Tab-Delimited File
and
Compound Object – Monographs

(Not Postcards, Documents, or Cubes)

See help sheet: Tab-Delimited File and Compound Objects– Postcards, Documents, or Cubes)

Content

Content

Why use Tab-delimited Files in CONTENTdm Project Client to upload items?

•Tab-delimited files are an easy way to upload many single objects, a single compound object, or multiple compound objects along with their associated metadata.

•Tab-delimited files leave the creator of the metadata with a record of the metadata in case the record is needed again.

•Multiple creators of metadata can work in the same file to create the metadata record.

How to create a Tab-delimited File:

- Use Excel to create the file – this is the working file.

- All cells must be formatted for **text**.

- The Excel file must have specific information in specific columns and rows for single items and compound objects (documents, post cards and cubes – one format and monographs – 2nd format).

- Once all the metadata has been entered, the file is saved as an Excel file and then as a Tab-delimited file (txt).

- Best practice is to always work in the Excel file and then save the file as a Tab-delimited file (txt).
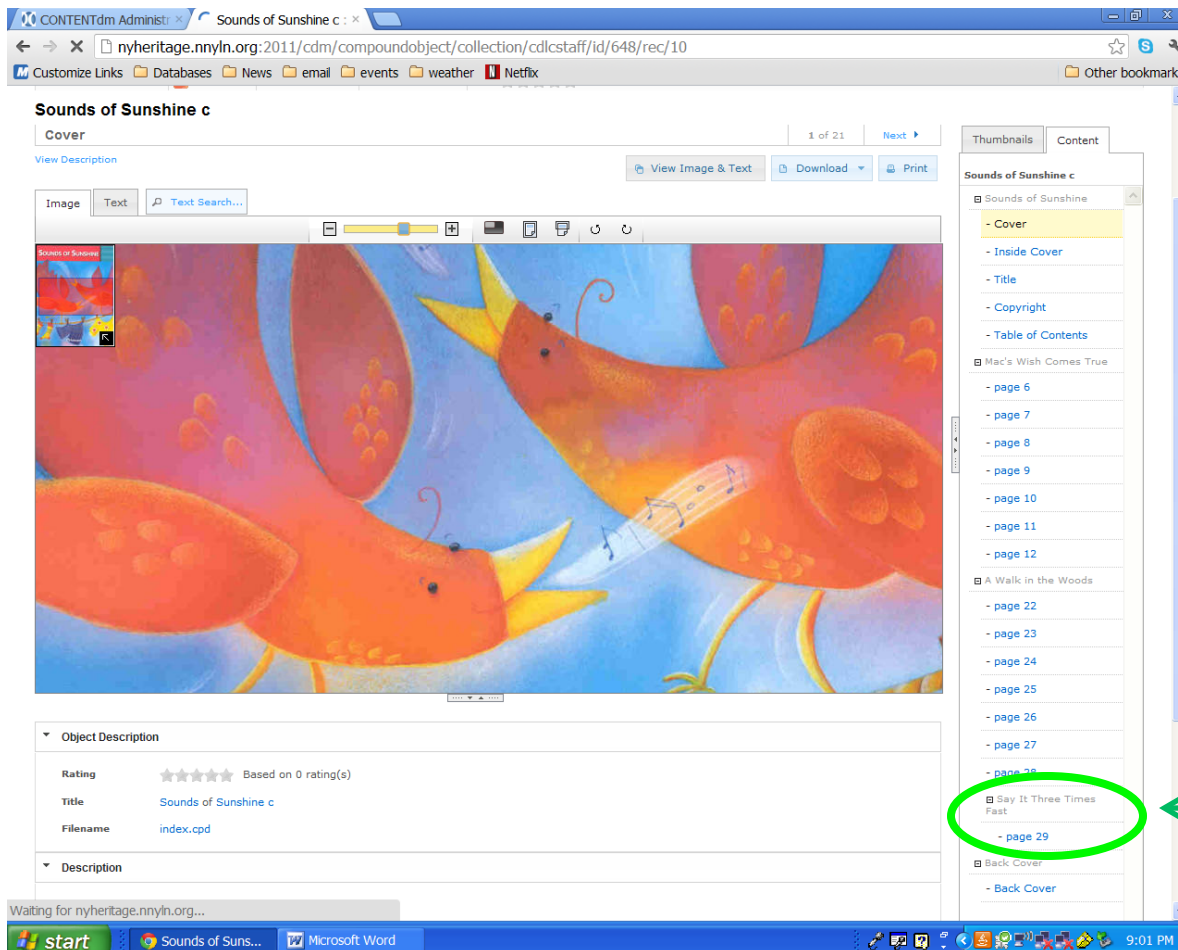
Work in Excel to create the Tab-delimited File:

•**One Excel column must contain the file name of the images being imported and mapped to Object File Name** - Best practice, the field that contains the file names of the imported files is placed in the last field column used in the Excel file.

•All **object metadata** field names must appear in the first record of the file.

•The file names must exactly match the contents of the upload directory for the files. One single mismatch causes the entire upload to fail.

•At least one field must map to the **Title** field.

•Remember for New York Heritage collections the **Collection ID** metadata field must be filled in for upload to the CONTENTdm server. This content can be added in the Tab-delimited file or in the Project Client.

- Use a tab for the delimiting character.

- End each record with a carriage return.

- Do not use carriage returns or tabs within a field.

- The end of the document should be the very end of the last word in the last line, with no extra blank lines or spaces.

- Remove special characters from file names and collection metadata field names. Special characters are: \ / : * ? " < > |

- When entering file names, be sure to include the extension, even if your operating system hides the extension type. Common extension types are JPG, TIF, GIF, and PDF.

- When entering file names, enter only the file name of the item, such as *item.jpg*. Do not enter full path names such as *c:\windows\item.jpg*.

- When importing files, use unique file names for each item. Using capital letters in the file name does not make file names unique. The file name *item.jpg* is treated the same as *ITEM.jpg* and *ITEM.JPG*.

- **Store all of the items referred to in the Tab-delimited file in one folder or sub-directory.**

Example: Compound Object - Monograph

This monograph has 3 chapters and one sub-chapter in chapter 3.

Folder setup for a monograph
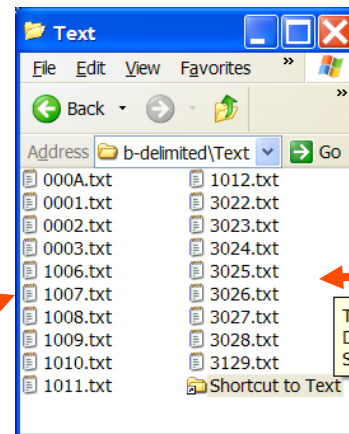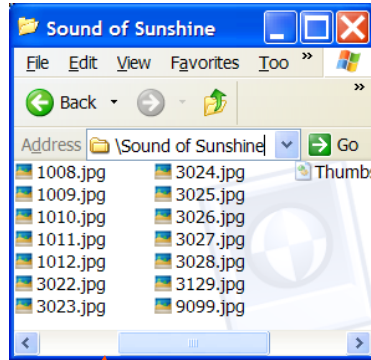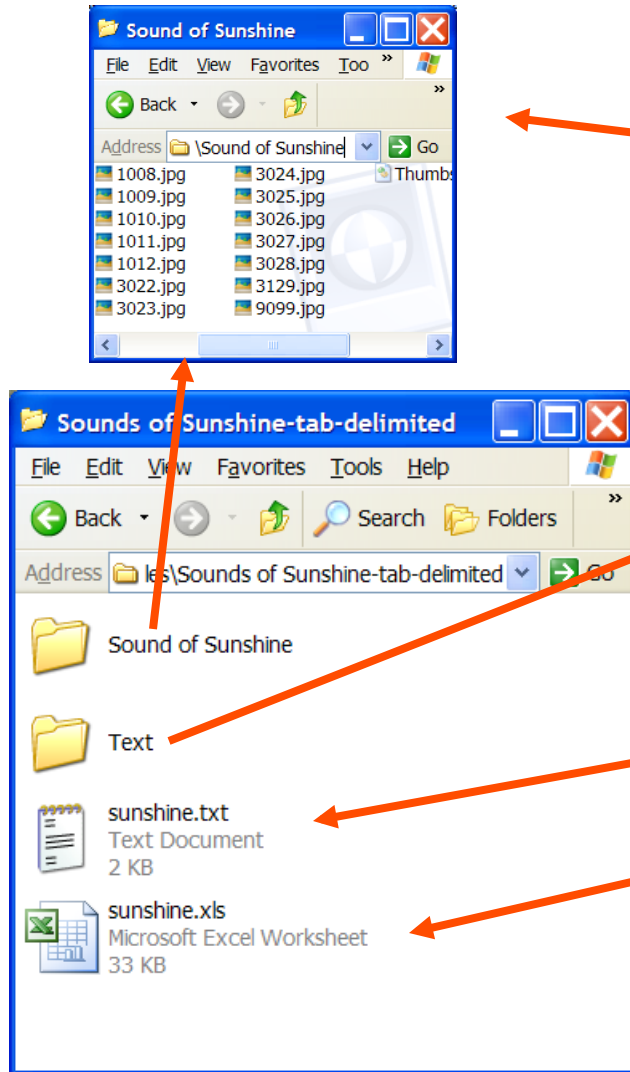
Folder contains:
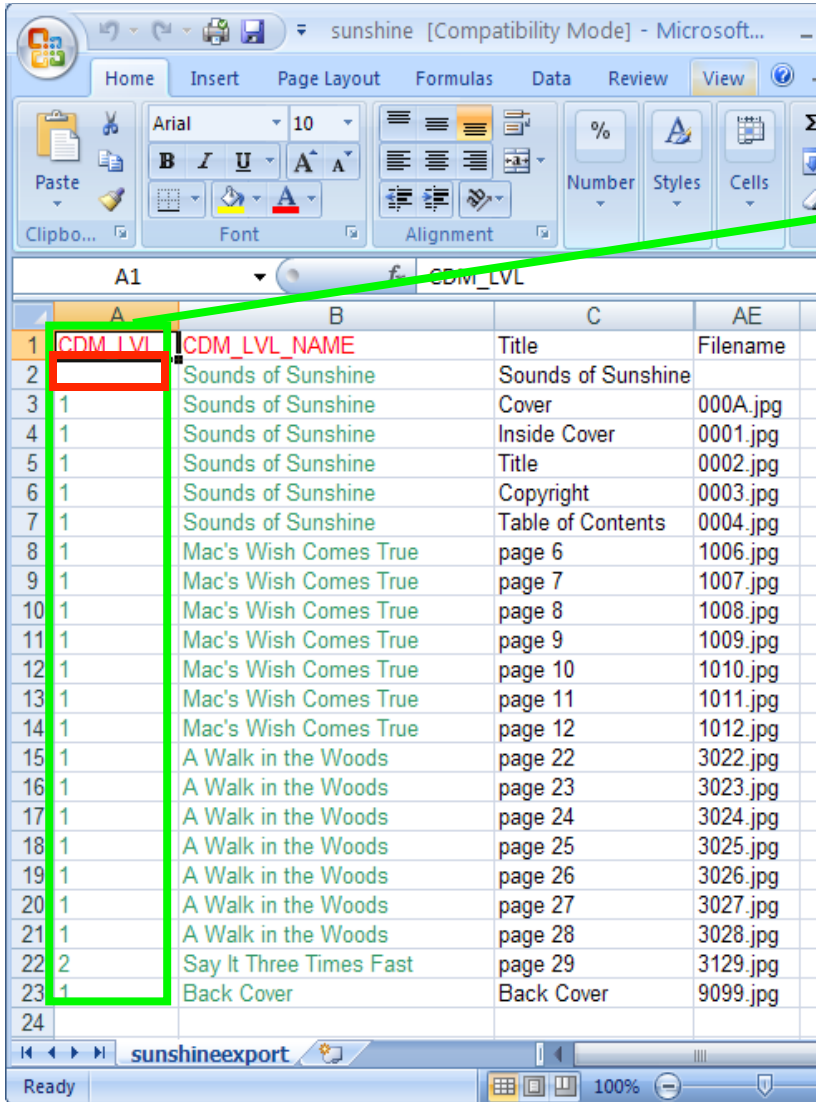
--Folder of images

--Folder of transcripts

--Tab-delimited file

--Excel file

Note: a folder containing the tiff files can also be in this folder.

