Inc. ^ to top ^'.

Figure 25: Edit Item confirmation page

Step 3: Index the collection

You can build a collection text index if you have rights to approve items and build indexes. For more information, see **Building the Text Index** in the Help files.

After the collection is indexed, the edited metadata is searchable and viewable in the collection.

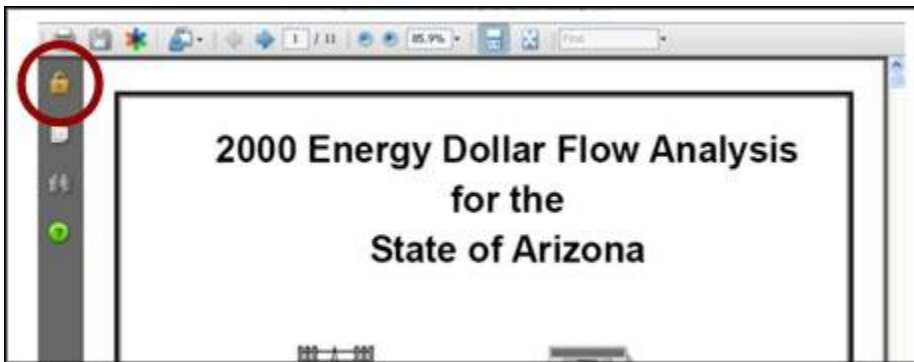
Section V Troubleshooting

This section provides suggestions for troubleshooting issues with using PDF files in CONTENTdm.

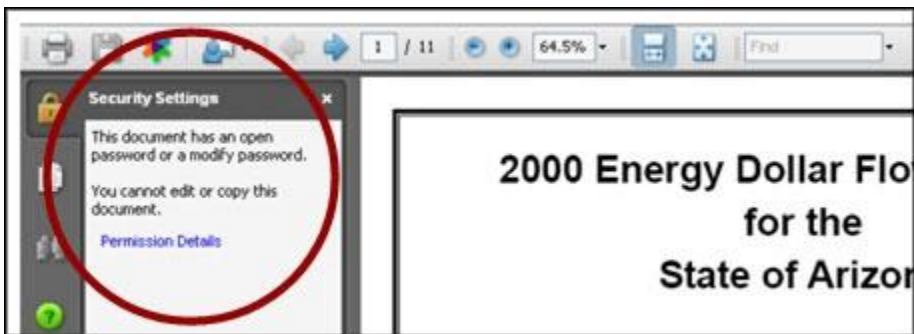
Multiple-page PDF file does not display as a compound object and collection setting for PDF conversion was enabled when the file was imported.

If you try to import a multiple-page PDF file and the collection is configured correctly to enable PDF conversion, the PDF file may have security settings that prevent processing.

PDF files can be encrypted, locked, or password-protected in many ways. You may still be able to open them without noticing security settings. One way to check whether your PDF file has security settings is to look for the lock icon in Adobe Reader, as shown in the following (this may vary depending on your version of Adobe Reader).



To learn more about the specific settings, click the lock symbol to display the Security Settings pane.



To see all of the permission information, click **Permission Details** to display the Document Properties dialog box.

To convert the PDF file to a compound object, you will need to disable the security settings.

Remove the file with security settings from the collection, and then check whether you need permission to change the settings from the document owner. Use Adobe Acrobat to edit the security settings and save the PDF file. Then **import the multiple-page PDF file as a compound object.**