

#### Data In Libraries: The Big Picture

# 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







William R. Adams Zooarchaeology
Laboratory
(Anthropology)



Indiana University Herbarium (Biology)



IU Paleontology
Collection
(Geological Sciences)

### This is important for "big data" research







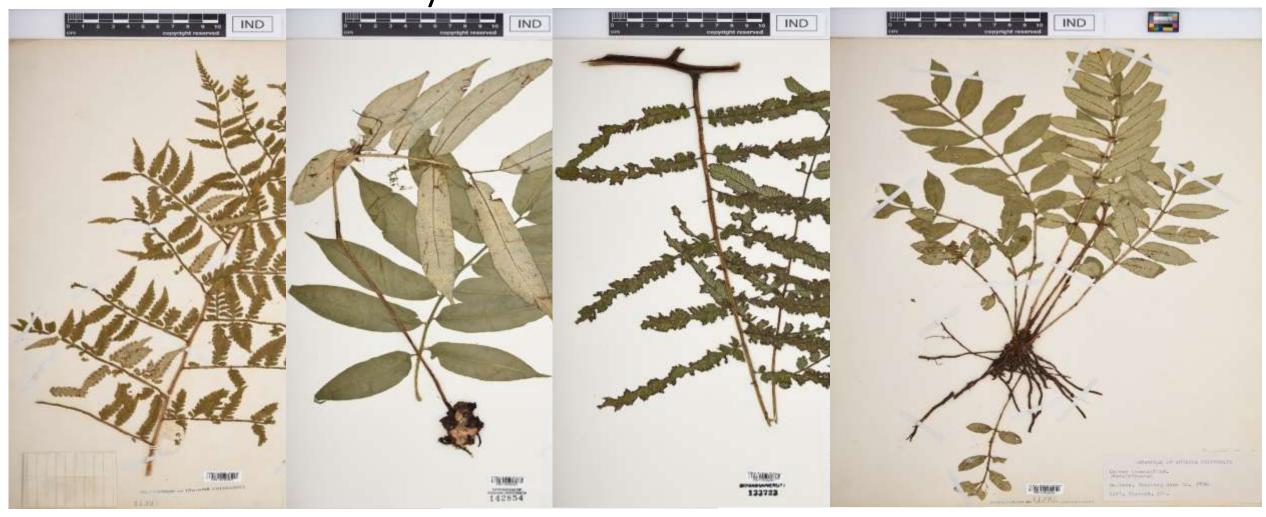


Ellen Ketterson, PhD: IU Biology

### Indiana University Herbarium



Indiana University Herbarium



### William R. Adams Zooarchaeology Lab



>10,000 vertebrate skeletons



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



Indiana University Paleontology Collection



### Indiana University Paleontology Collection



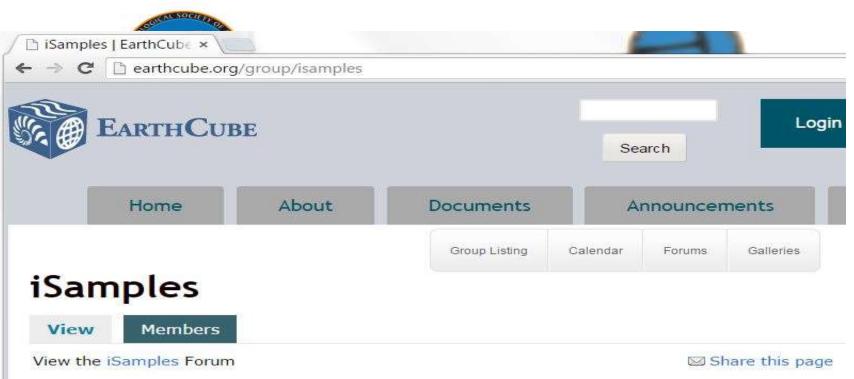
- ~4600 35mm slides
- ~1750 sheets of film
- 30 rolls of film
- ~2250 SEM negatives
- ~7300 photo prints
- ~750 lantern slides
- ~35 field notebooks