Monographs – Excel file- Columns

Monograph: Needs two additional columns.

--The first column, labeled CDM_LVL in the first row.

This column indicates the **level** of the hierarchical structure in the monograph. CONTENTdm supports up to 9 levels in a monograph.

There is **NO** level of the hierarchical structure for the object metadata – thus **CDM_LVL** cell in row 2 is blank.

In this example: there are 2 levels of hierarchical structure – everything is level 1 except for the one sub-chapter.

--The second column is labeled: CDM_LVL_NAME in the first row.

This column contains the **title for each hierarchical structure of the monograph.** All records within the same level must contain the same CDM_LVL_NAME.

In this example: there are 4 different **Level 1** hierarchical structures– so four different titles. There is one **Level 2**, which has its own title too.

Monographs – Excel Files - Columns

The **third column** contains the **titles for the object and each page of the object**.

The **last column** contains the image file names (mapped to "Object File Name") that make up the object.

Remember, the object metadata row does not have an image file so it is **blank**.

Monograph – Excel file - Rows

The **first row** contains the
COM_LVL in **cell 1** and
CDM_LVL_NAME in **cell 2**
with the names of the metadata
fields for the metadata being
imported following.

All cells are formatted as text.
The first row must contain **Title**
and **File Name** fields.

Monograph – Excel file - Rows

The **second row** of the excel file contains the **object metadata**.

There is **NO** level of the hierarchical structure for the object metadata – thus **CDM_LVL** cell in row 2 is blank.

There is **NO** image file for the object metadata – thus the **File Name** cell in row 2 is blank.

The thumbnail will be created for the compound object by using the first image file in the **File Name** column.

Once the Excel file is created, completed, and saved as an Excel file, **save the file again as a Tab-delimited file** (txt).
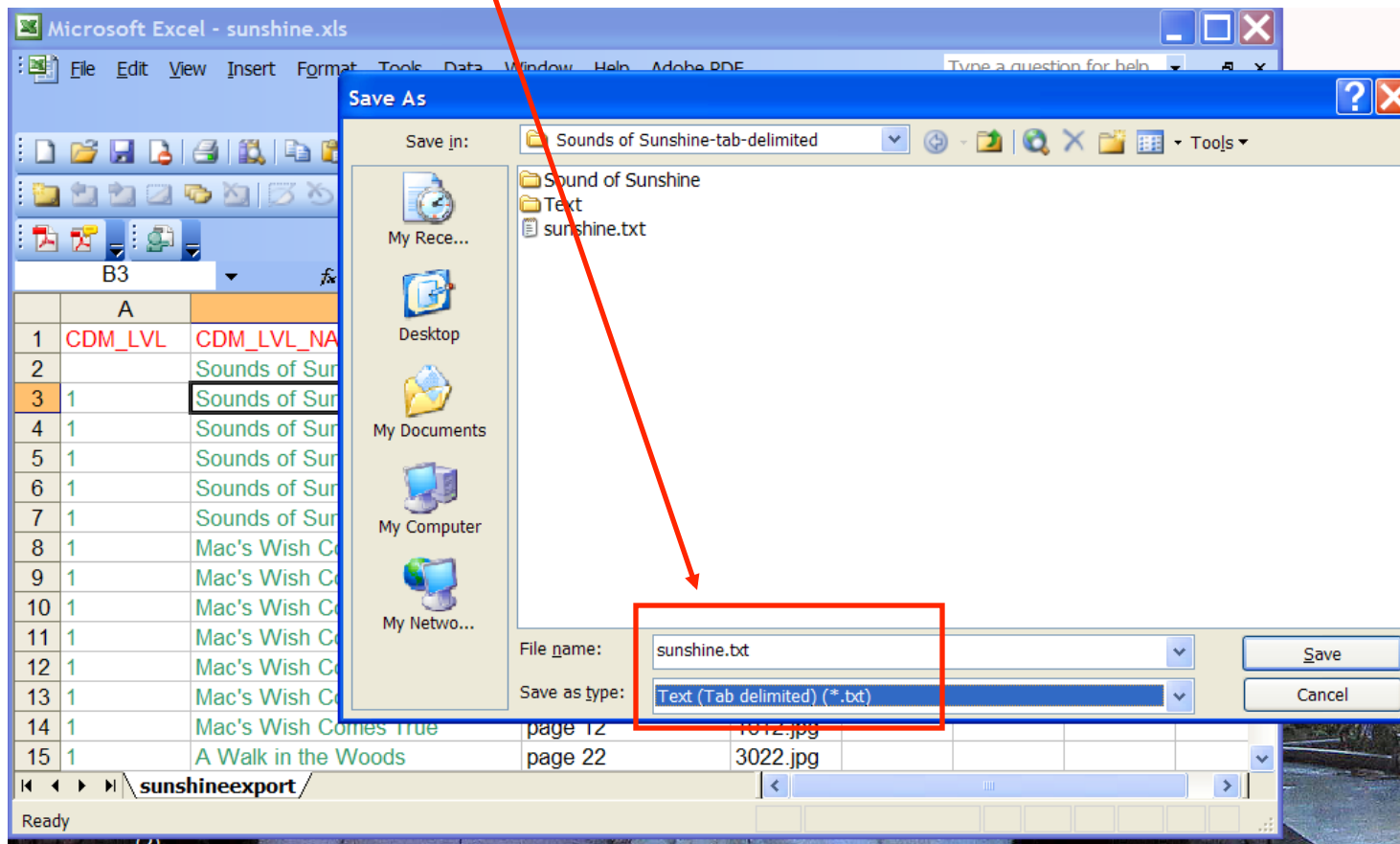
Note: Tab-delimited format does not support Excel workbooks that contain multiple worksheets. To save the active workbook sheet (the sheet you have open) <u>click</u> on **OK**.
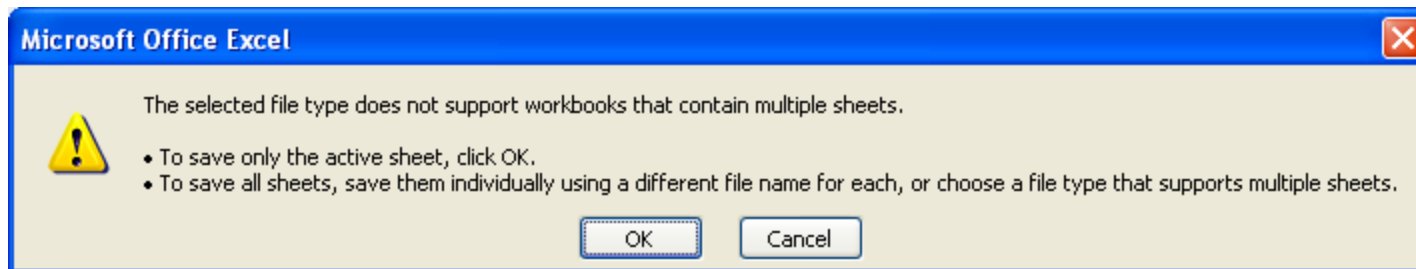

You want the worksheet in Tab-delimited format, so <u>click</u> on **Yes**.
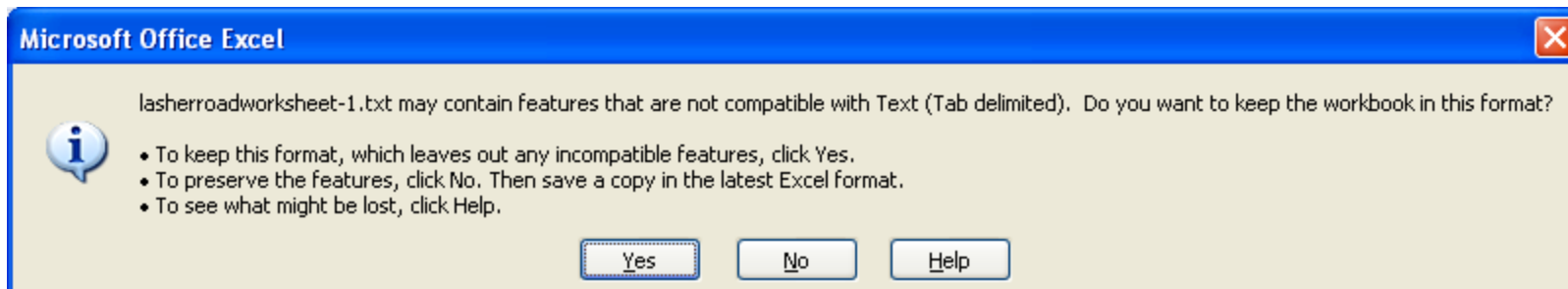
If changes need to be made in the Tab-delimited file, make the changes in the Excel file, save it, and then save it again as a Tab-delimited file. <u>**It is best not to work in the Tab-delimited file.**</u>

Import Tab-delimited file into CONTENTdm Project Client

- Open **CONTENTdm Project Client.**

- Open **Project** or create a **new Project.**

- The **Project** tab view opens.
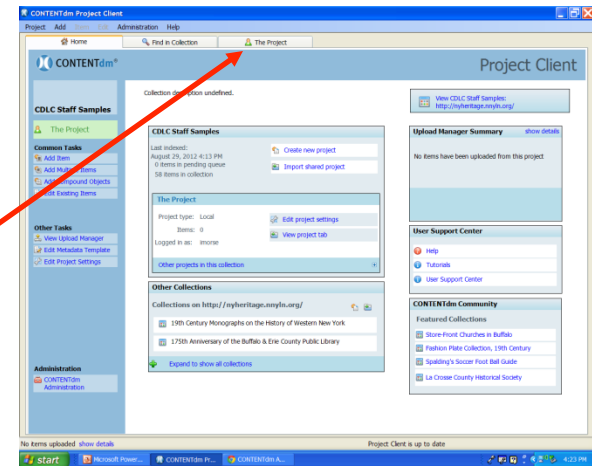
- The next 5 steps are **very important** because the **File Name** field in the tab-delimited file will be mapped to "**Object File Name**" not to the metadata field **File Name**. The items being imported into **CONTENTdm Project Client** will not be uploaded to the **CONTENTdm server** without the **File Name** metadata field information for New York Heritage collections.

