

# oVert FMNH Archiving Workflow 2018-20

Date Created: Autumn 2018

<b>Specimens loaned to scanner</b>	<b>2</b>
In EMu	2
Catalogue Records	2
Transactions	2
<b>Specimen scanned</b>	<b>2</b>
At Scanning Institution	2
<b>Derivative (Image Stack) Uploaded to Morphosource</b>	<b>3</b>
In MorphoSource	3
Specimen record located	3
Stack uploaded to MorphoSource	3
Drive with data/files returned to Field Museum	3
<b>Data Archived in EMu</b>	<b>4</b>
Data transfer to FMNH	4
In EMu	4
Record Structure	4
Preparation	5
File names	5
Check Group	5
Create 3 Multimedia records	5
Raw Stack / Raw data	5
Derivative Stack / Image Stack	11
The URL Record	13
Make the catalogue attachments	16
Repatriate the multimedia Identifier (UUID) to Morphosource	18
<b>End of project tasks</b>	<b>18</b>
<b>Notes</b>	<b>18</b>
1-April-2019 (kwebbink)	18
Files to transfer from UofC to FMNH:	19
13-May-2019 (kwebbink)	21
1-Nov-2019 (meeting notes)	21

# Specimens loaned to scanner

## In EMu

### 1) Catalogue Records

- a) There must be a catalogue record in EMu for any specimen that goes to be scanned.
- b) In order for iDigBio to pick up the record it must have "Publish in Internet = Yes"

### 2) Transactions

- a) Catalogue records must be attached to a transaction record in EMu.

# Specimen scanned

## At Scanning Institution

- 1) Include the scanning workflow here for future reference if we have it.
- 2) Create the preview jpg for the RawData
- 3) Create the RawData and ImageStack tar files
- 4) Create the ReadMe file

Note: When to do separate head scans? Eg a 2nd scan is considered for long thin things.

Rough/Rule of thumb is that if the Head < 10% of total size of scan then this is a decision point as to whether a head scan is taken. Particularly important for specimens where the bone structure is very delicate. See file naming convention.

# Derivative (Image Stack) Uploaded to Morphosource

## In MorphoSource

- 1) Specimen record located
  - a) Cross-check OccurrenceID (UUID)
  - b) Attached to relevant Main Field Museum Project
  - c) Attached to oVert Project
  - d) Ownership of specimen record assigned to Field Museum Project
- 2) Stack uploaded to MorphoSource
  - a) Attach to specimen record
  - b) Include/check citation text
  - c) Set to private
- 3) Drive with data/files returned to Field Museum
  - a) File names consistent (triples)
  - b) No nested scans
  - c) Tar each "Raw Data" and Zip "Image Stack" directory using 7-zip

# Data Archived in EMu

## Data transfer to FMNH

As of October 2018 - transferring via portable SSD (SanDisk Extreme 500GB)

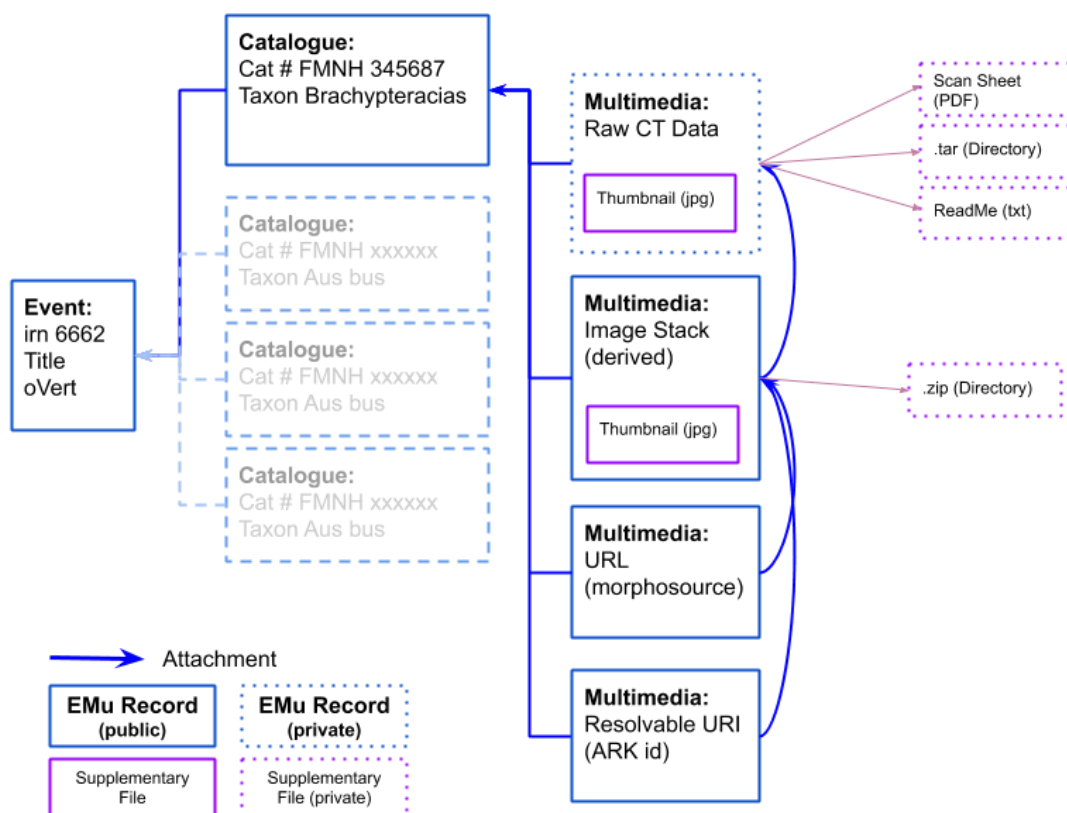
As of spring 2019 - setting up FMNH Globus endpoint

## In EMu

### Record Structure

LINKED: 3D CT EMu data structure 2018

<https://docs.google.com/drawings/d/1-HDE8zyllMnu3FjzHs24vAco-vQ1RuivGeeCDKYorIo>



#### 4) Preparation

- a) Go to event record 6662
  - i) Events 1: Add a description.
  - ii) Objects tab: Check and attach specimen catalogue records that have been loaned to oVert for scanning to the oVert event irn 6662. If there is no catalogue record check with the collection manager and get records created as appropriate. You can also check the loan record in transactions.

