

## Running Batch Import of Image Metadata in Archon

When inputting the metadata of multiple images into Archon's Digital Library (10-15+), it may be beneficial to create a comma separated values (csv) spreadsheet and import that data into Archon directly. This process results in the importation of descriptive metadata including affiliated record series (Collection), identifier number, date, title, description (scope), physical description, contributor, publisher, rights statement, and content URL. This process does NOT allow for the batch importing of image files, subjects, or creators. These items must be added by hand after the batch import is complete. The following guide covers the steps involved in importing csv spreadsheets into Archon using the database management tool. This includes using the correct template, inputting the information, saving the file as a comma separated values spreadsheet, and having a University Archives supervisor import the file.

To see examples of how these imports come to fruition in a practical sense, consult the entries for images 0005145 - 0005163. The descriptive metadata for these digital library records were all imported using the following steps. Please use the saved csv files used to create these imports, found in the DLImportFiles directory of the ER Working Folder, as templates if you wish.

*\*\*Please note, the actual import into Archon detailed in step 4 should be handled by a University Archives supervisor with "database manager" permissions in Archon. Please contact Chris Prom for further information or guidance.*

### 1: Opening the Proper Template

*When attempting to import a csv spreadsheet file of digital library metadata into Archon it is crucial that the correct template is employed. If the wrong template is employed, the information will not be organized properly and will need to be re-input into the spreadsheet and re-imported.*

Find the file titled 'Digital-Content-CSV-Sample' found in W:\archon\incoming or in G:\Archives\AhxAsst\ER Working Folder\Image Scan Resources. Copy this file and paste it to the selected directory that you wish to work from\*\* (most likely G:\Archives\Ahxasst\ER Working Folder\Order\XXXXXXX (where XXXXXXX is the name of the order folder for the research who has requested the images) but could be H:\ Drive or something else that makes sense), then name the file according to some sort of standard. The past standard has been the name of the collection (if all images are from one place) or the name of the researcher requesting the images - Ask Chris or Bill for guidance if you have questions.

**\*\*DO NOT WORK DIRECTLY WITH THE SAMPLE FILE FOUND IN THE W:\ DRIVE OR IMAGE SCAN RESOURCES DIRECTORIES - THESE ARE THE TEMPLATES AND SHOULD NOT BE SAVED OVER.**

Open the copied file and note the example template that has already been started. Each column is formatted in a specific order that must be maintained i.e. column A must be “web output” and column G must be date, etc.). Each column is titled as to what the cells in that column will contain (folder, title, date, creator, etc.). Before beginning to input data, confirm that this template includes all of the fields necessary to import the information that you will include in the finding aid. If you find that there is a problem, please consult with Chris.

## **2: Inputting the Information into the Digital Content Template**

Begin by deleting all of the example data that you do not wish to use as a resource. If you find that you would like it to be visible for consultation, start entering your data in a row below the example data and you can erase the sample data later. From here you should input the information that you know simply by looking at the image. This means that a column such as “collection identifier” will be left blank and addressed later. A few notes about some of the fields:

1) “web-output enabled” should be “yes” unless you are attempting to hide all of the image metadata from public view, in which case it should be “no.” There are no other options.

2) When entering the descriptive information, you can manually enter a carriage return in a cell (in excel) by pressing alt+enter when typing. This will enable you to follow the Archives website format of having a line break in between the narrative description and the reference information in the “scope” field and the word tiff and the image measurement information in the “physical description” field.

3) When entering “titles” and “identifiers,” Archon cannot handle duplicate data. Therefore, to make each “title” unique, follow the name subsequent matching titles with a parenthetical enumeration (e.g Alma Mater, Alma Mater (1), Alma Mater (2), etc.). It is best to do a search of the Archives website for the image titles that you wish to use to see if duplication may be an issue. For identifiers, always ensure that the numbered scan enumeration being used is available.

4) If you know the copyright information, use the formatted statements found in the document “Copyright for Photos in Archon” found in the Image Scan Resources Directory or the WhitledgeDoc directory.

When the data is completed entered into the spreadsheet regarding the descriptive information, date, contributor, etc., discuss with your supervisor any information about the image that you cannot deduce (copyright statement, date, etc.). For the collection identifier DO NOT use the RG/SG/RS format or the seven digit standard UA format. Rather, search for the record series in Archon and then copy the identifier found at the end of the url. E.g. the url on the Archives website for RS: 39/2/20 is:

<http://www.library.illinois.edu/archives/archon/index.php?p=collections/controlcard&id=4014>

The digits 4014 at the end of this are the collection identifier that should be entered into the

“collection identifier” column. Each series has a unique identifier and you should search for all of them the correspond to the images for which you are batch importing the metadata. See the example template for guidance (in the sample 2126 is for Theatrical Production Photographs, rs: 12/11/6).

### **3: Creating a comma separated value spreadsheet (csv) file**

After the data is completely input and the organization is satisfactory, save the spreadsheet in the preferred format. Then delete any rows of data that is not directly related to the information to be imported - This includes the column labels at the top of the sheet and all example entries. To delete rows, right-click on the row number and select delete - merely pressing the delete key with all of the cells highlighted will not delete the cell, only the contents.

After this is complete, select “save as” from the “file” dropdown and select format as “csv (comma delimited).” The file should be named the same as the spreadsheet, but should also include the suffix csv as well as the extension (e.g. radauerimport.xls vs. radauerimportcsv.csv). The files should be saved to G:\Archives\Ahxasst\uasfa. After reading the prompts and selecting the appropriate responses, you are complete and the file is ready to be imported.

### **4: Importing CSV Files into Archon**

*\*\*Please note, this step should be handled by a University Archives supervisor with “database manager” permissions in Archon. Please contact Chris Prom for further information or guidance.*

With the csv file saved to the proper directory, it is now time to upload the file into Archon. To do this, enter the Archon Administrative interface and select “database management” from the Archon Administration package. Then select the “Digital Content Metadata csv” import module and press “launch.” You are then prompted to select a file or browse to upload a new file, which is the choice you want. Then find the file with the proper name and the csv suffix and then click on “import.” Archon will now import the file, which may take some time depending on the number of cells that Archon is reading through. After it is complete a report will be generated which will note any errors. The most common errors occur because duplicated information is being entered (i.e. you have titles that are the same) or because the collection identifier cannot be found (most likely because the actual series number was used and not the collection identifier assigned by Archon found in the url for the series).

After the import is complete, open each digital library file (usually search by identifier number) and check the metadata to confirm that it is correct. If there are minor errors, those can be fixed in Archon using the edit pencil. If there are substantial errors throughout all of the entries, it may be necessary to delete all of the unique entries, edit the original spreadsheet, re-save the csv and then run the import again. While confirming the metadata, enter line breaks that have been missed in the descriptive fields, associate subjects, and link the embossed jpg image.

## **5: Saving the Digital Library Import Files**

Upon completion of the import, the excel file used to create the batch import CSV file should be saved. These files should be saved to the DLImportFiles folder and named with a unique identifier followed by DONE. This DONE indicates that the file has been imported and there is no need to upload again. These files will remain in this directory for the foreseeable future.

W:\staff\documents\analog\csv.metadata.wpd