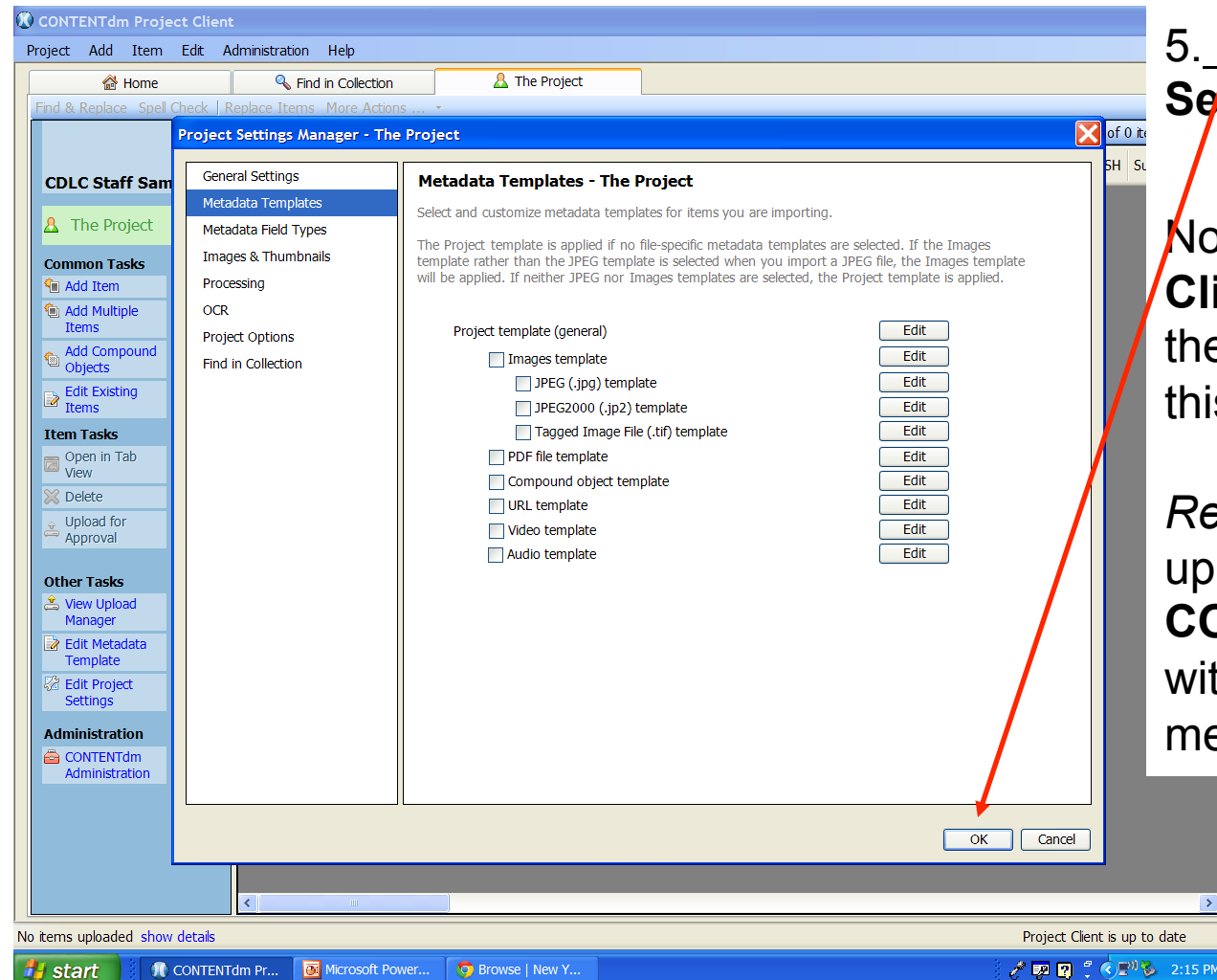To automatically collect the File Name metadata information follow the next 5 steps.
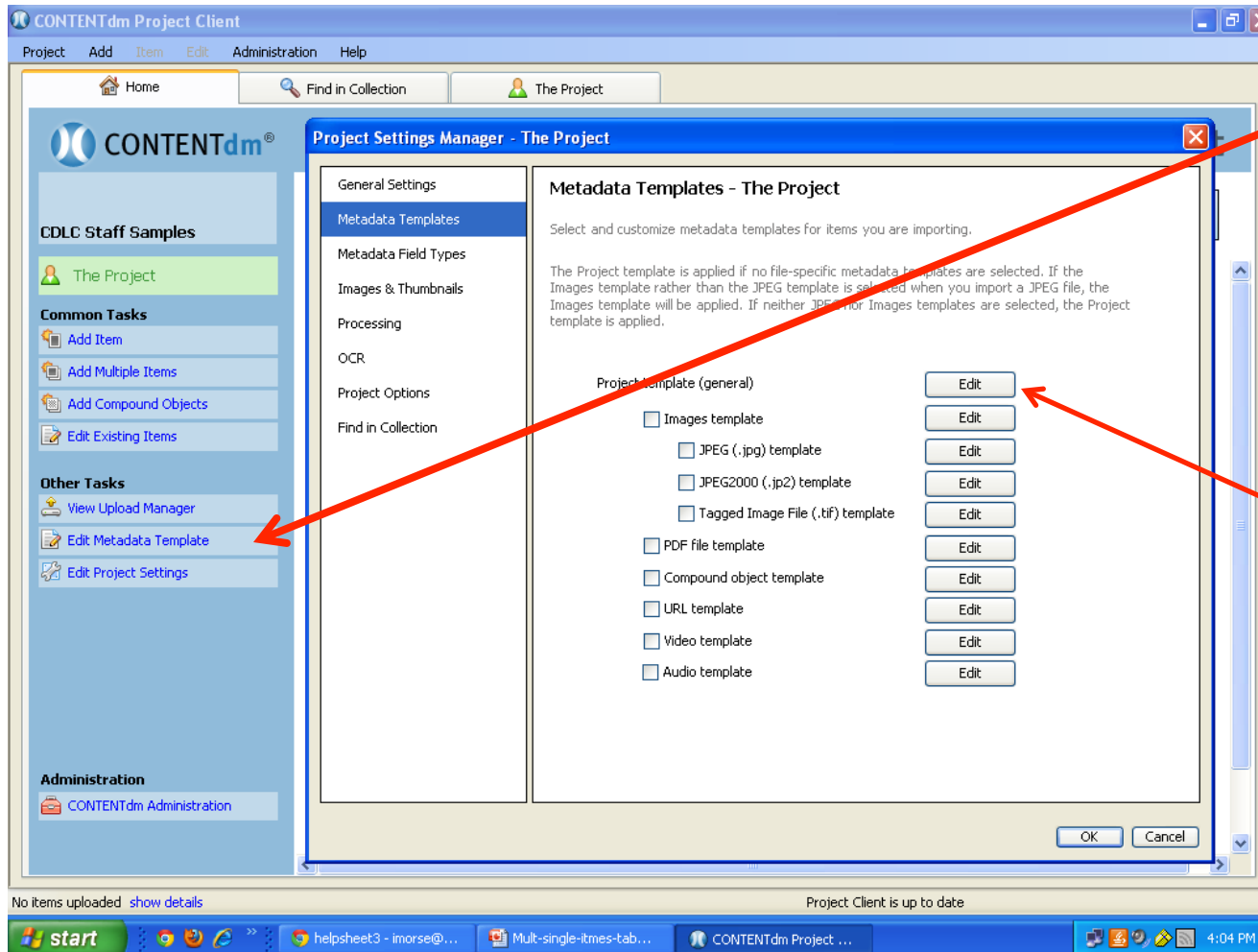
1.Click on **Edit Metadata Template** – the **Project Settings Manager** window opens.

2. Click on **Edit** for the **Project template (general)**
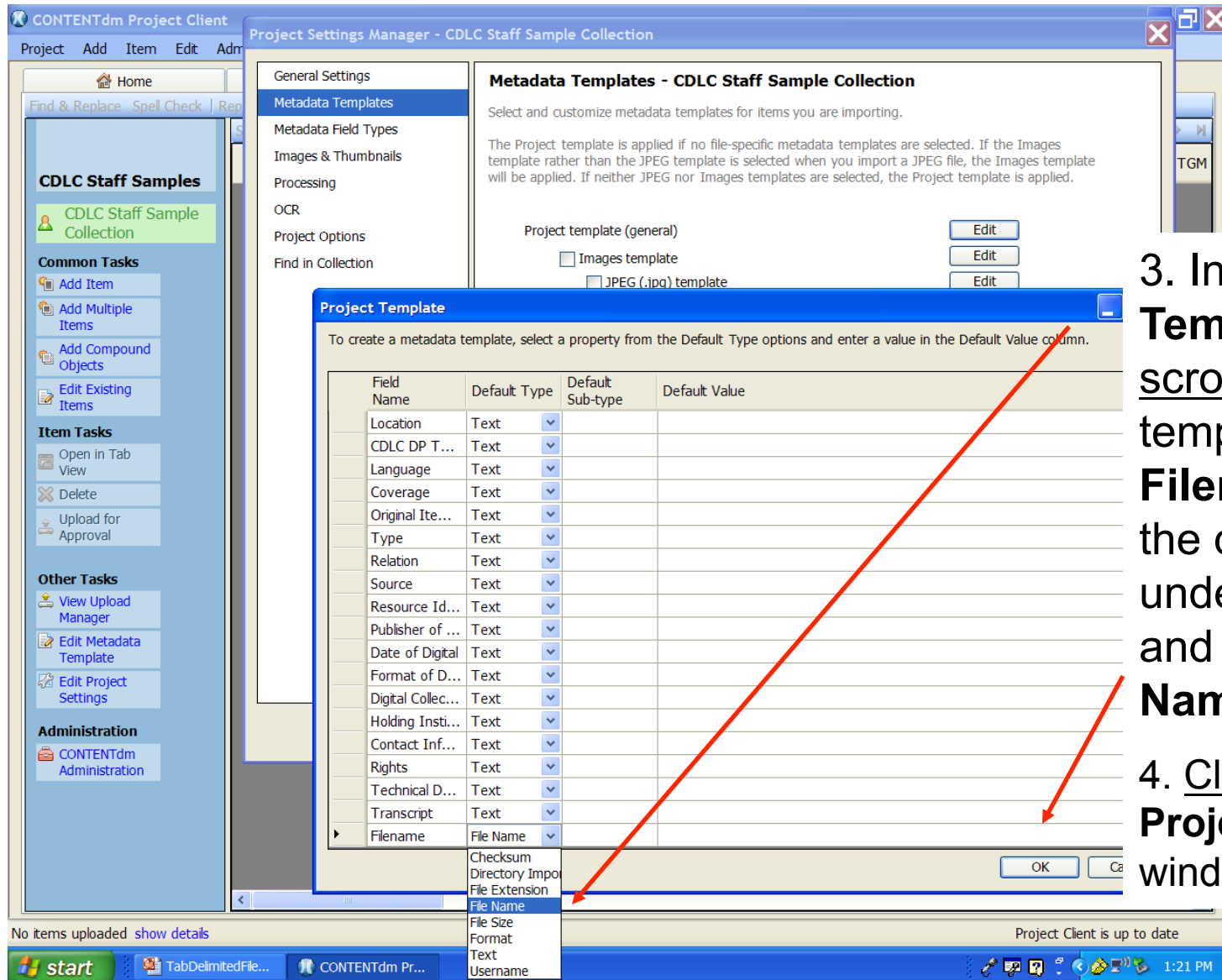
3. In the **Project Template** window, scroll down the template to **Filename**, click on the drop down box under **Default Type**, and highlight **File Name.**

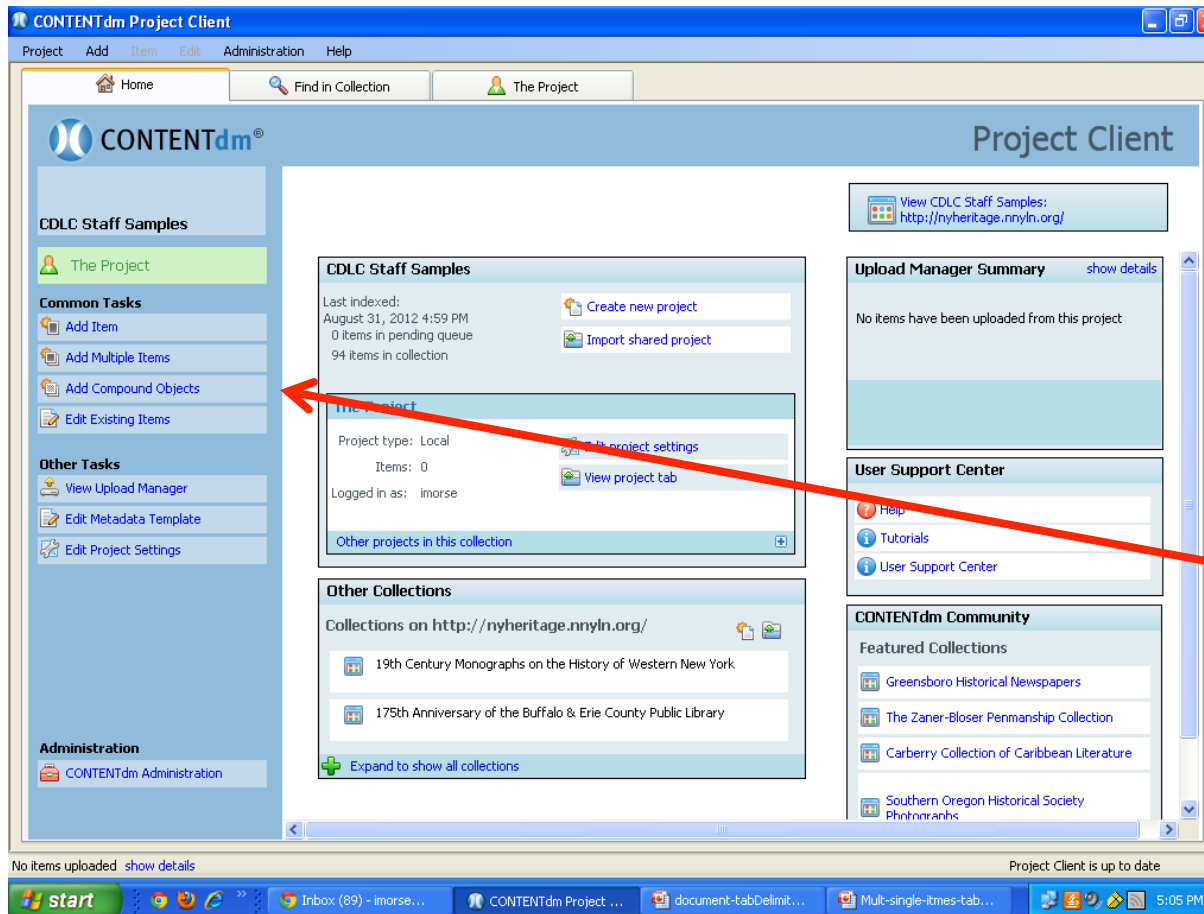4. Click on **Ok** in the **Project Template** window.