#### 5) File names

- a) Check that files/directories are named correctly document naming using this convention
  - i) FMNH\_H\_123456\_Genus[\_subject part]
    - (1) First scan has no subject part suffix
    - (2) Subsequent scans include a suffix Subject part = e.g., "body" or "head"
  - ii) Zoology: H - herps, M - mammals, B - birds; F - fishes
  - iii) Fossils: PR - herps, , PM - mammals, P - birds, PF - fishes

#### 6) Check Group

Make sure that you are logged into the appropriate user group for the records you will be editing/uploading

- a) BirdsMgr for Birds records.
- b) HerpsIntern for Amphibians & Reptiles records.

#### 7) Create 3 [Multimedia records](#)

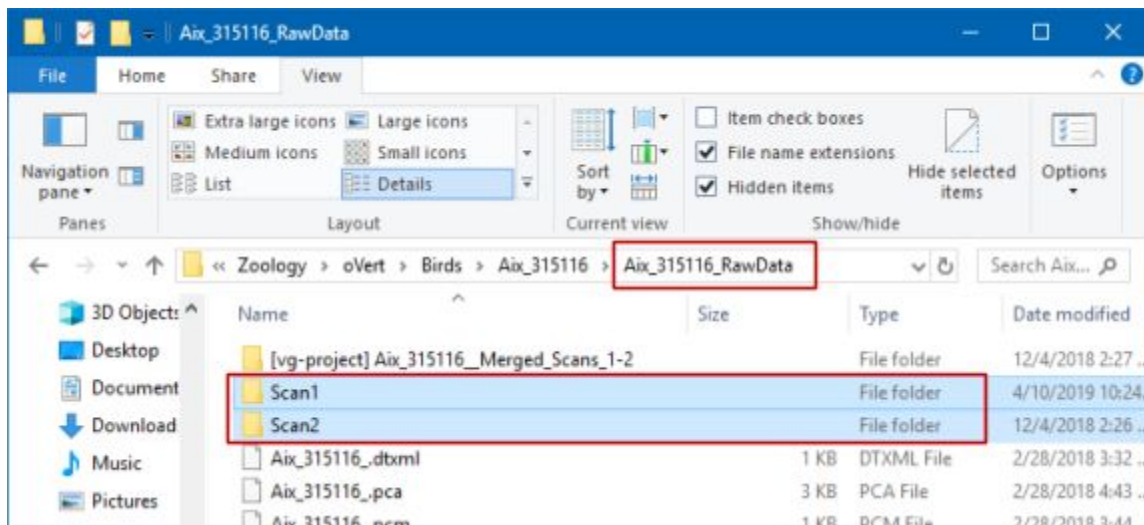
Open the Multimedia module

- a) Raw Stack / Raw data
  - i) Navigate to wherever the assets are. Either network storage or the hard-drive

This is for the Raw Scan Stack transferred from UofC to FMNH  
[ old path: \\serenity\Imaging\_Scratch\Zoology\oVert ]  
**new path: \\serenity\globus-oVert\Zoology**

**Note:** For larger specimens whose resulting Image Stack is stitched together from multiple scans, make sure the multiple scan folders are nested inside the “RawData” folder.

e.g., 2 scans for FMNH B 315116 (Aix sponsa):



ii) Select the “oVert CT Raw Data” defaults

- (1) New Record
- (2) If you have already selected defaults then things will be filled in otherwise
  - (a) Go to Edit Default values
  - (b) Change and select “oVert CT Raw Data” defaults

iii) Fill in the Multimedia tab:

- (1) Title = “scan name” CT Raw Data
  - (a) Replace “[ENTER SCAN NAME HERE]” with “Genus SpecimenNo”
- (2) Description = “scan name” CT Raw Data. Scanned as part of the oVert project. Additional data is on the supplementary tab.
  - (a) Replace “[ENTER SCAN NAME HERE]” with “Genus SpecimenNo”
- (3) Creator:
  - (a) individual who scanned the specimen (noted on the scan sheet (Role = Scanner)
  - (b) Collection Area (Role = Creator)
- (4) Other Numbers = Catalogue number
- (5) Other Number Source = “FMNH Catalogue Number”

iv) Fill in the Details tab:

- (1) Keywords = Add "CT scan" and "oVert"
  - (2) Date = (DD Mon YYYY) for when the raw data was generated
  - (3) Date Description = "Created"
  - (4) Source = "Zoology Birds Collection" or "Zoology Amphibians and Reptiles Collection"
  - (5) Resource Type = "Image | CT Data"
  - (6) Rights - CC-BY-NC
- v) Security tab:
- (1) Department = [Collection Area]
  - (2) Publish on Internet & Intranet = No
- vi) SAVE
- vii) Representative jpg from the set (take a decent slice)
- (1) Select the tiff you want in something like MSPaint and save it as a jpg on the Desktop
  - (2) Filename = "scan name"\_"frame number\_"preview".jpg
    - (a) Eg Centropus\_345651\_00056\_preview.jpg
    - (b) If multiscan, Centropus\_345651\_Scan1\_00056\_preview.jpg
  - (3) Navigate to the "Multimedia" tab of the Raw Data multimedia record. Drag the preview .jpg into the large white box on the left-hand side of the window.
- viii) SAVE
- (1) After saving the EMu record, you can delete the jpg from the Desktop
- ix) Supplementary Tab
- (1) Rename the file as "scan\_name"\_"ScanSheet".pdf
  - (2) Add the pdf's of metadata scan sheet(s) to the Supplementary table
  - (3) Add the text file of directory contents/structure
    - (a) Name the file as "scan\_name"\_"ReadMe".txt

## 0. Copy RawData directory to desktop

Navigate to the RawData directory:

1. On a PC, open **cmd.exe**

(Press "Start/Windows"-button;  
type "**cmd**";  
press Enter)
2. In the cmd.exe interface, use the "**cd**" ("change directory") command to navigate to the Raw Stack directory.

