



UNIVERSITY OF  
MICHIGAN

# Developing an Institutional CT Data Repository: An Example from Michigan



Cody Thompson, PhD - Mammal Collections Manager, UMMZ  
Susan Borda - Data Workflows Specialist, UM Library

# Outline

1. Project background - Cody
2. DCN case study with UMMZ/oVert - Susan
3. Deposit structure for CT data and manifests - Susan

# Project Background

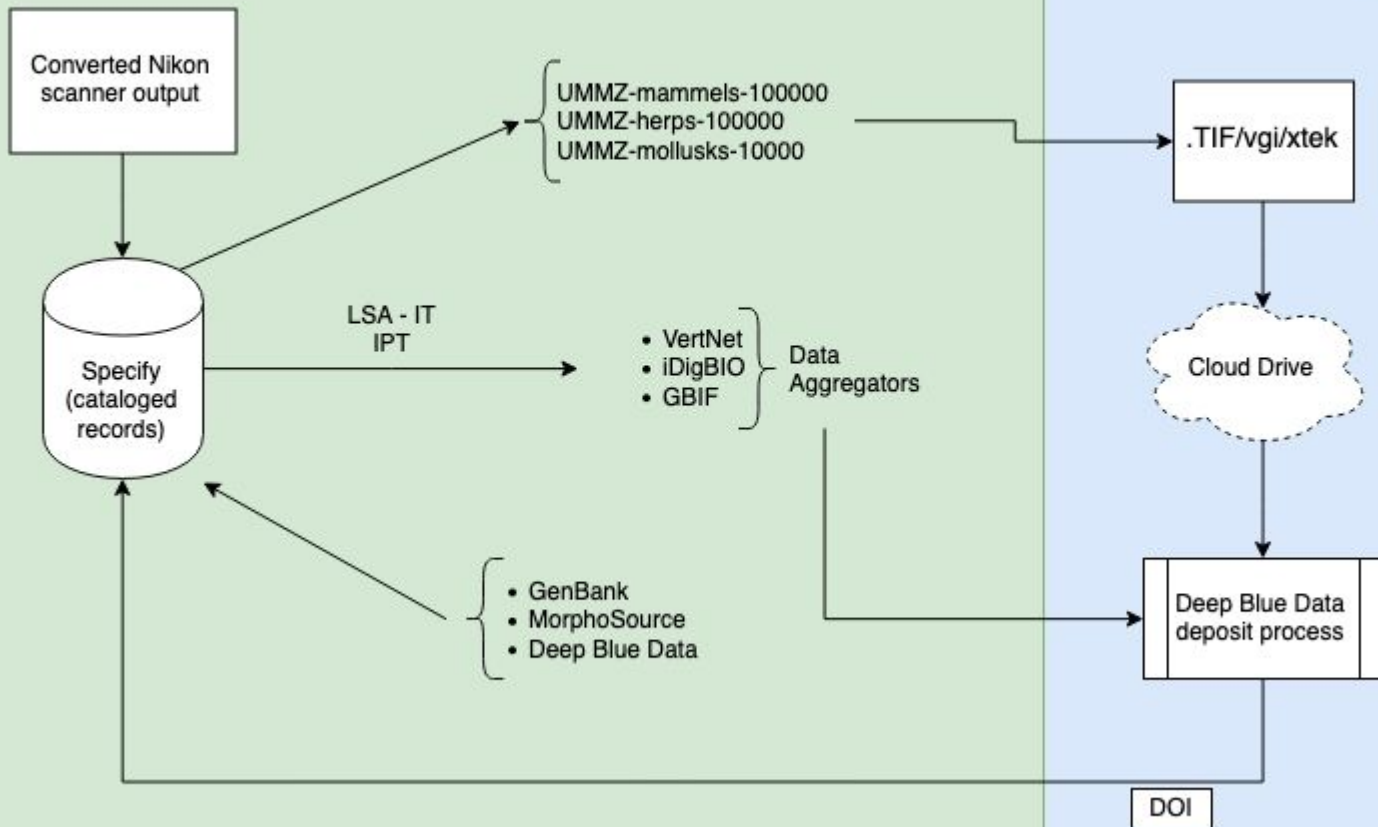
- In 2016, UMMZ became involved with oVert project.
  - Potentially producing 4,000-5,000 scans
  - The UMMP already had begun doing 3D photogrammetry, and had developed UMORF
- During oVert proposal preparation process, we ran into problems identifying long-term storage solution for CT data.
  - No college support beyond the length of the grant.
  - Proposed solutions were not cost effective.
- Contacted UM science librarian for assistance
  - Integrated library into data management plan

# Project Background (cont.)

- In Summer 2018, UMMZ reached out to UM Library about their data needs.
  - Deep Blue and Deep Blue Data serve as UM repositories.
- Library and UMMZ created a workflow for transferring the data from UMMZ to the UM Library's data repository.
- This workflow involved copying the data, re-organizing it, pulling descriptive metadata from GBIF, creating a yaml file for ingest, ingesting the data, and returning DOI information.