5. Click on **Ok** in the **Project Settings Manager** window.

Now, **CONTENTdm Project Client** software will collect the information needed for this field automatically.
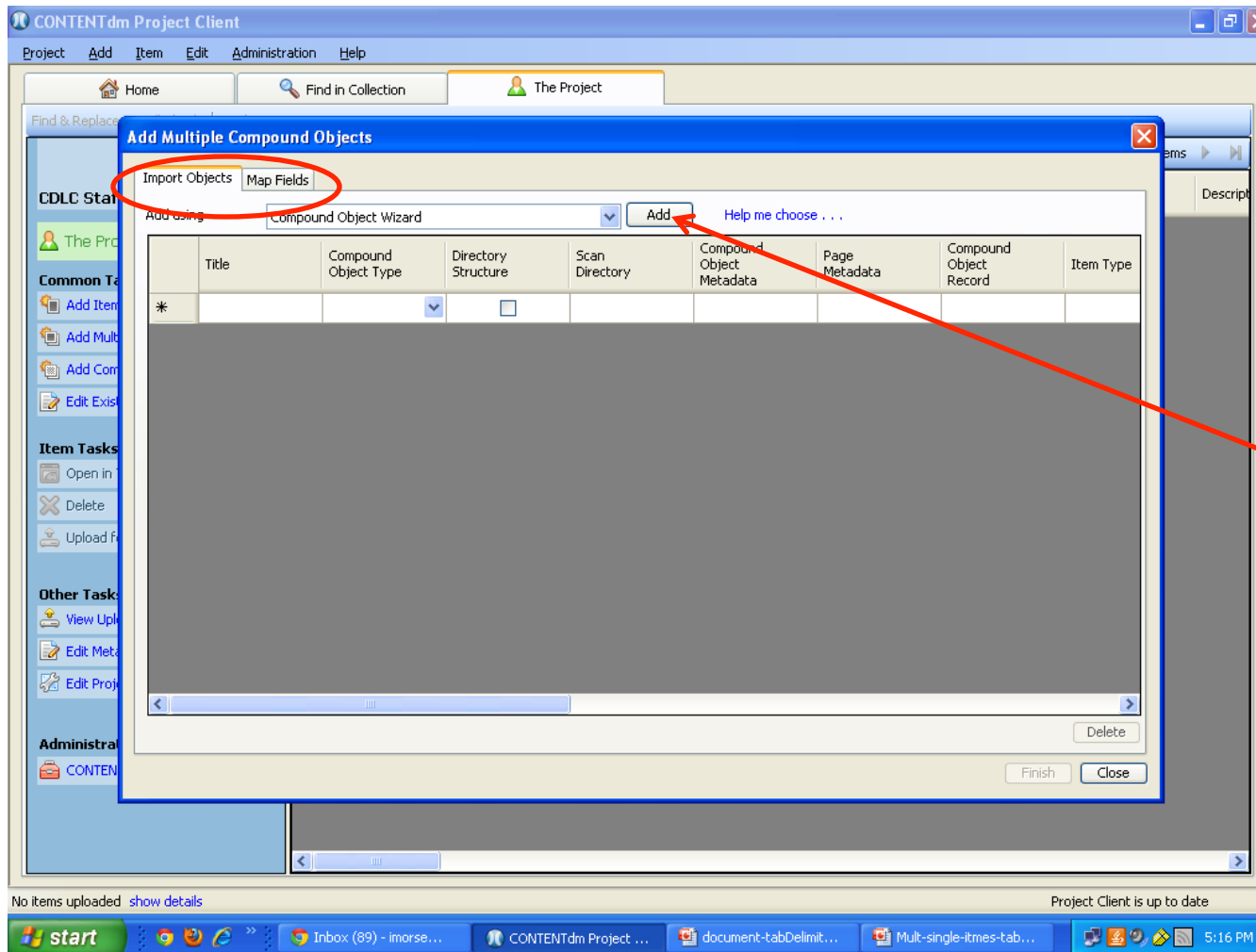
Remember, items will not be uploaded to the **CONTENTdm server** without the **File Name** metadata field information.
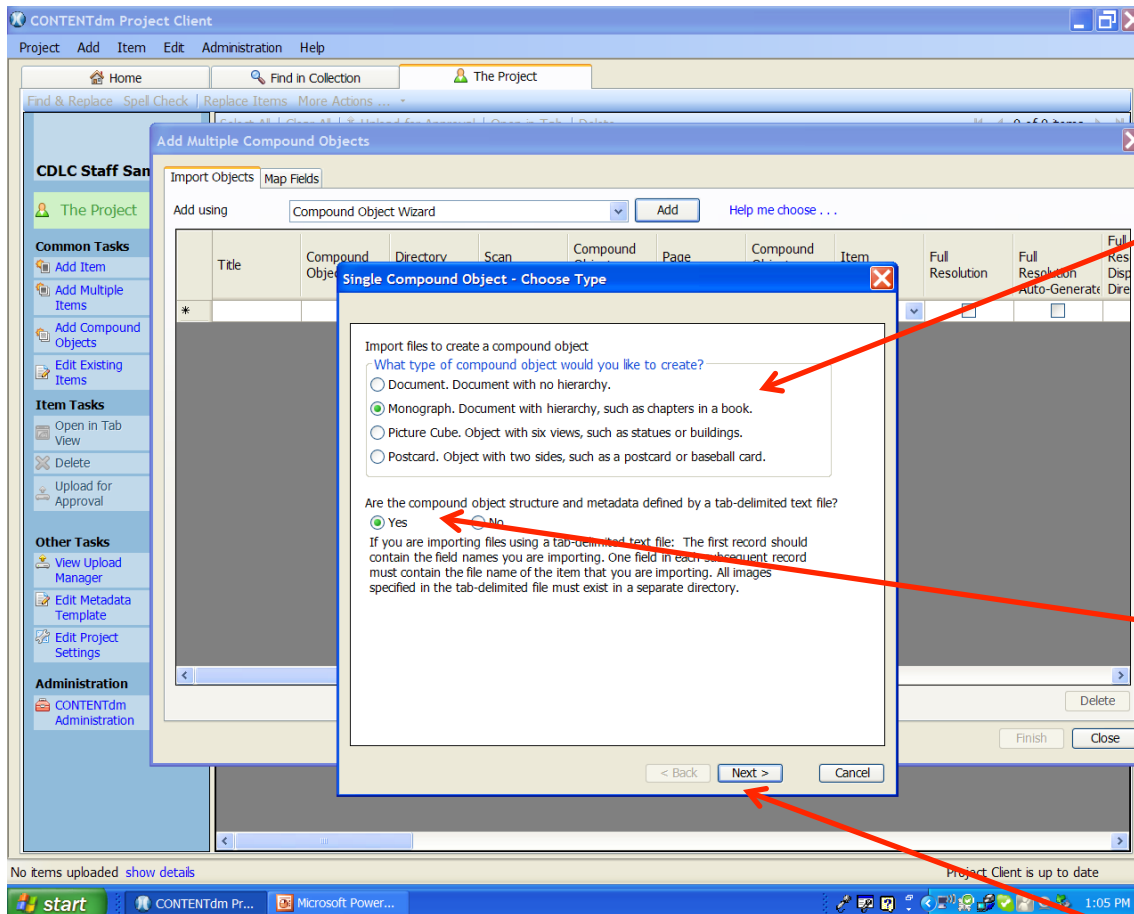
Import Files

You are now back in the **Project tab** view and ready to <ins>import the Tab-delimited file, the monograph, and the transcript files.</ins>

<ins>Click</ins> on **Add Compound Objects**.

The **Add Multiple Compound Objects** window opens in the **Import Objects tab** view.

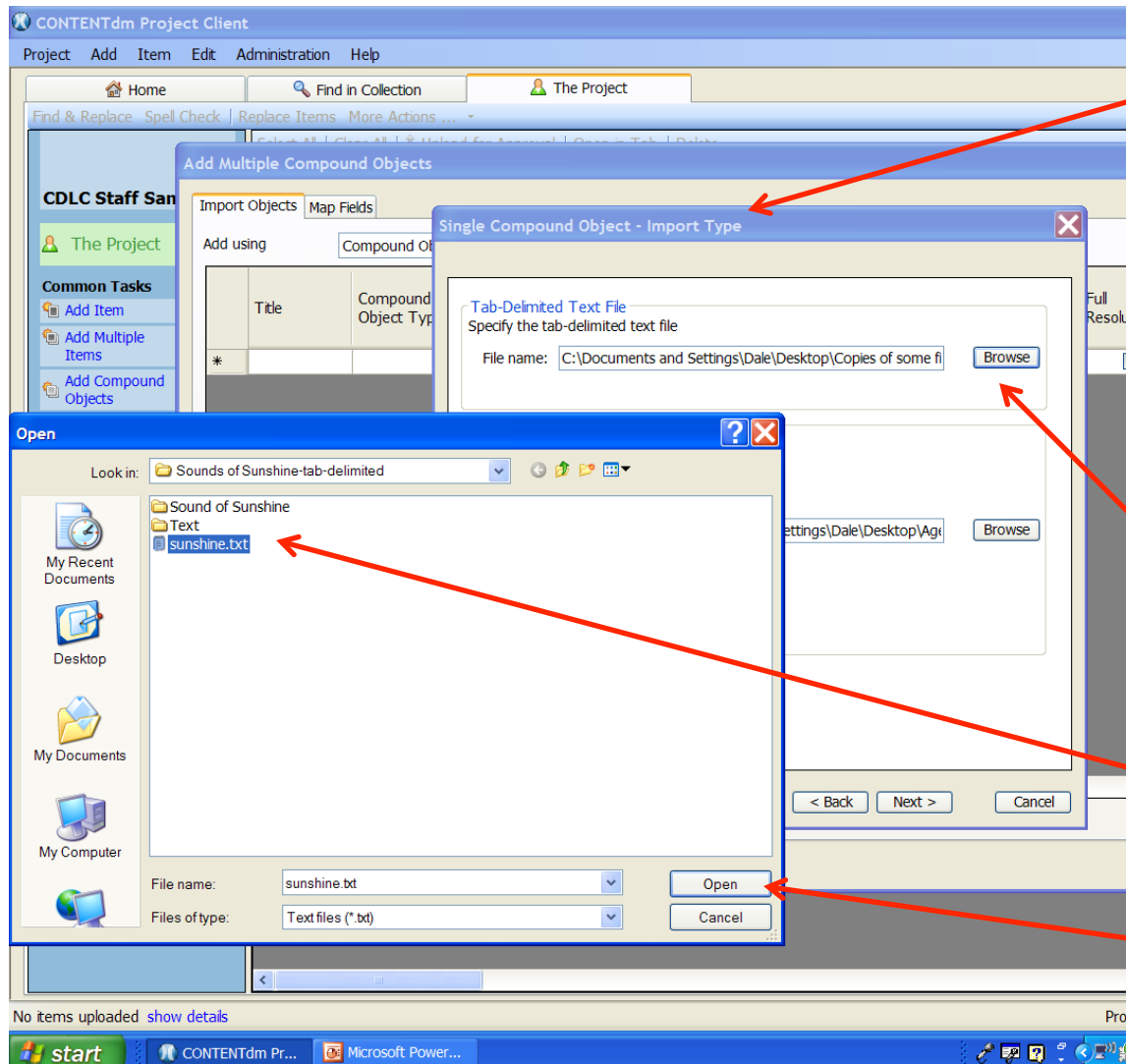Click on **Add** using the **Compound Object Wizard**.