Type this:

**cd Desktop\NameOfDirectory**

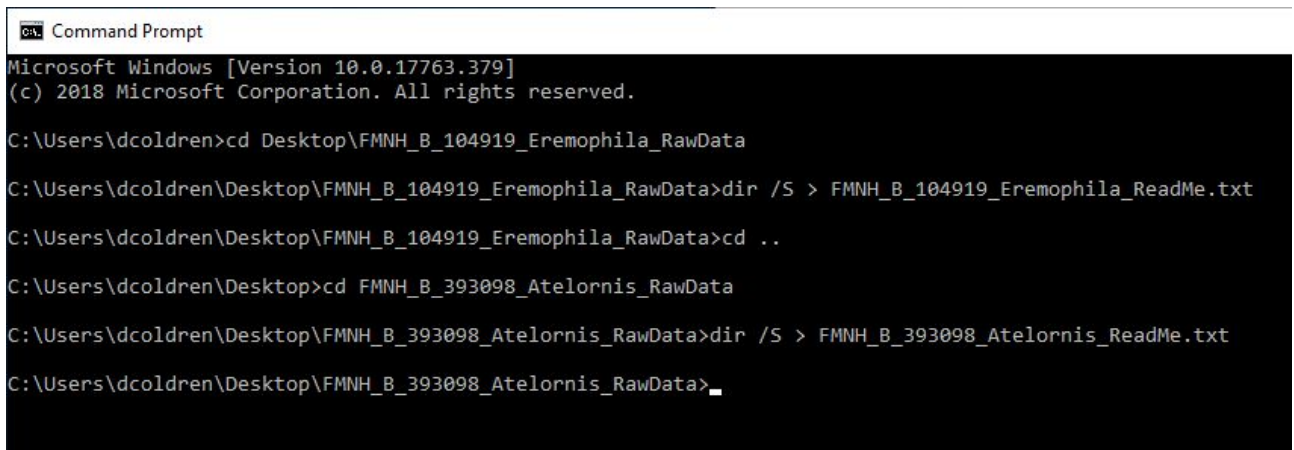
Or if you need to navigate to a different drive (i.e., if the folder's path doesn't begin with "C:\"), type:

**cd /d Y:\FMNH-Birds [etc]**

3. Then use the "**dir**" command to print the list of files in that directory to a "ReadMe" text file:

**dir /S > "scan\_name"\_ReadMe.txt**

*The command "**dir**" lists all files in a directory;  
- Its argument **/S** includes all files in subdirectories;  
The command **>** prints the output of preceding  
commands into a new file (whose name is specified  
after the **>**)*



```
Command Prompt
Microsoft Windows [Version 10.0.17763.379]
(c) 2018 Microsoft Corporation. All rights reserved.

C:\Users\dcoldren>cd Desktop\FMNH_B_104919_Eremophila_RawData
C:\Users\dcoldren\Desktop\FMNH_B_104919_Eremophila_RawData>dir /S > FMNH_B_104919_Eremophila_ReadMe.txt
C:\Users\dcoldren\Desktop\FMNH_B_104919_Eremophila_RawData>cd ..
C:\Users\dcoldren\Desktop>cd FMNH_B_393098_Atelornis_RawData
C:\Users\dcoldren\Desktop\FMNH_B_393098_Atelornis_RawData>dir /S > FMNH_B_393098_Atelornis_ReadMe.txt
C:\Users\dcoldren\Desktop\FMNH_B_393098_Atelornis_RawData>_
```

**Shortcut:**

**cd /d Y:\FMNH-Herps\FMNH\_H\_344305\_Anolis\_Body**

**dir FMNH\_H\_344305\_Anolis\_Body\_RawData /S >  
FMNH\_H\_344305\_Anolis\_Body\_ReadMe.txt**

**You stay in the specimen's main folder, while telling the computer to go look in the RawData folder and list its contents into a ReadMe.txt generated into the main folder. Then you don't**



**have to move the .txt file after generating it.**

**[Example 2] - run the following two lines:**

**cd Desktop\Centropus\_345651**

*(This would go to a “Centropus\_345651” folder on the Desktop)*

**dir /S > Centropus\_345651\_ReadMe.txt**

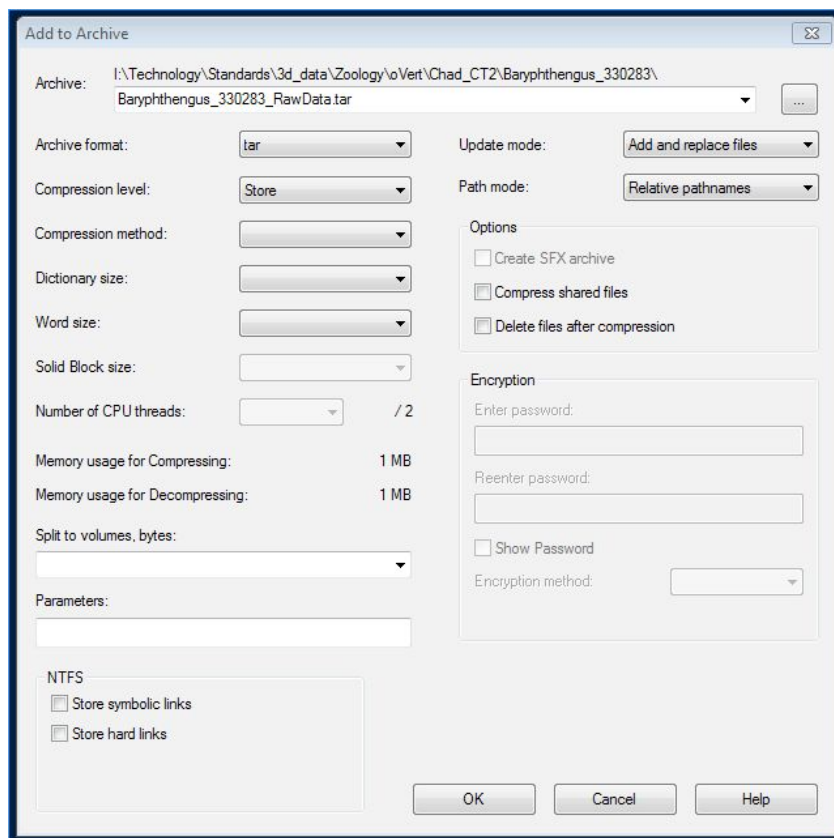
*(This would make a new file called “Centropus\_345651\_ReadMe.txt” inside of the Centropus\_345651 directory)*

- On a Mac / from terminal, navigate to Raw Stack directory, and run:  
`ls -a -R > “scan_name”_ReadMe.txt`

x) **SAVE**

xi) **Make a .tar directory using 7-Zip (<https://www.7-zip.org/download.html>):**

- (1) In your file browser, right-click the “RawData” directory
- (2) Choose 7-Zip → “Add to archive...”
- (3) Fill in the window as shown below (name the archive as appropriate) & click OK:



xii) For reference the directory should contain these file types. Add in any new ones as you come across them.:

- (1) tiff/dng/bmp(s)
- (2) vg-project directory
  - (a) dtxml - scanner-generated file
  - (b) pca - use this to render and reconstruct raw data into a .vgl/.vol
  - (c) Pcj - scanner-generated file
  - (d) pcp - scanner-generated file
  - (e) pcr - scanner-generated file
  - (f) Vgl - open with VG Studio to reconstruct into a .tiff stack
  - (g) Vol - associated with .vgl
- (3) Reference tiffs - maybe a calibration file?

