

# **Making the SHiFt:**

## **Using Sufia with Hydra/Fedora for collection management and access**

James Halliday

Programmer/Analyst, Library Technologies

Juliet L. Hardesty

Metadata Analyst, Metadata Resources & Systems

Indiana University Libraries

OR2015, June 11, 2015

# About IU Libraries Repo System

- Fedora 3 – digital special collections, digitized books
  - 2,000,000+ objects
  - Various custom sites and services sites
- Dspace – Institutional Repository (IR), document-centric
  - IU Scholarworks (IUSW), Digital Library of the Commons, Archives of Institutional Memory
  - IUSW – 11,000+ items
- No Fedora 4 prod (yet)

# CBRC

- Center for Biological Research Collections
- Specimen collections
  - Herbarium – 2D images of plants
  - Zooarchaeology – 3D images of bones
  - 10's of K's of items
- Specify Software
  - Specimen management and cataloging
  - Darwin Core, geospatial info

# CBRC Needs

- Cataloging and inventory management covered
- No place to store/preserve images
- No way to provide public access to digitized specimens

# Born-digital docs on finding aids

- Spring 2014 survey
  - 4 out of 12 repositories responded
  - Almost 500,000 born-digital documents (not including images)
  - PDF, Word and other text files, video, audio, Excel, PPT, ZIP and other

# Born-digital docs on finding aids - Needs

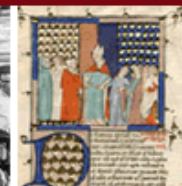
- No workflow to ingest non-image items into Fedora
- No way to offer search and discovery or access to non-image items

# Sufia 6/Fedora 4

- Problems solved! (JK)
- Workflow for multiple file types
- Chance to try Fedora 4 (aka, THE FUTURE)
- Possibility of system mostly out of the box that lets collection managers actually manage collections in Fedora

# Archives Online

at Indiana University




[Home](#) | [Browse](#) | [Search](#) | [Advanced Search](#) | [Help](#)

Search Archives Online

Search

## View options:

- [Standard](#)
- [Entire document](#)
- [Printer-friendly](#)
- [XML](#)
-  [Items online](#)

## Search within this document:

go

## Contents:

### Summary Information

[Administrative Information](#)

[Administrative History](#)


[Scope and Content Note](#)

[Arrangement](#)

[Indexing Terms](#)

[Separated Material](#)

 [Content List](#)

 [Administration, 1960-1992](#)





[Annual Reports, 1960-1973](#)

["Archives," 1960-1972](#)

[Course Offerings 1965-1988](#)

[Governance 1978-1989](#)

## Indiana University African Studies Program Directors' records, 1957-1995

 BOOKMARK   

### Summary Information

#### Repository

Indiana University Office of University Archives and Records Management  
1320 East Tenth Street  
Herman B Wells Library E460  
Bloomington, IN 47405-7000  
Phone: 812-855-1127  
[Email: archives@indiana.edu](mailto:archives@indiana.edu)  
<http://www.libraries.iub.edu/archives>

#### Creator

Indiana University, Bloomington. African Studies Program.

#### Extent

9.3 cubic feet

#### Language

Materials are in English unless otherwise indicated.

#### Title

Indiana University African Studies Program Directors' records, 1957-1995

#### Collection no

C217

#### Abstract

The collection consists of the administrative records and correspondence of directors J. Gus Liebenow and Patrick O'Meara related to the development of the African Studies Program.



# University Archives Sufia Test - Results

- Awesome
  - Self-deposit
  - Permissions controls
- Not Awesome
  - No hierarchy (collection of objects that contain objects that have multiple items)
  - No view, only downloadable
  - Separate Sufia instance needed for each repository?
- Sufia doesn't match needs yet

# Sample Item – Canis familiaris mandible

- Sample item from one CBRC collection
- 3D images created using FARO ScanArm equipment
- Each item will be digitized and stored in Sufia / Fedora along with associated metadata
- Metadata will also be stored in Specify, an open-source Java-based platform for storing information about biological research collections (see <http://specifyx.specifysoftware.org/>)

# Darwin Core metadata

- All CBRC collections use the Darwin Core metadata standard
- Extension of Dublin Core to add metadata specific to biological collections
- See <http://rs.tdwg.org/dwc/>


# IU CBRC Repository

1jhallida@indiana.edu

[Home](#)[About](#)[Help](#)


Enter search terms Go All

[Back to search results](#)



## R38\_mandible.obj Institution Name

### Descriptions

<b>Resource type:</b>		<b>State Province:</b>	Nunavut
<b>Creator:</b>	garymotz@indiana.edu	<b>Locality:</b>	South Hampton Island
<b>Keyword:</b>	3D	<b>Decimal Latitude:</b>	
<b>Rights:</b>	Attribution-NonCommercial-ShareAlike 3.0 United States 	<b>Decimal Longitude:</b>	
		<b>Scientific Name:</b>	Canis familiaris
		<b>Scientific Name Aut...</b>	
<b>Identifier:</b>		<b>Kingdom:</b>	Animalia
<b>Collection Code:</b>	WRAZL	<b>Phylum:</b>	Chordata
<b>Catalog Number:</b>	1900-10-0840	<b>Class:</b>	Mammalia
<b>Other Catalog Num...</b>	R38	<b>Order:</b>	Carnivora
<b>Continent:</b>	North America	<b>Family:</b>	Canidae
<b>Country:</b>	United States	<b>Genus:</b>	Canis
<b>County:</b>	Kivaliq County	<b>Specific Epithet:</b>	Canis familiaris
<b>Higher Geography:</b>	Kivaliq, Nunavut, Canada	<b>Infraspecific Epithet:</b>	
		<b>Type Status:</b>	