The **Single Compound Object – Choose Type** window opens.

Under, **Import files to create a compound object**, select **Monograph, Document with hierarchy, such as chapters in a book.**

Select **Yes**, for **Are the compound object structure and metadata defined by a tab-delimited text file?**

Click on **Next**.

The **Single Compound Object – Import Type** window opens.

In the **Tab-Delimited Text File** section of the window.

Click on **Browse** and find the tab-delimited file for this object.

Highlight the tab-delimited file.

Click on **Open**.

Back in the **Single Compound Object – Import Type** window

go to the **Item Type** section, and under **What do you want to import?**

Select **Import from a directory.**

Click on **Browse** to find the object files.

Highlight the name of the file folder.

Click on **OK**

Click on **Next** in the **Single Compound Object – Import Type** window.
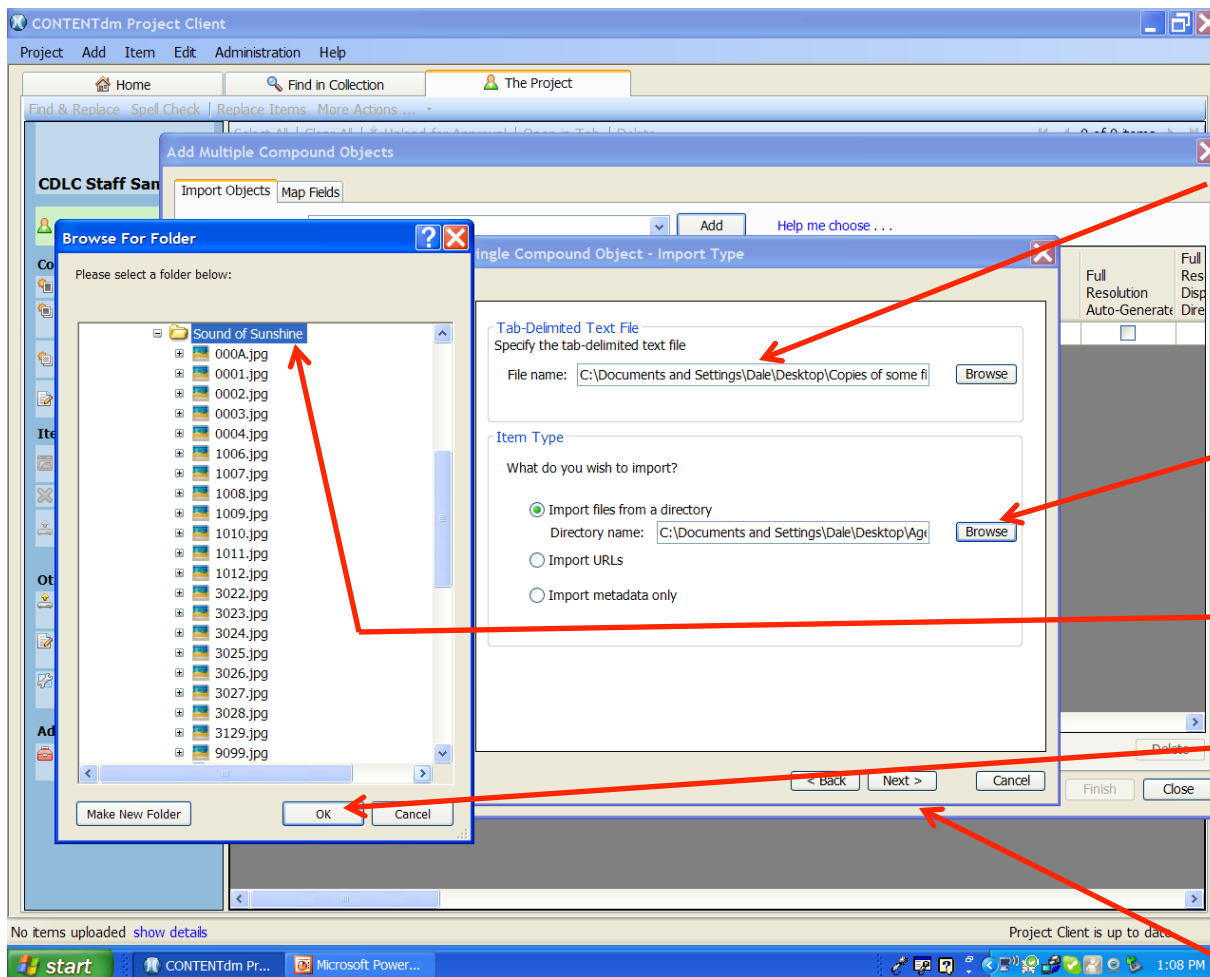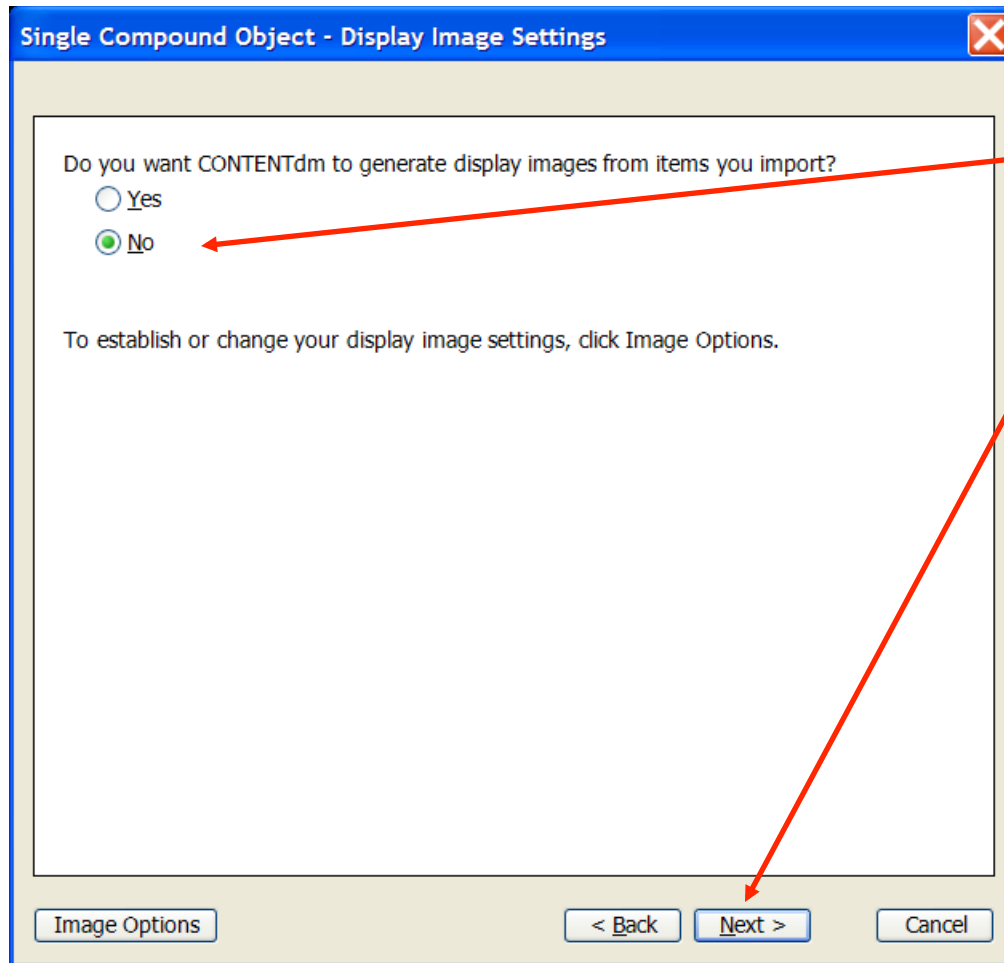
The **Add Multiple Items – Display Image Settings** window opens.

Select **No**.

Click on **Next**.

Note: CONTENTdm Project Client can automatically create a jpg or jpg2000 file. This is what the **Display Image Settings** window is referring to. If you are interested in finding out more about this feature of the Project Client please contact your council.

Single Compound Object - Display Image Settings

Do you want CONTENTdm to generate display images from items you import?

○ Yes
● No

To establish or change your display image settings, click Image Options.

Image Options   < Back   Next >   Cancel
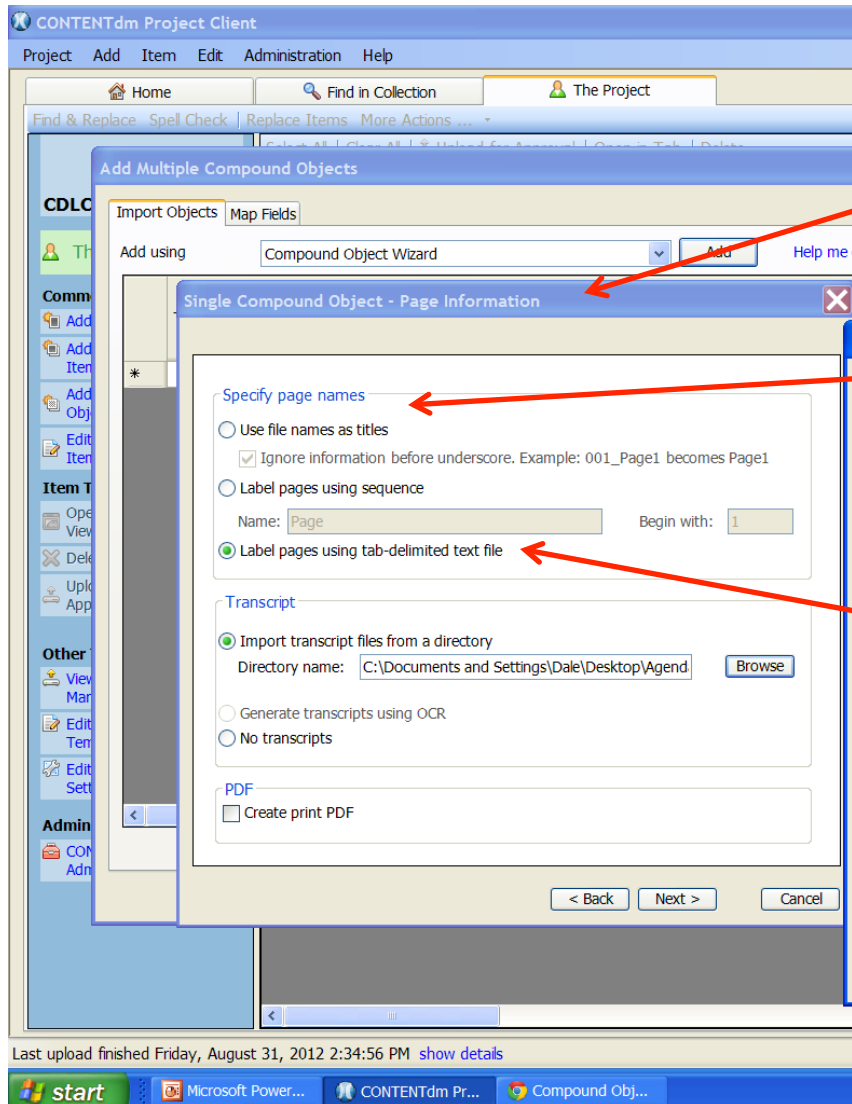
The **Single Compound Object – Page Information** window opens.