Use 7Zip

For a ~5GB directory:

- Tarring will take ~1 minute
- Adding to the Supplementary tab will take ~3 minutes
- Saving completed Multimedia record will then take ~6 minutes.

Now add the files to the supplementary tab in the multimedia record

Add ScanSheet.pdf

- Highlight the first row in the supplementary table
- Click the green “+” to add a supplementary file
- Navigate to the [scan name]\_ScanSheet.PDF

Add ReadMe.txt

Add .tar

## b) Derivative Stack / Image Stack

This is for the derivative Image Stack that was uploaded to MorphoSource. Its parent will be the Raw Data.

### i) Create the Image stack record

Make sure that you are on the correct Raw Data multimedia record then Ditto the Raw Data Record

- (1) Edit → Ditto → Use existing record
- (2) File → New record
- (3) Edit → Ditto → All fields

Open up the Morphosource page for the specimen

Eg [https://www.morphosource.org/Detail/MediaDetail/Show/media\\_id/22055](https://www.morphosource.org/Detail/MediaDetail/Show/media_id/22055)

- ii) Copy the Morphosource preview jpg to your desktop
  - (1) Right click the preview and save the image do not change its file name.
- iii) Drag the jpg into the multimedia window
- iv) Change the Title = "scan title" CT Scan Image Stack
- v) Change the Description
- vi) Add the Morphosource Media Number to the Other Number table
  - (1) Other Number Source = MorphoSource Media No

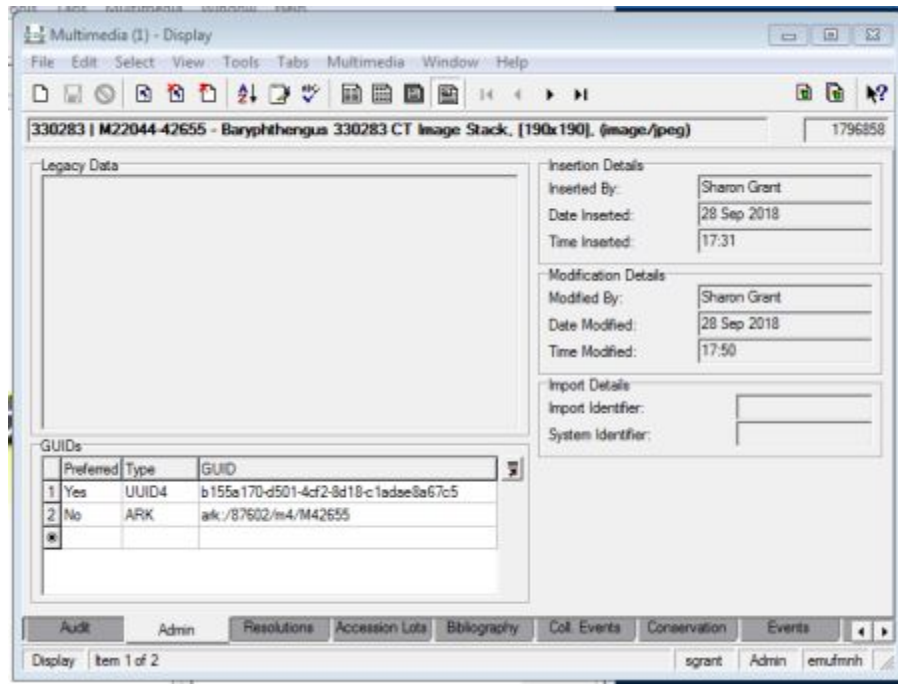
SAVE

- vii) Associations tab -- Set the record for the raw data as this record's parent.
  - (1) In the Parent Media field, enter "[scan name] Raw Data" and click the green "+"
  - (2) On the View Menu select Refresh
- viii) On the Supplementary tab, add the **.ZIP** directory of the derivative image stack jpg/tiff
  - (1) Highlight the first row
  - (2) Upload the **.ZIP** (~2 min) by clicking the green cross and select the Image Stack
- ix) On the **Security** tab
  - (1) Set "Publish on Internet" and "...Intranet" to **Yes**
  - (2) Save the record.

Go to the Admin Tab

- x) Morphosource ARK ID and DOI, and Media number
  - (1) On Admin tab, add ARK id (if available) to the GUID table





xi) Save the derivative record (~3 min)

c) The URL Record

This is for the Morphosource URL of the derived image stack.

Make a new EMu record

i) Select the “oVert CT URL record” insert default:

<http://intranet.fieldmuseum.org/emu/node/6333>

ii) On the Multimedia tab,

(1) Go to Multimedia → Add Resource → URL, and paste the MorphoSource URL for the scan into the window.

(2) Update the title to: “[scan name] URL for CT Image Stack”

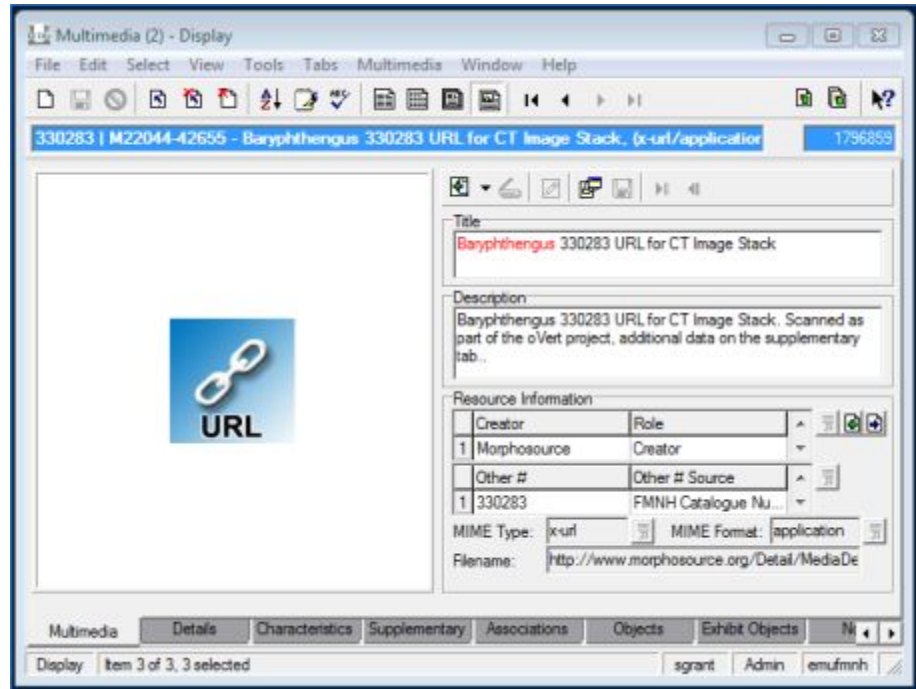
(a) Replace [ENTER SCAN NAME HERE] with the scan name eg “Centropus 345651”

(3) Update the Description: “[scan name] URL for CT Image Stack. Scanned as part of the oVert project, additional data on the supplementary tab.”