### Actions

[Download](#)

[Export to: EndNote](#) | [Zotero](#) | [Mendeley](#)

### Collections

This file is in the following collections:

3D Collection Objects

## Sample Item in Fedora

ns002: **title**  
R38\_mandible.pdf

ns003: **isPartOf**  
<http://poplar.dlib.indiana.edu:8260/fedora4/rest/dev/b5/64/4r/75/b5644r755>

ns004: **dpt**  
garymotz@indiana.edu

ns006: **relativePath**

ns007: **catalogNumber**  
1900-10-0840

ns007: **class**  
Mammalia

ns007: **collectionCode**  
WRAZL

ns007: **continent**  
North America

ns007: **country**  
United States

ns007: **county**  
Kivaliq County

ns007: **family**  
Canidae

ns007: **genus**  
Canis

ns007: **higherGeography**  
Kivaliq, Nunavut, Canada

ns007: **kingdom**  
Animalia

ns007: **locality**  
South Hampton Island

ns007: **order**  
Carnivora

ns007: **otherCatalogNumbers**  
R38

ns007: **phylum**  
Chordata

ns007: **scientificName**  
Canis familiaris

## Update Properties

<<http://www.w3.org/ns/auth/ac#>>

PREFIX ns007:  
<<http://rs.tdwg.org/dwc/terms/>>

PREFIX dc:  
<<http://purl.org/dc/elements/1.1/>>

PREFIX foaf:  
<<http://xmlns.com/foaf/0.1/>>

PREFIX fedoraconfig:  
<<http://fedora.info/definitions/v4/config#>>

PREFIX image:

Update

# Sample Item in Specify

Specify 6.6.01

File Edit Data System Tabs Help

Welcome Data Trees Reports Interactions Statistics Query Symbiota Workbench SGR Plugins Lifemapper Attachments

Create/Update

Collection Object

Taxon

Agent

Geography

Locality

More Forms

Record Sets

Beth's photos

Bovidae

Canidae

Deer photo record

Husky Kelsey loan

IUPII

IUPII 2

Jesualdo

Jesualdo Scans

NORTH POLE

Polly 2011 Loan

Rowe loan 01-12-2014

SAA Femur

Sufia 3D export

Tuttle Furs locality

WRAZL Primates

Zoarch Scans

Collection Object

Catalog Number: 1900-10-0840

Original Cat Number: R38

Accession: 1900-ZL-000

Object Condition:

Specimen Description:

Surface Features:

Remarks: Skull and mandible

Determinations

Common Name: DOMESTIC DOG

Taxon: Canis familiaris

Preferred Taxon: Canis familiaris

Determined Date:

Determiner:

Type: None

Method:

Collecting Information

Context:

Locality: South Hampton Island, Kivaliq County, Nunavut, Canada

Collectors

Last Name	First Name	Remarks
-----------	------------	---------

Preparations

Prep Type	Count	Is on Loan	Preparation Attachments	Description	Solution	Storage
Skull	1	Yes		NH40H		7D, IUZL, Stud...

Col Obj Attribute

Sex:

Sex Basis:

Cause of Death:

Age:

Age Basis:

Date Born:

Specimen Length:

Weight:

Tissues Taken:

Specimen Source: W. R. Adams

Attachments

Modified By Agent: Motz, Gary

Timestamp M...: 2015/02/10

Cataloged D...: 2010/12/03

1 of 2

Edit

Trash

Zoarchaeology | William R. Adams Zooarchaeology Collection | garymotz

[illegible]

# Specify and Sufia work together

- All metadata stored in both Specify and Sufia
- Specify provides data entry and analysis tools that would be hard to replicate
- Data synchronization is an issue
- Sufia serves as the primary archival location for both data files and metadata



# Darwin Core metadata and the RDF-Vocab project

- At our request, Darwin Core metadata was added to the available vocabularies within the RDF-vocab project
- See <https://github.com/ruby-rdf/rdf-vocab>
- This enables us to easily reference DWC metadata fields from within the application and link them appropriately

# Sufia fields to DWC metadata code snippet

```
property :order, predicate: ::RDF::Vocab::DWC.order do |index|  
  index.as :stored_searchable, :facettable  
end  
property :family, predicate: ::RDF::Vocab::DWC.family do |index|  
  index.as :stored_searchable, :facettable  
end  
property :genus, predicate: ::RDF::Vocab::DWC.genus do |index|  
  index.as :stored_searchable, :facettable  
end
```

# Future work – 3D viewer integration

- 3D scans can currently be viewed using solutions such as SketchFab – see <https://skfb.ly/CUqJ>
- We hope to instead integrate a viewer solution with Sufia using technology being developed at IU's Advanced Visualization Lab

# Future work – batch ingest and workflow

- We need automated scripts between the digitizers workstations, Sufia and Specify
- Digitizers will enter metadata directly into Specify, and drop files for deposit into Sufia.
- We will leverage existing infrastructure for image processing and backup as much as possible

# Future work – need for hierarchy

- We will soon need the ability to have multiple levels of hierarchy within the repository
- 3D imaging produces multiple files per object

# Future work – generalizing the repository

- For now, our Sufia repository is specific to these biology-related collections
- It also serves as a pilot project for Fedora 4
- Eventually, we might use a single Sufia instance to hold a wider variety of content
- Each collection might have different metadata needs
- Editing metadata fields through the interface, and not within code would be desirable

# Thanks! Question? Comments?

- Github repo

<https://github.com/IUBLibTech/cbrc>

- Jim Halliday

[jhallida@indiana.edu](mailto:jhallida@indiana.edu)

- Julie Hardesty

[jlhardes@iu.edu](mailto:jlhardes@iu.edu)