UMMZ

DBRRDS/  
Library



# Deep Blue Data

Data Repository at the University of Michigan Library

Started in Fall of 2016, based on Fedora Hyrax (Samvera)

Used by researchers (faculty and graduate students) from all areas of U-M (Humanities → STEM)

Deposits are not restricted by type/format -- geared toward machine readable files

File sizes range from KBs → 3+TBs

# Data Model Iterations

Everything in one deposit:

\*\* 1 DOI

UMMZ-herps-246849:

Skull - Raw.zip

WholeBody-Raw.zip

Skull-Recon.zip

WholeBody-Recon.zip

Each processing stage and ROI in it's own deposit:

\*\*4 DOIs

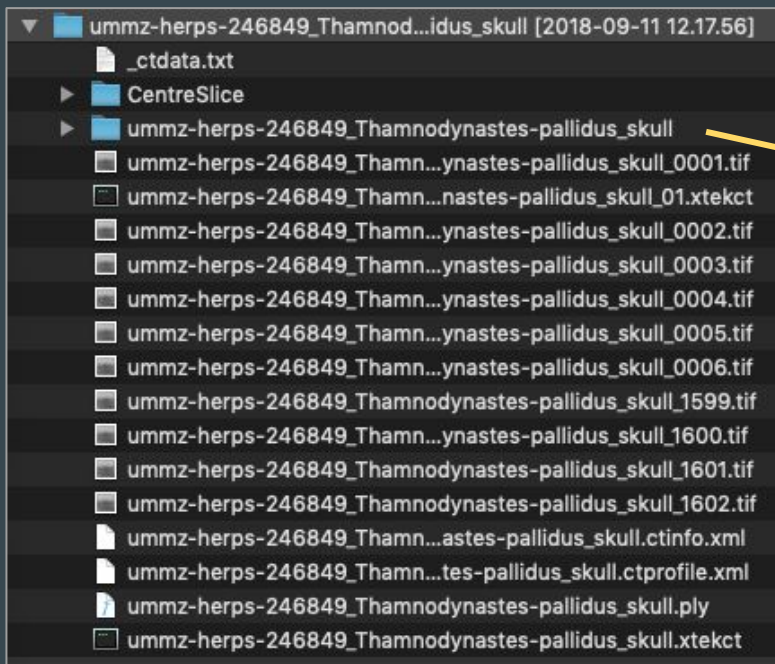
UMMZ-herps-skull-246849-Raw  
Raw.zip

UMMZ-herps-skull-246849-Recon  
Recon.zip

UMMZ-herps-body-246849-Raw  
Raw.zip

UMMZ-herps-body-246849-Raw  
Recon.zip






# Data - Before and After Re-org





# Data Model - Final

Files (Count: 5; Size: 12.4 GB)

	Title	Original Upload	Last Modified	File Size	Access	Actions
	<a href="#">Raw-Skull-ummz-herps-246849_Tham...7.zip</a>	2019-11-21		10.8 GB	<a href="#">Open Access</a>	Select an action ▾
	<a href="#">ummz-herps-246849_Thamnodynastes...l.ply</a>	2019-11-21		91.7 MB	<a href="#">Open Access</a>	Select an action ▾
	<a href="#">ummz-herps-246849_Thamnodynastes...0.tif</a>	2019-11-21		7.63 MB	<a href="#">Open Access</a>	Select an action ▾
	<a href="#">Recon-Skull-ummz-herps-246849_Th...7.zip</a>	2019-11-21		1.49 GB	<a href="#">Open Access</a>	Select an action ▾
	<a href="#">ummz-herps-246849_Thamnodynastes...0.tif</a>	2019-11-21		2.47 MB	<a href="#">Open Access</a>	Select an action ▾

# Metadata

## From xtek files:

Description	<p>Scan of specimen ummz:herps:246849 (THAMNODYNASTES PALLIDUS) - Skull. Raw - Dataset includes 1601 TIF images (each 1416 x 914 x 1 voxel at 0.0150099083890765 mm resolution, derived from 1601 scan projections), xtek and vgi files for volume reconstruction.</p> <p>Scan of specimen ummz:herps:246849 (THAMNODYNASTES PALLIDUS) - Skull. Reconstructed - Dataset includes 654 TIF images (each 1416 x 914 x 1 voxel at 0.015010 mm resolution, derived from 1601 scan projections), xtek and vgi files for volume reconstruction.</p>
-------------	--

## From GBIF:

Keyword	<a href="#">Animalia</a> <a href="#">Chordata</a> <a href="#">Reptilia</a> <a href="#">OPHIDIA</a> <a href="#">COLUBRIDAE</a> <a href="#">THAMNODYNASTES PALLIDUS</a> <a href="#">1987213257</a> <a href="#">computed tomography</a> <a href="#">X-ray</a> <a href="#">3D</a>
Date coverage	2018-09-11
Citations to related material	For more information on the original UMMZ specimen, see: <a href="https://www.gbif.org/occurrence/1987213257">https://www.gbif.org/occurrence/1987213257</a>

# Close the Loop

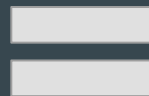
MorphoSource metadata:

```
file_name,ms_id,media_type,ingestable,  
parent_file_name,parent_ms_id,parent_identifier  
,occurrence_id,institution_code,collection_code  
,catalog_number,scandate,device_model,  
device_manufacturer,device_modality,  
device_description,device_organization_name,  
x_spacing,y_spacing,z_spacing,unit,  
number_of_images_in_set,  
processing_activity_type,creator,part,license,  
media_ct_series_type,  
media_ct_number_of_images_in_set
```



UMMZ deposit info:

```
ID,Title,DOI,Create  
Date,Modified Date,URL,File  
Name,File Set ID
```



```
MorphoSource file manifest
```

# Future Plans

The Library would like to use this as a model for working with the other museums (Paleontology for instance) and other similar entities as well.

# Questions?

Cody Thompson - [cwthomp@umich.edu](mailto:cwthomp@umich.edu)

Susan Borda - [sborda@umich.edu](mailto:sborda@umich.edu)

Code - <https://github.com/mutanthumb/ummzgbif>

Deep Blue Data repository - <https://deepblue.lib.umich.edu/data/>