(a) Replace [ENTER SCAN NAME HERE] with the scan name eg “Centropus 345651”

SAVE

(4) Update the Creator to “Morphosource”, with Role = “Creator”

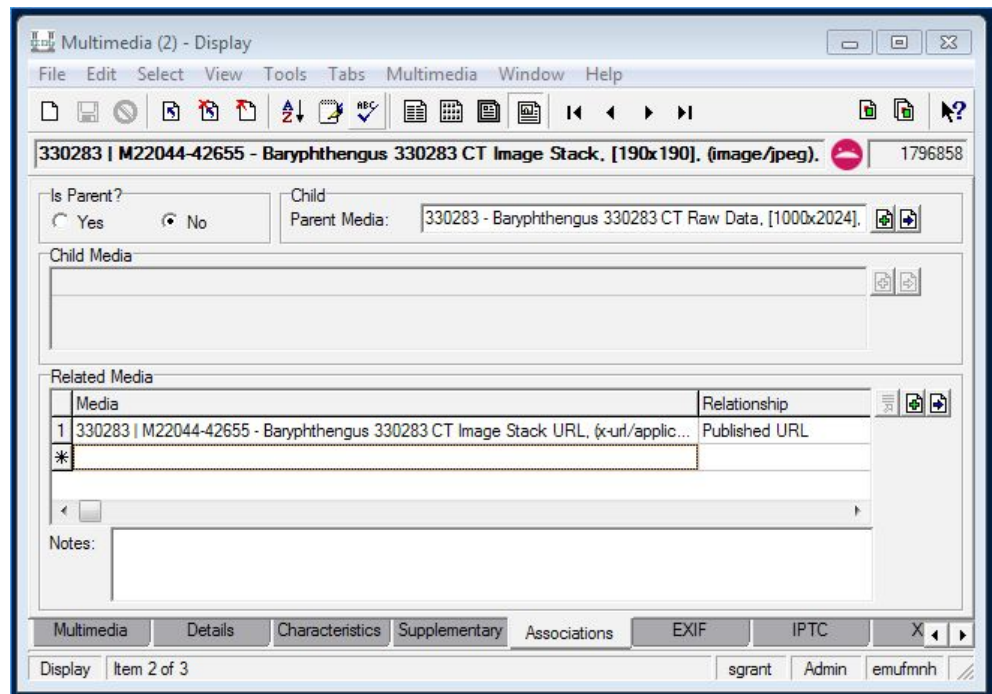


- iii) On Details tab,
- (1) Set Resource Type = “**URL**”
  - (2) Erase the Date
  - (3) Source = MorphoSource
  - (4) In the Rights field, erase the current value, and type “**URL**” and hit tab.  
 (“**[Public Domain] - URL**” should appear in the field after a moment.)
  - (5) Save the record.

- iv) On the Associations tab,
- (1) Delete the Parent Media attachment.
  - (2) Attach the corresponding CT Image Stack in “Related Media”
  - (3) For Relationship, enter “**Published Dataset**”

- d) On the **Security** tab
- i) Set “Publish on Internet” and “...Intranet” to **Yes**
  - ii) Save the record.

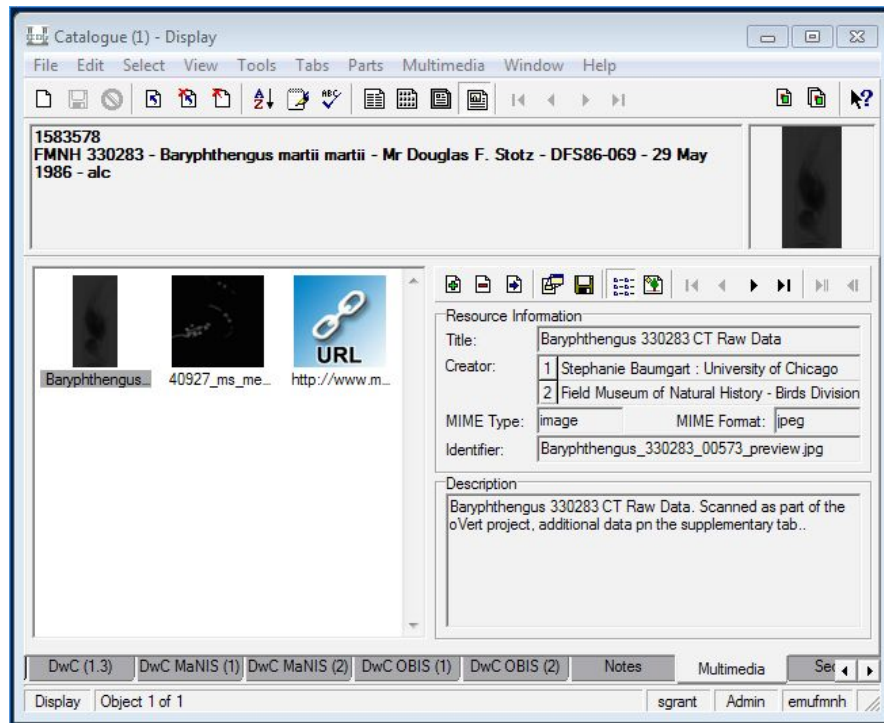
- e) Return to the Derivative CT Image Stack Multimedia record
  - i) On the Associations tab, fill in the “Relationship” for the attached URL:  
**“Published URL”**



