Boundless Use of Indiana University's Biological Research Collections Possible in Partnership with IU Libraries

Jennifer Laherty & Gary J. Motz
Creating Opportunities for Specimen Collections

• Libraries and academic units in partnership
  • Co-curators of research data
This is important for “big data” research

Ellen Ketterson, PhD: IU Biology
Indiana University Herbarium

~150,000 plant specimens
Indiana University Herbarium
William R. Adams
Zooarchaeology Lab

• ~20 student volunteers per semester
• 1 graduate assistant
• FARO laser scanner
• Dedicated scanning coursework

>10,000 vertebrate skeletons
Indiana University Paleontology Collection

It’s not just 1.3 million fossils...
Indiana University Paleontology Collection

- ~4600 35mm slides
- ~1750 sheets of film
- 30 rolls of film
- ~2250 SEM negatives
- ~7300 photo prints
- ~750 lantern slides
- ~35 field notebooks

~160 years of research at IUB on half a billion years of the history of life on Earth
The Internet of Samples in the Earth Sciences (iSamples)

The iSamples RCN seeks to advance the use of innovative cyberinfrastructure to connect physical samples and sample collections across the Earth Sciences with digital data infrastructures to revolutionize their utility for science. The ultimate goal of this RCN is to dramatically improve the discovery, access, sharing, analysis, and curation of physical samples and the data generated by their study for the benefit of science and society.
Technology Innovation

- Fedora 4, Sufia 7, Specify
- Systematic digitization
- Application of metadata
- Repository of ANY digital object
- Built in Sufia (a Ruby on Rails engine)
- Powered by Hydra and Fedora 4

- Jim Halliday, Julie Hardesty, Randall Embry
Sufia Digital Content Repository

### R38_mandible.obj

#### Descriptions

<table>
<thead>
<tr>
<th>Resource type:</th>
<th>Creator: <a href="mailto:garymetz@indiana.edu">garymetz@indiana.edu</a></th>
</tr>
</thead>
<tbody>
<tr>
<td>Keyword: 3D</td>
<td></td>
</tr>
<tr>
<td>Rights: Attribution-NonCommercial-ShareAlike 3.0 United States (C)</td>
<td></td>
</tr>
<tr>
<td>State Province:</td>
<td>Nunavut</td>
</tr>
<tr>
<td>Locality:</td>
<td>South Hampton Island</td>
</tr>
<tr>
<td>Decimal Latitude:</td>
<td></td>
</tr>
<tr>
<td>Decimal Longitude:</td>
<td></td>
</tr>
<tr>
<td>Scientific Name:</td>
<td>Canis familiaris</td>
</tr>
<tr>
<td>Scientific Name Authority:</td>
<td></td>
</tr>
<tr>
<td>Kingdom:</td>
<td>Animalia</td>
</tr>
<tr>
<td>Phylum:</td>
<td>Chordata</td>
</tr>
<tr>
<td>Class:</td>
<td>Mammalia</td>
</tr>
<tr>
<td>Order:</td>
<td>Carnivora</td>
</tr>
<tr>
<td>Family:</td>
<td>Canidae</td>
</tr>
<tr>
<td>Genus:</td>
<td>Canis</td>
</tr>
<tr>
<td>Specific Epithet:</td>
<td>Canis familiaris</td>
</tr>
</tbody>
</table>

#### Actions

- Download
- Export to: EndNote, Zotero, Mendeley

### Collections

This file is in the following collections:

- 3D Collection Objects
Specify collections management

- mySQL application with Java frontend GUI
- Linked data using Darwin Core metadata schema
Repositories working together

**Indiana University Herbarium**
- Founded in 1885
- Houses 150,000 plant specimens
- Official botanical repository for the state of Indiana
- [http://go.iu.edu/1jCd](http://go.iu.edu/1jCd)

**Scholarly Data Archive**
- Distributed storage
- ~42 PB of storage
- [https://kb.iu.edu/d/aiyi](https://kb.iu.edu/d/aiyi)

**Consortium of Midwest Herbaria Symbiota Portal**
- Building a skeletal metadata architecture
- Enhancing metadata via optical character recognition (OCR)
  - We’re also using handwritten text recognition (HTR)
- Georeferencing of specimen localities using GEOLocate
- Crowdsourcing modules to engage citizen scientists
- [http://go.iu.edu/1jBr](http://go.iu.edu/1jBr)

**Sufia Digital Content Repository**
- Linked framework that preserves and makes digital content available to outside community
- Able to ingest ANY digital content
- Robust repository infrastructure that provides URI for objects
- [http://go.iu.edu/1jCe](http://go.iu.edu/1jCe)
Herbarium Digitization Workflow
A metadata standard for biodiversity informatics

A body of standards, including a glossary of terms (in other contexts these might be called properties, elements, fields, columns, attributes, or concepts) intended to facilitate the sharing of information about biological diversity by providing reference definitions, examples, and commentaries.

Sponsored by Biodiversity Information Standards (TWDG), the current standard was last modified in October 2009.

| Mappings                  | ABCD
|---------------------------|----
| Related Vocabularies     | Audubon Core, GBIF Vocabularies
| Specification            | http://www.tdwg.org/standards/450
| Standard’s website       | http://rs.tdwg.org/dwc/index.htm

**Extensions**

**Apple Core**
Darwin Core documentation and recommendations for herbaria.

**Darwin Core Geospatial Extension**
A protocol-independent XML schema for a geospatial extension to the Darwin Core.

**DWC Germplasm**
An extension to the Darwin Core standard, it includes additional terms required to describe plant genetic resources and in particular germplasm seed samples.

http://www.dcc.ac.uk/resources/metadata-standards/darwin-core
Making data and images of millions of biological specimens available on the web.
Digital Content

**Pipistrellus pipistrellus**
Modern bat
London, England

3D files are the product of many composite files and hours of work (by both computers and humans)
Preserving Digital Assets

Trust

Future

Create
On the horizon...

https://skfb.ly/FH9I

https://skfb.ly/FGRz
Thank you!

- Jennifer Laherty
  - Librarian, Head of Sciences
  - jlaherty@Indiana.edu

- Gary Motz
  - Paleontologist & Digitization Coordinator
  - garymotz@Indiana.edu