Under **Specify page names** section,

for this example,

select **Label pages using tab-delimited txt file**

Our document example has transcript files so,

Under the **Transcript** section of this window,

Select **Import transcript files from a directory**.

Click on **Browse**.

Find the folder with the transcript files, highlight the folder and click on **OK**.

(Note: If the document does not have any transcripts then select **No transcripts**.)

Click on **Next** back in the **Single Compound Object – Page Information** window.

Single Compound Object - Confirm Settings

Summary of settings for import.

Importing single Monograph.
Import using tab delimited file.
Tab delimited file: C:\Documents and Settings\Dale\Desktop\Copies of some files from CDLC work\2012-07-20-TabDelimitedExamples\Sounds of Sunshine-tab-delimited \sunshine.txt
Import Item Type: Object

Title pages using tab delimited field mapped to the collection field Title.
Transcripts are contained in directory C:\Documents and Settings\Dale\Desktop\Copies of some files from CDLC work\2012-07-20-TabDelimitedExamples\Sounds of Sunshine-tab-delimited\Text.

< Back   Finish   Cancel

The **Single Compound Object – Confirm Settings** window opens.

Check the settings.

Click on **Finish**.

Summary

All | Summary (1) | Errors (0) | Warnings (0)

Added monograph containing 21 pages.
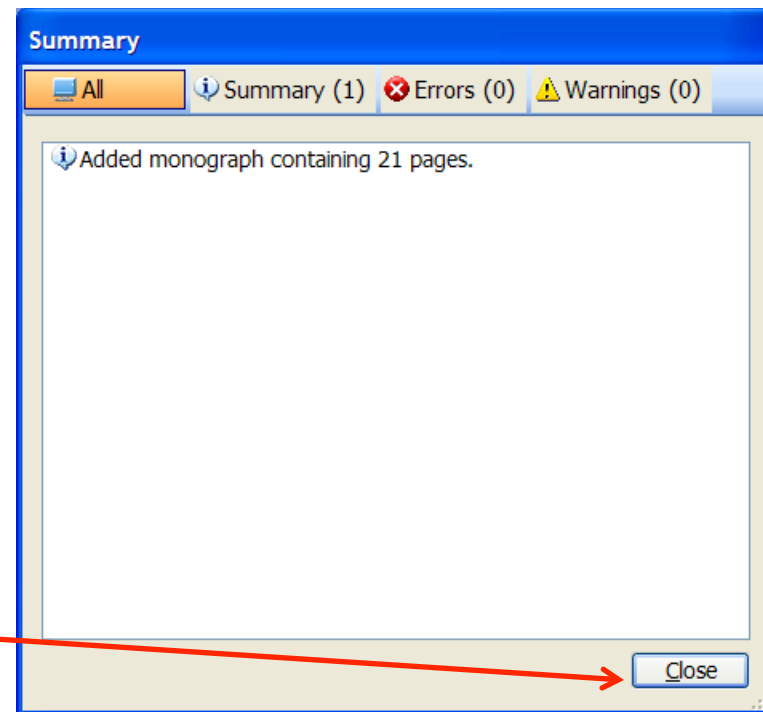
Close

Summary window opens.

If there are no **Errors** or **Warnings** then click on **Close**.

Map Metadata Fields
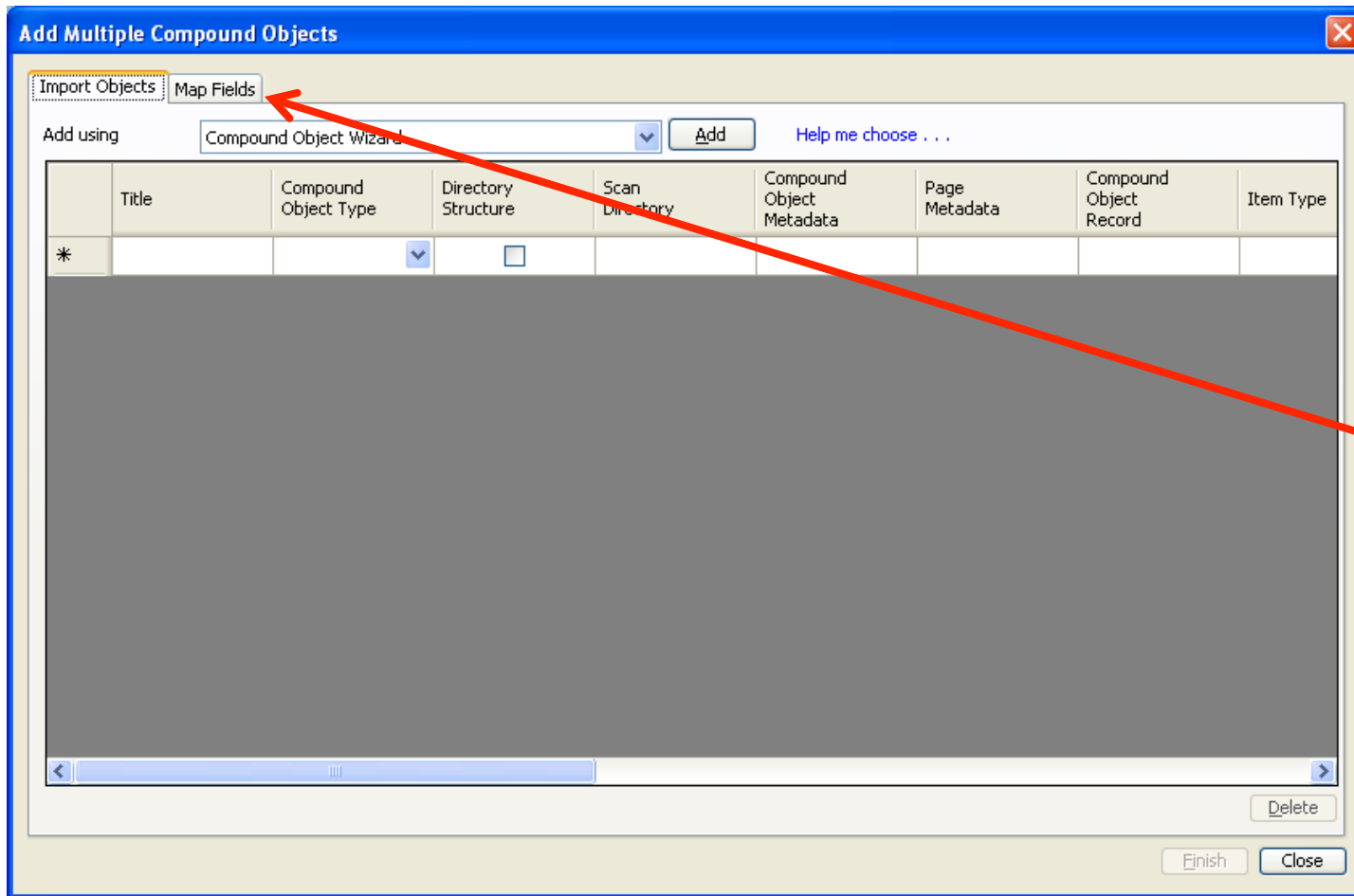
The **Add Multiple Compound Objects** window opens.

Click on the **Map Fields** tab.
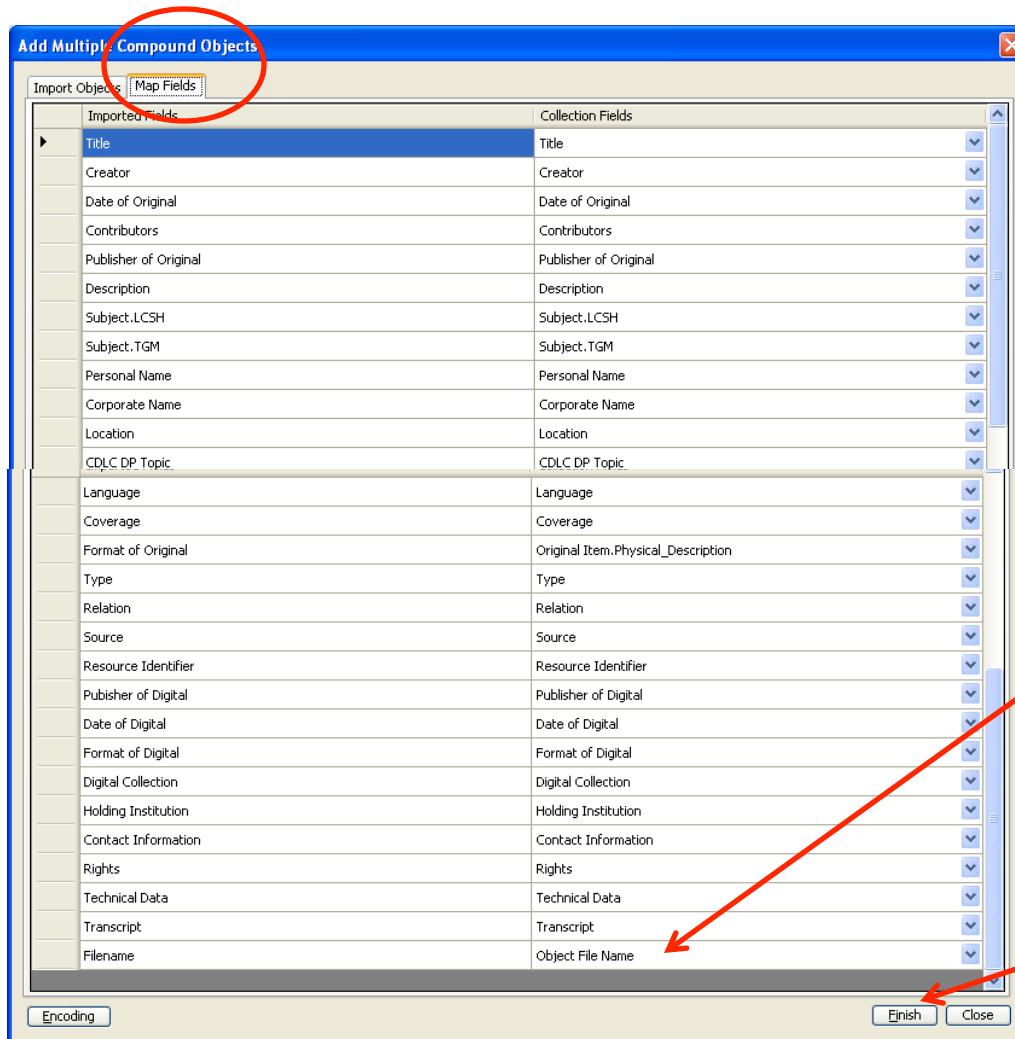
You are now in the **Add Multiple Items – Map Fields** tab view.

<u>This window is very important</u>. You must check to make sure each metadata field* that is being imported is matched to a collection field in the collection template.

Note: **Filename** is mapped to "**Object File Name**" not to the metadata field **File Name.**

<u>Click</u> on **Finish**.