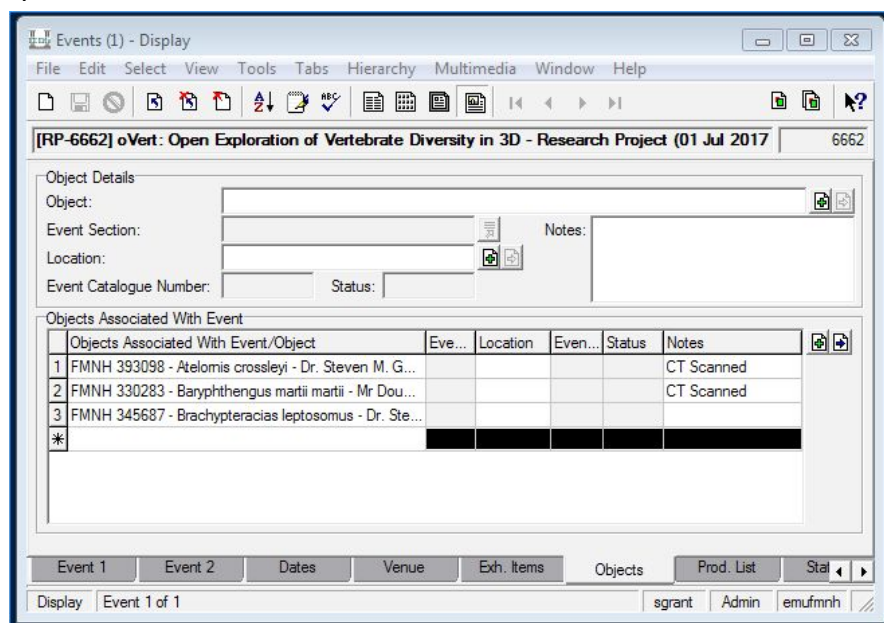
## 8) Make the catalogue attachments

- a) Attach all three multimedia records to catalogue record
  - i) In the Multimedia module, highlight the three Multimedia records (raw, image stack, and URL)
  - ii) In the Catalogue module, retrieve the specimen record, and go to its Multimedia tab.
  - iii) Drag the “attach” icon from the Multimedia module into the Catalogue record’s media box.





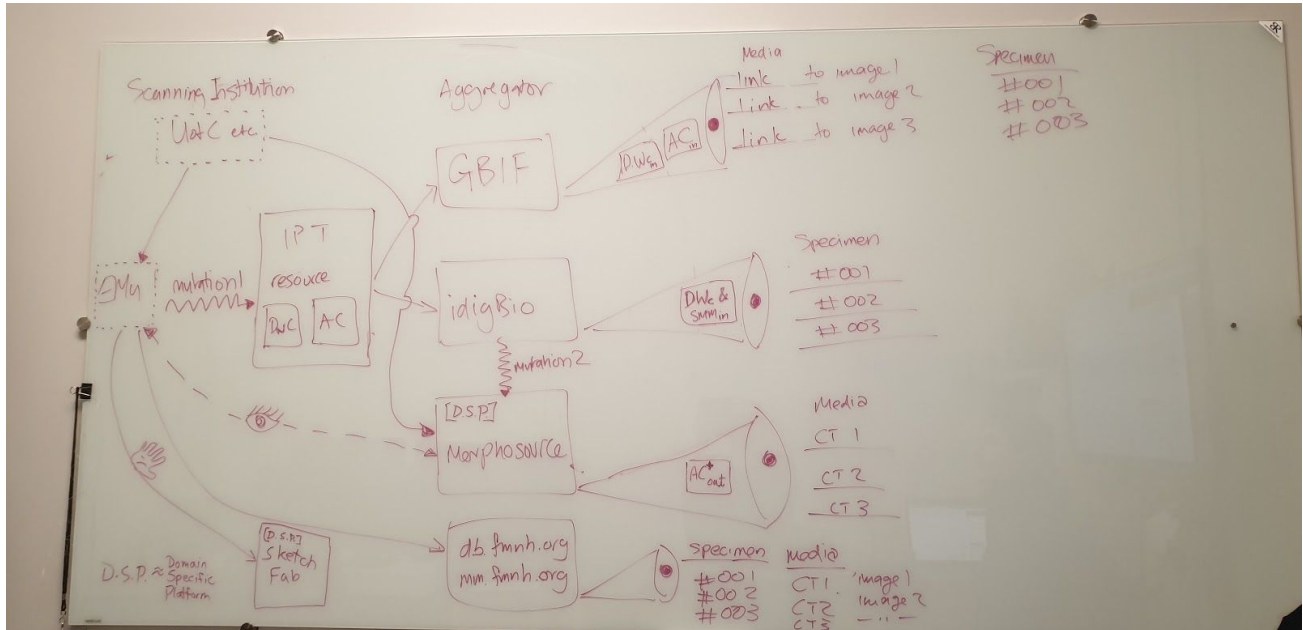
- b) Update the Event record
  - i) Make sure that the catalogue record is attached to the event.
  - ii) Add a note "CT scanned" to indicate that the MM records for that scan are all uploaded to EMu..