*(file names in example are based on old CDLC template)
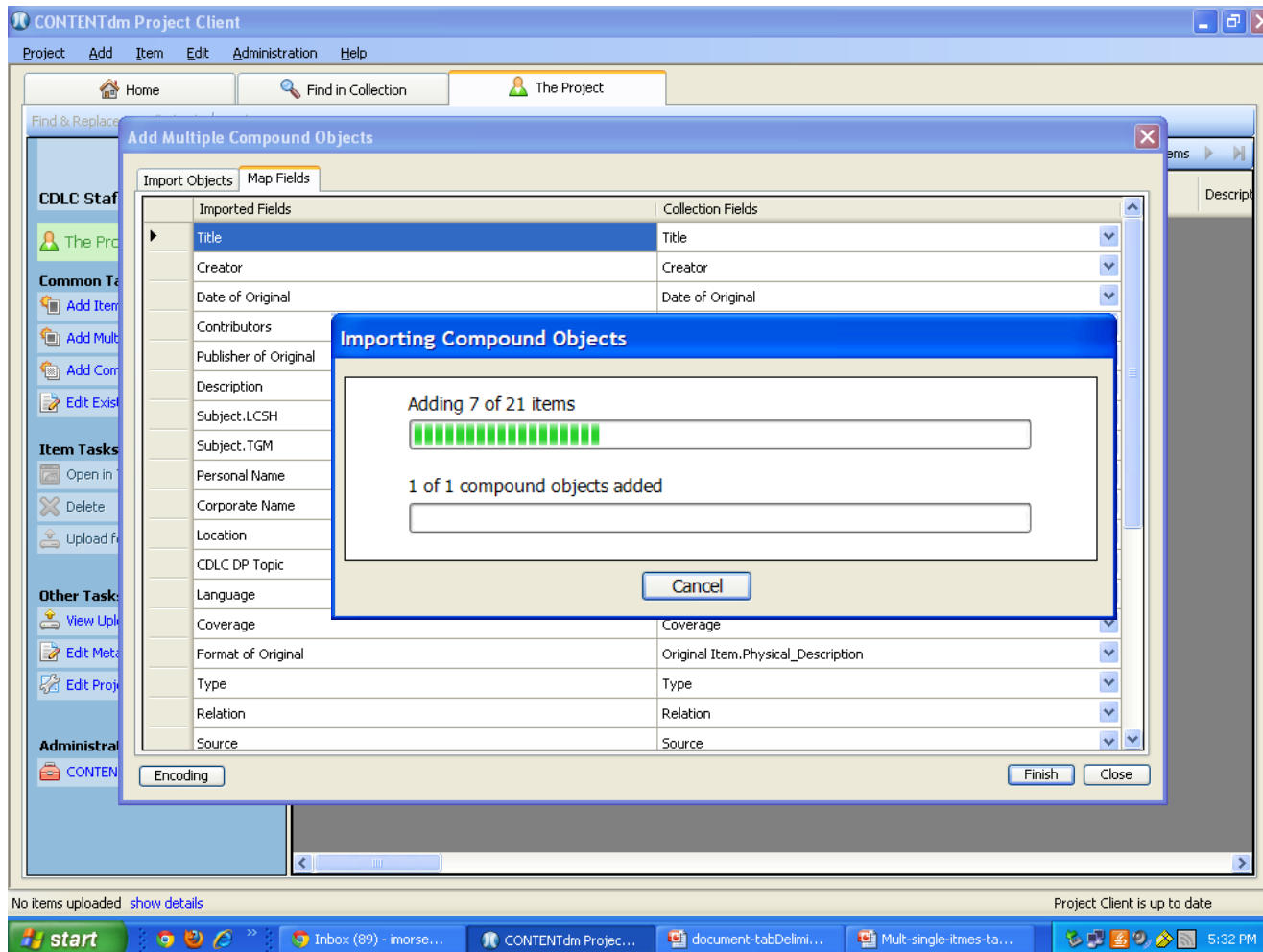
The **Importing Compound Objects** progress bar window opens.
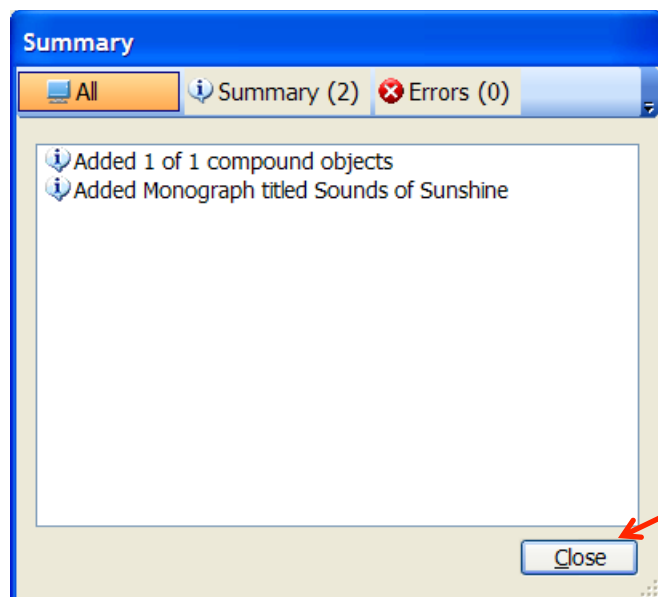
Summary window opens.

If there are no **Errors** or **Warnings** then click on **Close**.

In the spreadsheet view of the Project Client:
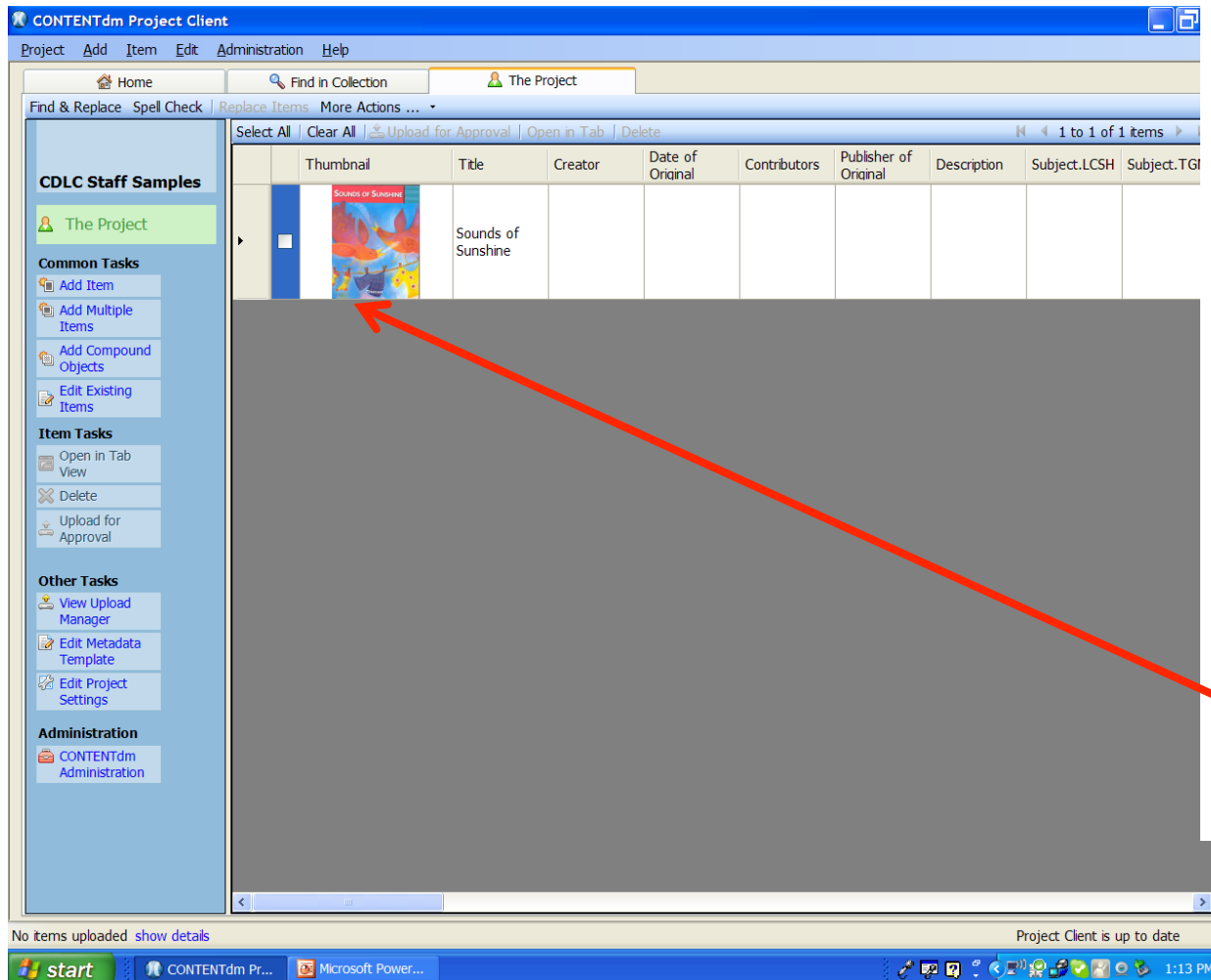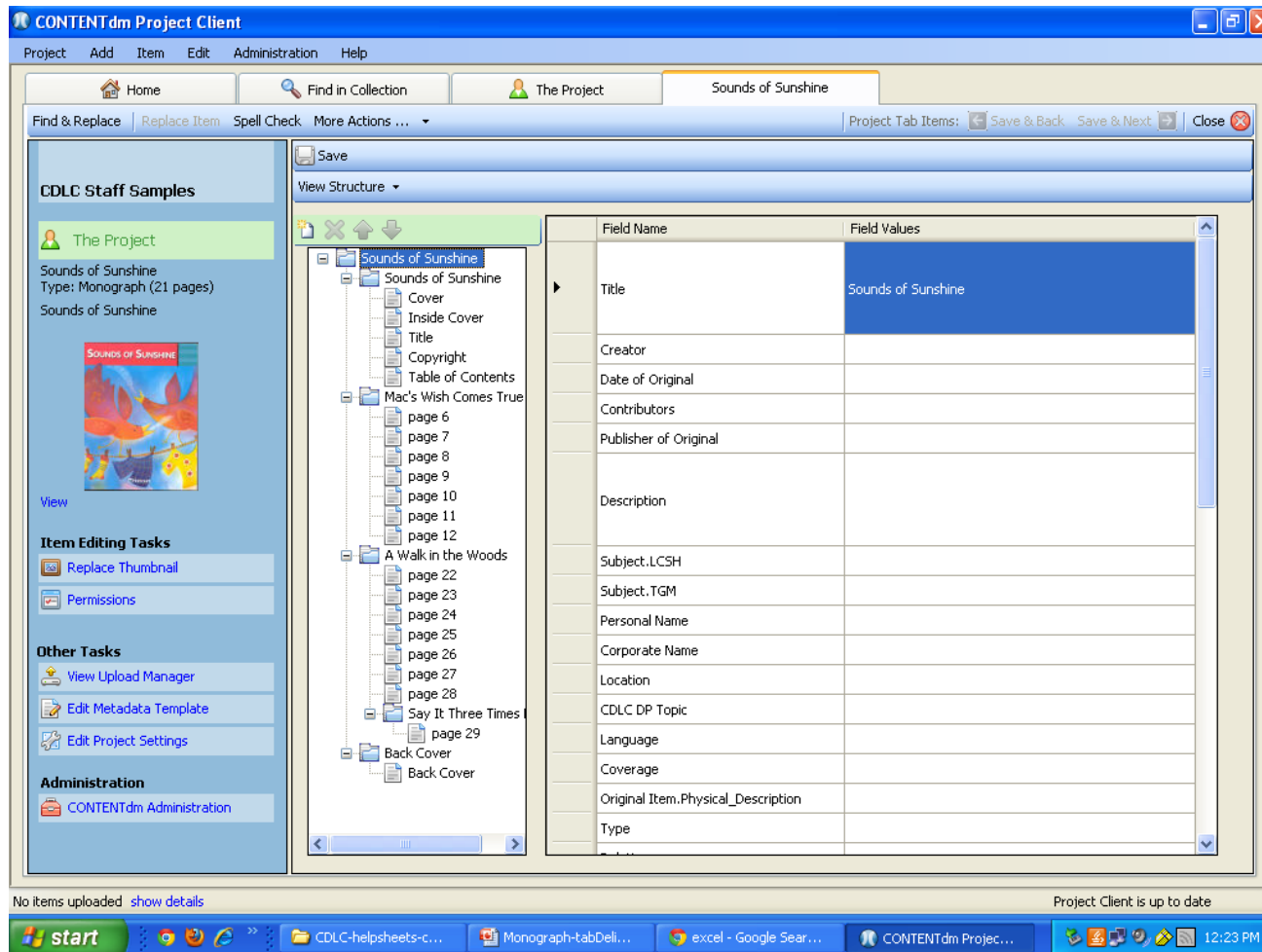
The document (a compound object) has been imported into the **Project Client** with all its associated metadata.

Additional metadata can be added if needed.

Double click the **thumbnail** of the document.

Views of the structure of the compound object:

This is the **View Structure** of the compound object.
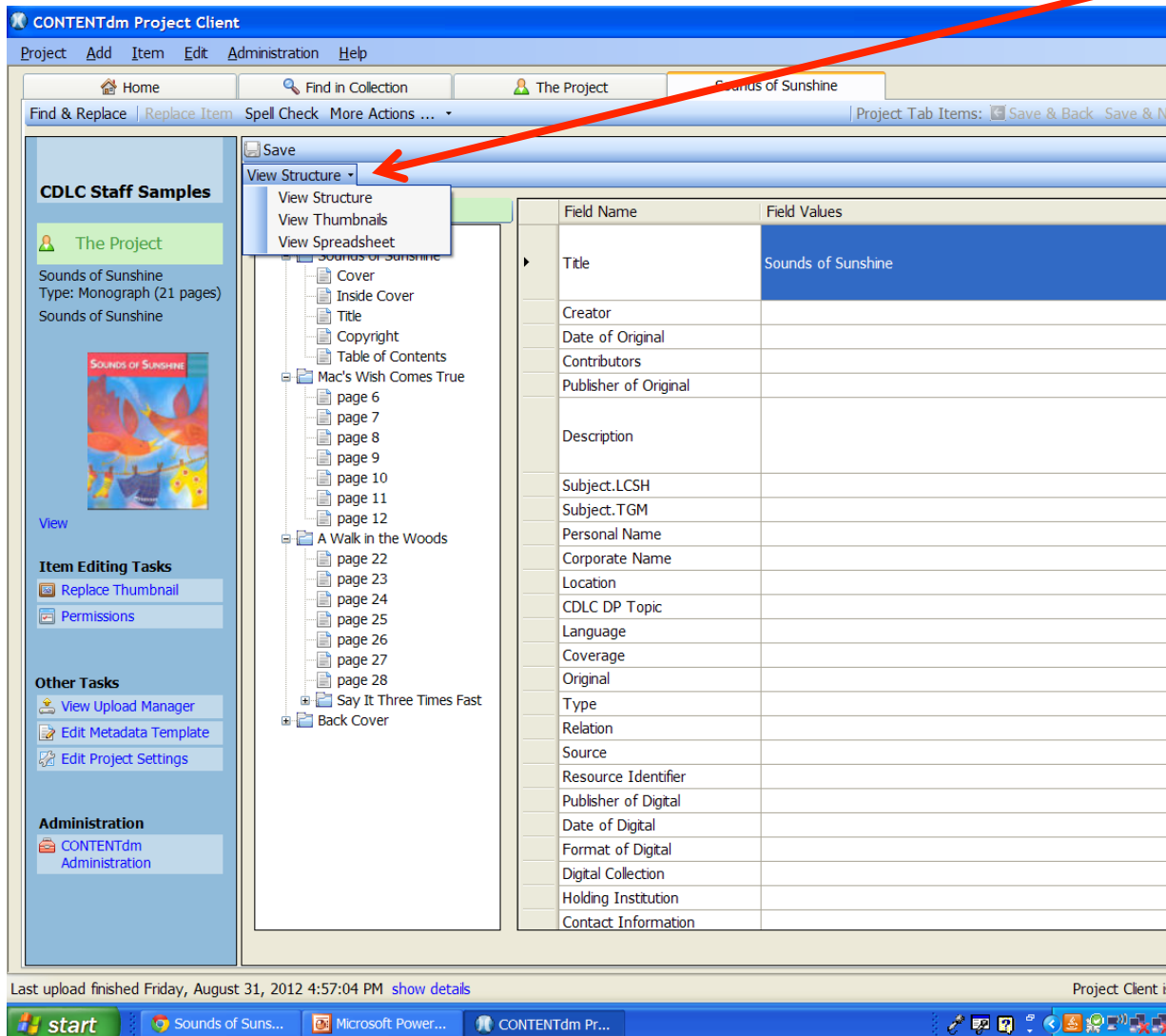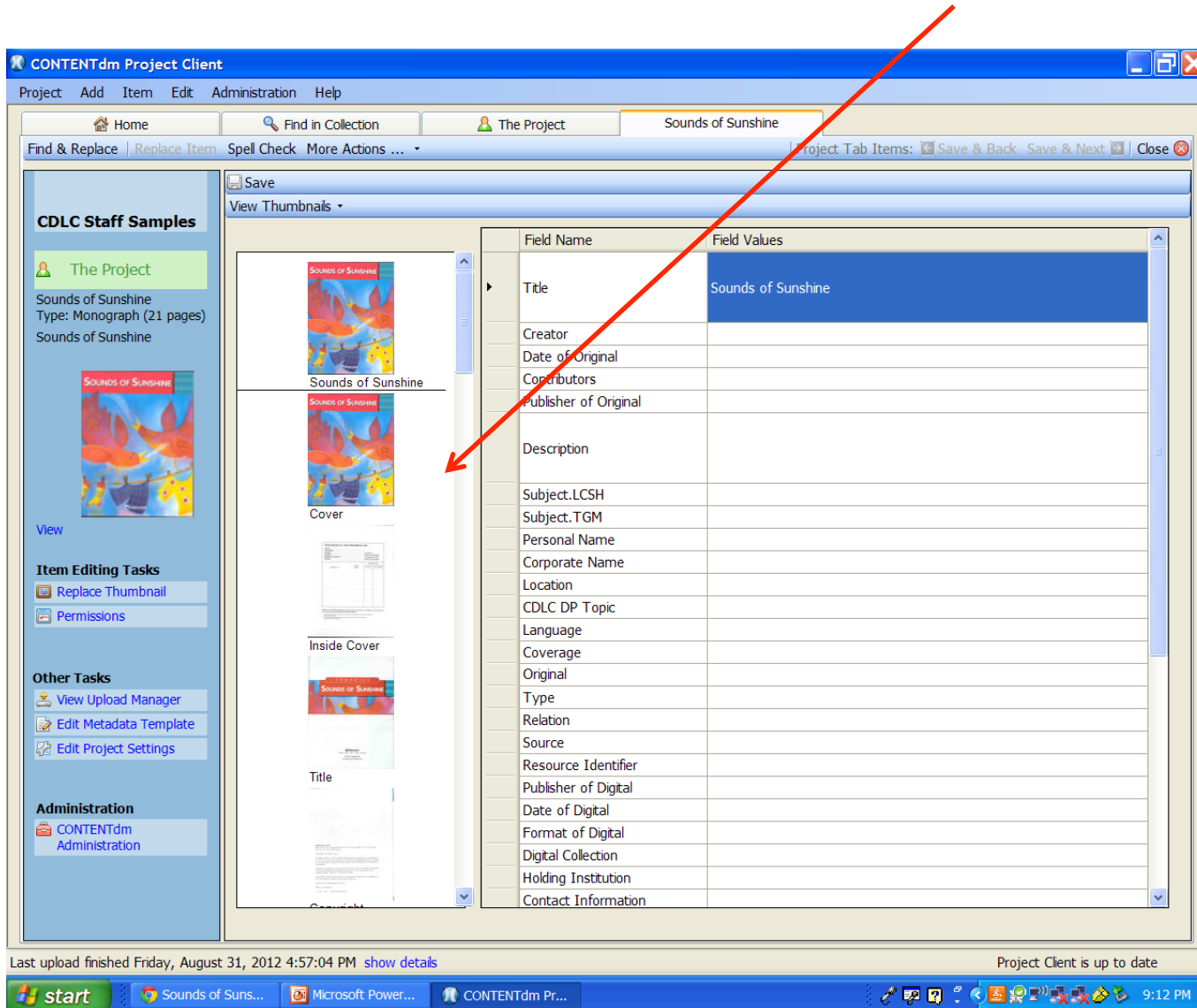
Click on **View Structure**.

There are three views of the structure of the compound object:

View Structure
View Thumbnail
View Spreadsheet

Metadata for the object and page metadata records can be added in any of these three structure views.

Thumbnail structure view:

Spreadsheet Structure view:
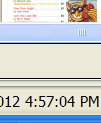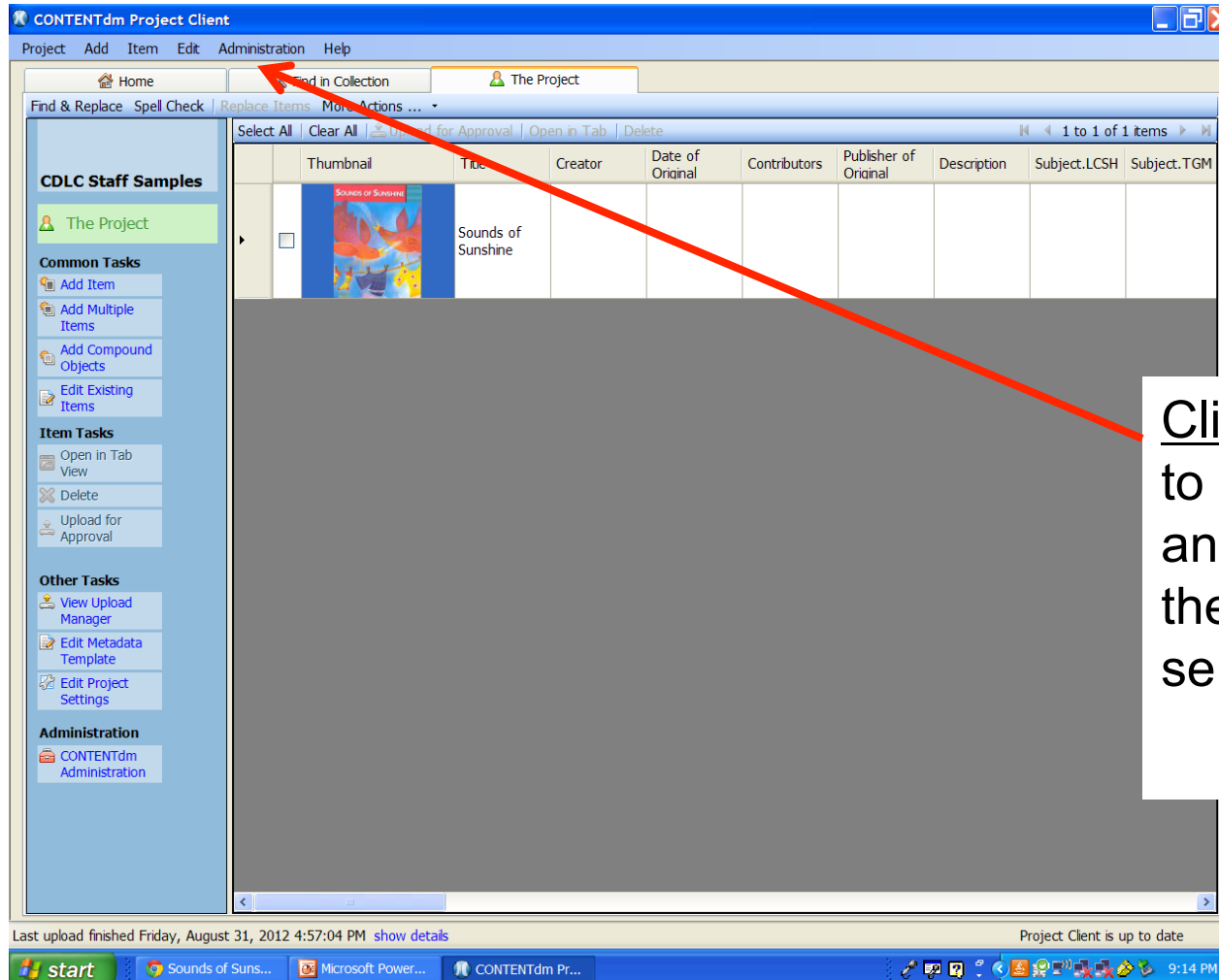
Once all metadata has been entered:

Click on **Save**

Click on **Close**

In **Project tab** view:

Click on **Administration** to **upload**, **approve**, and **index** the items to the CONTENTdm server.