## 9) Repatriate the multimedia Identifier (UUID) to Morphosource

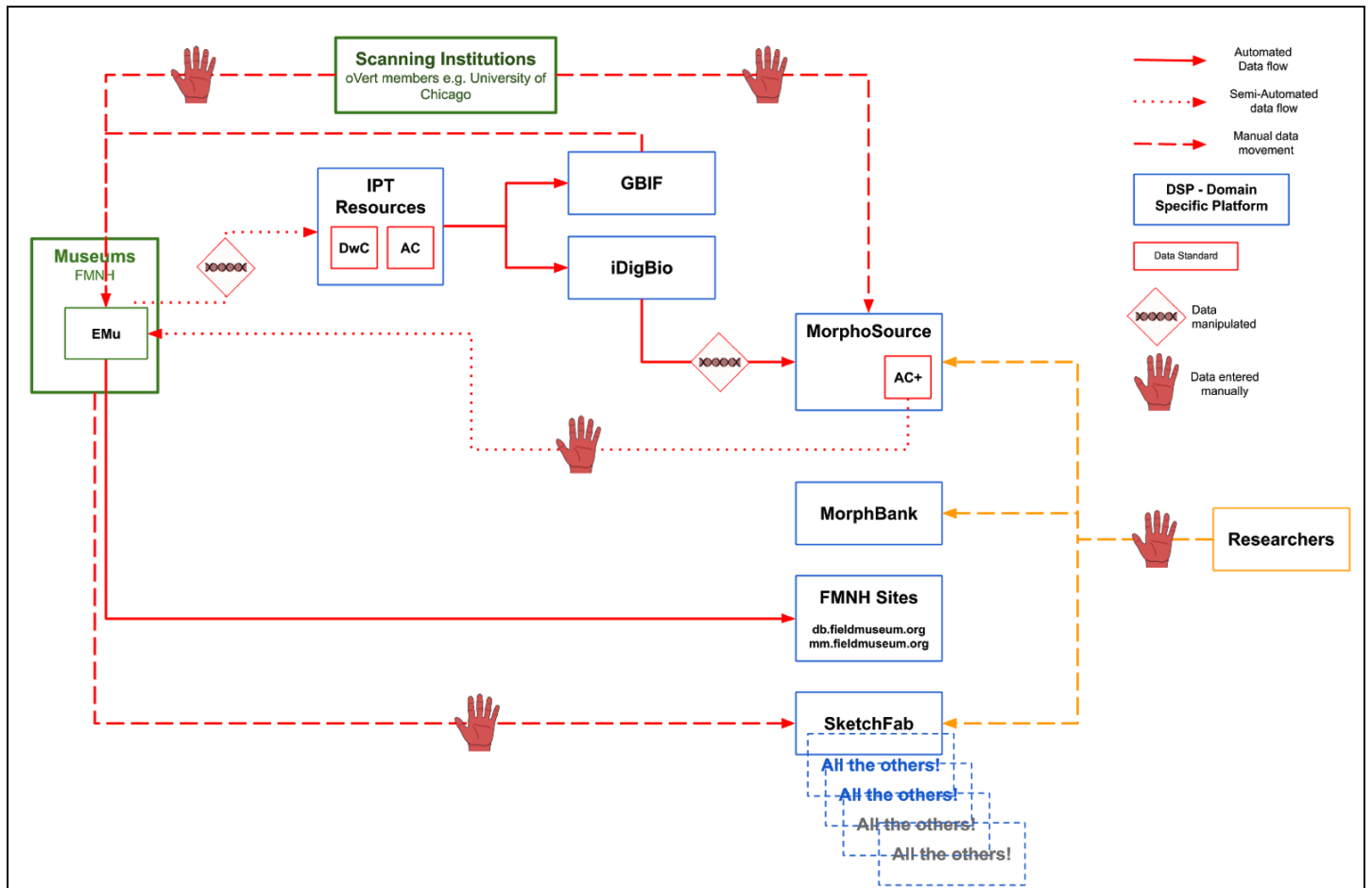
# Data Publication

<https://www.gbif.org/terms/data-publisher>



LINKED: Where does all the 3D data go?

<https://docs.google.com/drawings/d/136RSICSy1wnJk9KOH7DZixS85tL2AIXHI502OcrBv-4>



## End of project tasks

- Download Audubon Core archive for project from Morphosource, upload it to Multimedia in EMu, and attach it to Event irn 6662.
- Upload project documents as multimedia attachments (PDFs of Final project proposal, MOUs, Agreements, Releases, etc)
- Validate/Compare checksums between MorphoSource- & EMu-copies of scans.
- Delete scan file copies from Serenity after importing to EMu. (Give 2 weeks post-importing to EMu to allow time for checking files & records.)
- Add a "Y" in the [EMu column of this sheet](#) to indicate to UofC that scans for a given specimen have been transferred to EMu

## Notes

In MorphoSource

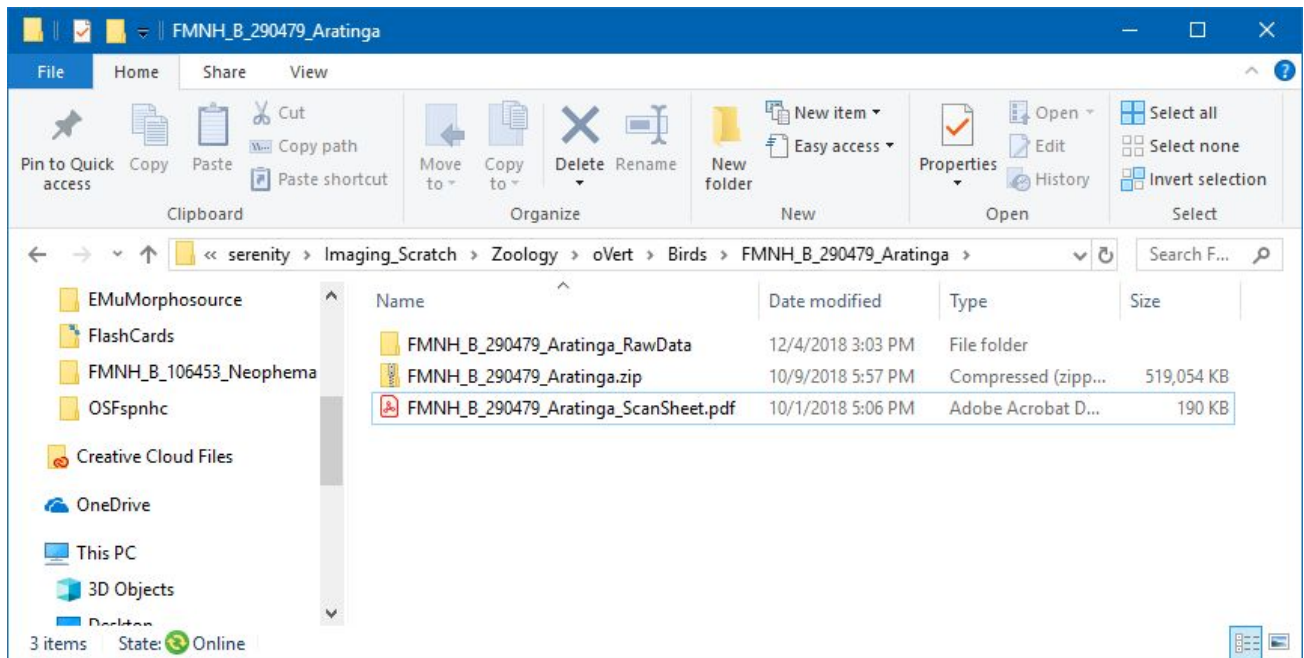
- Add Identifier to Stack record

- 11) Set to Public
- 12) In EMu
- 13) Make Bibliography record for the publication of the stack through MorphoSource
  - a) Include the ARK ID eg [ark:/87602/m4/M22152](https://n2t.org/ark:/87602/m4/M22152)

#### 1-April-2019 (kwebbink)

- EMu Multimedia records created on & before 1-April-2019 for CT Scan Image Stacks on MorphoSource contain the zip file for an Image Stack retrieved directly from MorphoSource. This was done to standardize and cross-reference the exact datasets stored on MorphoSource.
- Going forward, Image Stacks stored in EMu will be the zip files created by scanners at UiofC (April/Stephanie/Daryl), which will likewise be identical to those uploaded to MorphoSource

- Files to transfer from UofC to FMNH:



- 1 = Raw Data directory (can also include TAR)
- 2 = Zipped Image Stack
- 3 = ScanSheet PDF

9-March-2020

- Filenaming conventions - <http://intranet.fieldmuseum.org/emu/node/7069>
- MorphoSource naming convention:

[Institution code]\_[Collection Code]\_[Catalog #]\_[Genus]

e.g., "FMNH\_M\_12345\_Genus.zip"

## Accounting for lead in scans

- Might be worth putting this into the EMu overt event as a dataset reference doc. Check with April (is there a publication/rights/usage of diagram below etc.)

from: April Isch Neander (Basecamp) <notifications@3.basecamp.com>  
reply-to: april-isch-neander-usJHA2Yt6XqV@replies.3.basecamp.com  
to: sgrant@fieldmuseum.org  
date: 15 Apr 2019, 11:45  
subject: (Scanning Institutions) Lead in Specimens

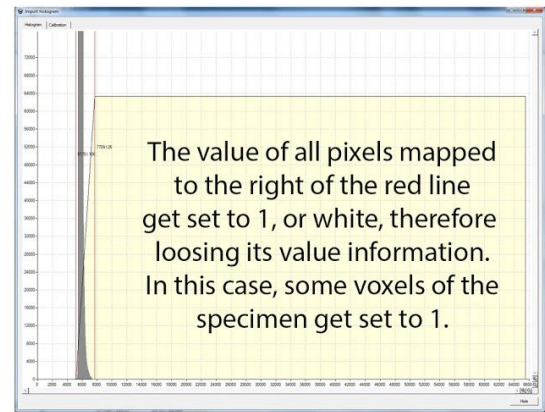
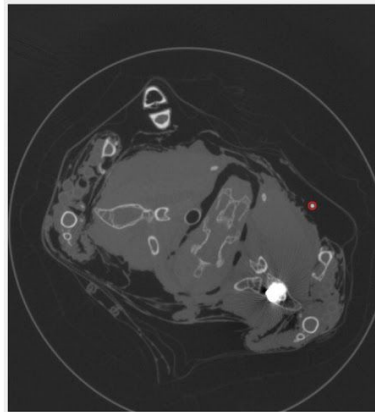
Hello All,

We discussed how to deal with lead in mostly bird specimens during the last call. Here is an image to describe why I think that when lead is present, you shouldn't crop the histogram or "Define Material" to the aluminum. Rather, you maintain more contrast of the area around the lead shot if you crop this histogram less or "Define Material" to the area just outside the lead shot.

Best,  
April

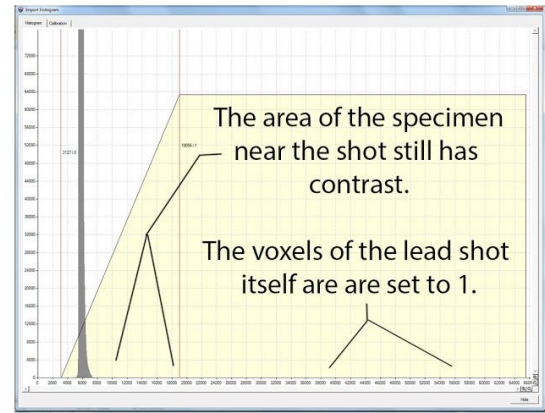
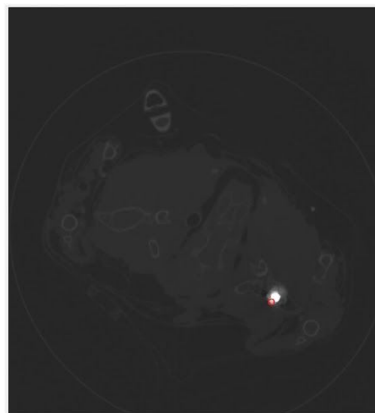
Aluminum selected for  
Define Material  
(see red circle)

Edges of the shot are  
poorly defined, and much  
of the contrast of the  
specimen near the shot  
is obliterated by the  
voxels all being set to 1.



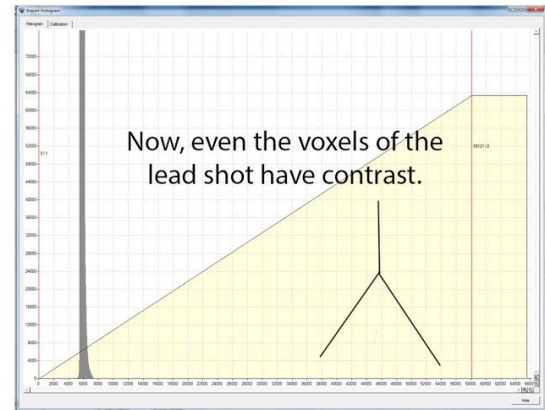
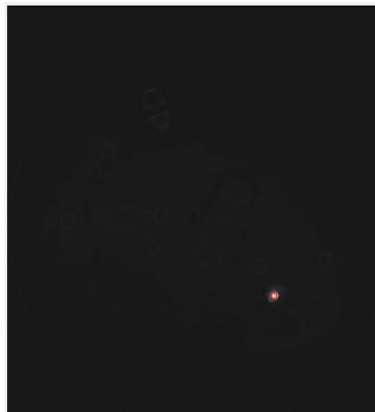
Bright area just outside  
lead shot selected for  
Define Material

I feel this is the best  
compromise between  
defining the shot so it  
can be more easily  
segmented out, while  
maintaining good enough  
contrast in the specimen.  
The entire shot is set to 1,  
but none of the surrounding  
area is.



Inner portion  
of lead shot selected  
for Define Material

The histogram now includes  
the lead shot, narrowing  
the range of contrast of the  
specimen. However, the user  
can still adjust the histogram  
during segmentation. Less  
information is lost, but it is also  
less user friendly.



13-May-2019 (kwebbink)

- FMNH-specimen scanning on hold until August 2019 at earliest (to scan Kansas specimens on loan)

## 1-Nov-2019 (meeting notes)

- 15-20 scans into EMu / week / volunteer
  - Keeps up with ~1 - 25 new CT scans per week
- Backlog - needs people / cross-checking with full list of FMNH specimens on MorphoSource
  - Training w/ media agreement for FMNH collections mgrs
  - 1000s of media assets to transfer to EMu
- To help track what's in + published:
  - [Collections-zoology site / search for "overt"](#)
  - <https://mm.fieldmuseum.org/search/overt?start=0>