## We need to talk based geoscien

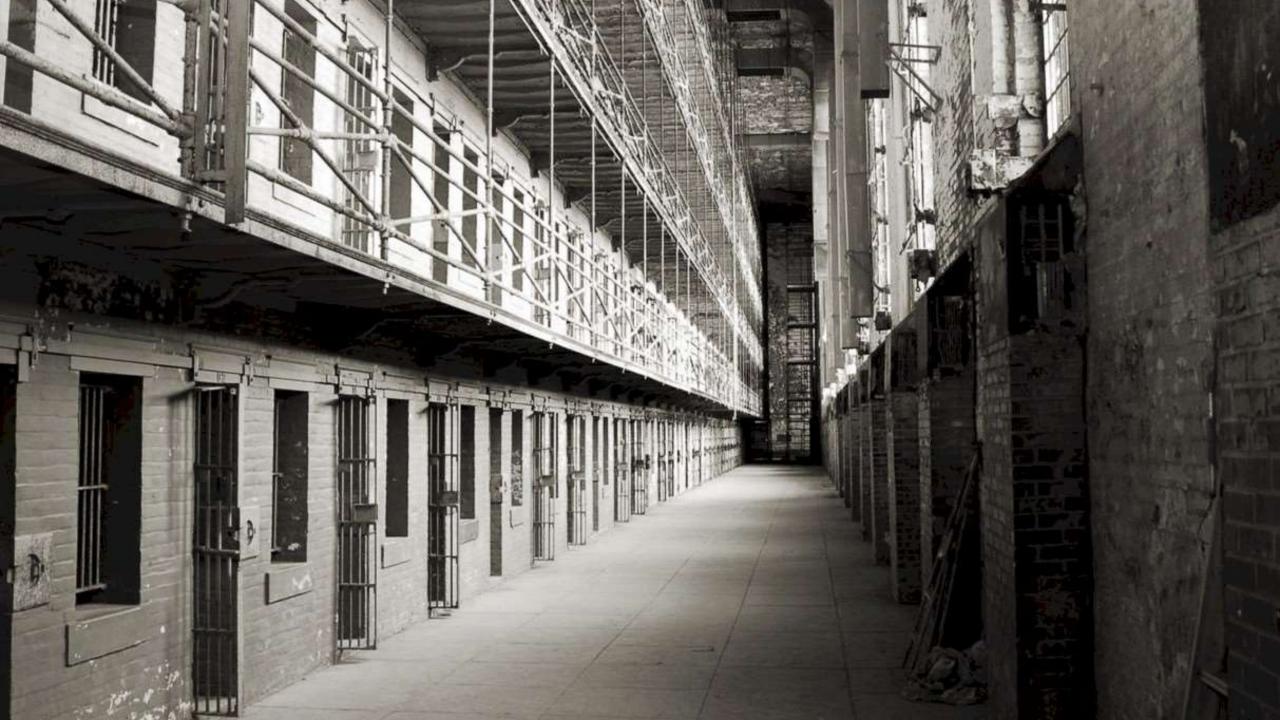
Matty Mookerjee\*, Daniel Vieira, University, Rohnert Park, Californi Dept. of Geology & Geophysics, T. City, Utah 84112, USA; Yolanda Gi. University of Southern California, USA; Charles Goodwin, Dept. of A California at Los Angeles, Los Ang Thomas F. Shipley, Dept. of Psych Philadelphia, Pennsylvania 19122, Geoscience, University of Wiscon Wisconsin 53706, USA

A UNIQUE OPPORTUNITY FOR



### 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

Biodiversity Information Standards





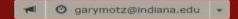




Promoting Bio-Collaboration







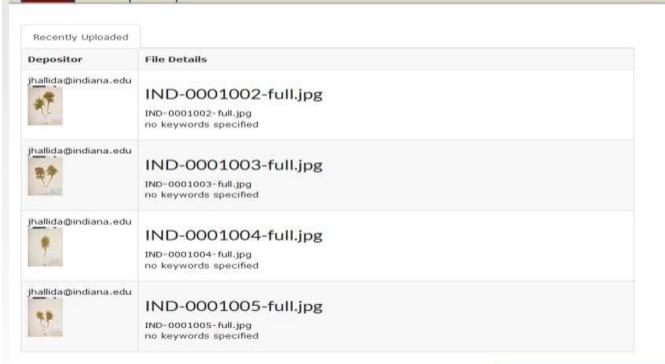
#### **IU CBRC Repository**

About

Help

Enter search terms





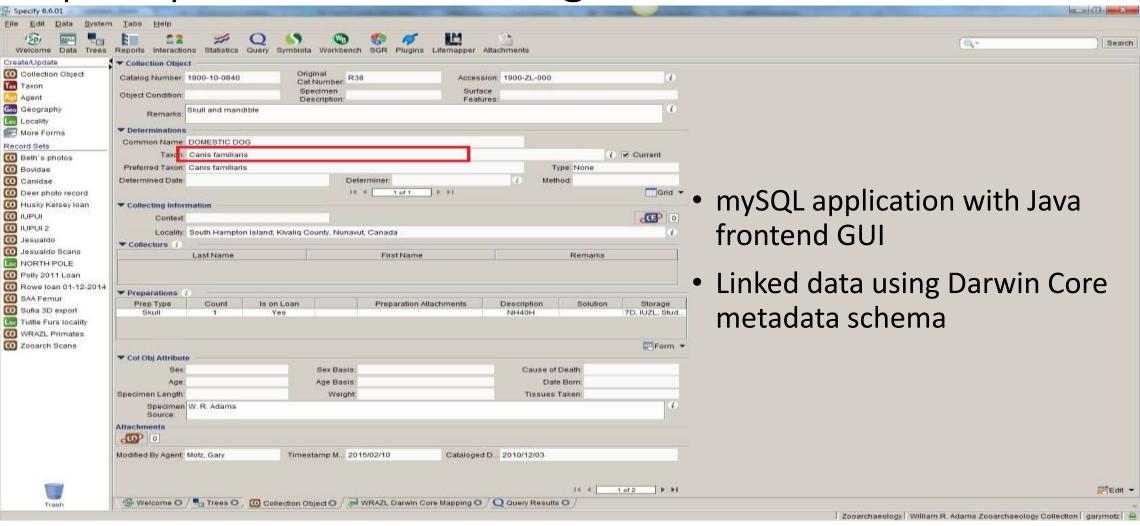
- Repository of ANY digital object
- Built in Sufia (a Ruby on Rails engine)
- Powered by Hydra and Fedora 4
- Jim Halliday, Julie Hardesty, Randall Embry

FULFILLING the PROMISE

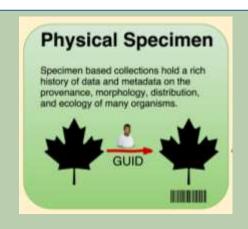
### Sufia Digital Content Repository



### Specify collections management



### 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

#### **Scholarly Data Archive**

- Distributed storage
- ~42 PB of storage
- https://kb.iu.edu/d/aiyi



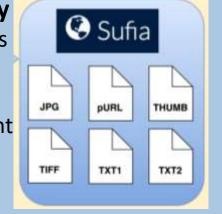
#### **Consortium of Midwest Herbaria Symbiota Portal**

- 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



#### **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

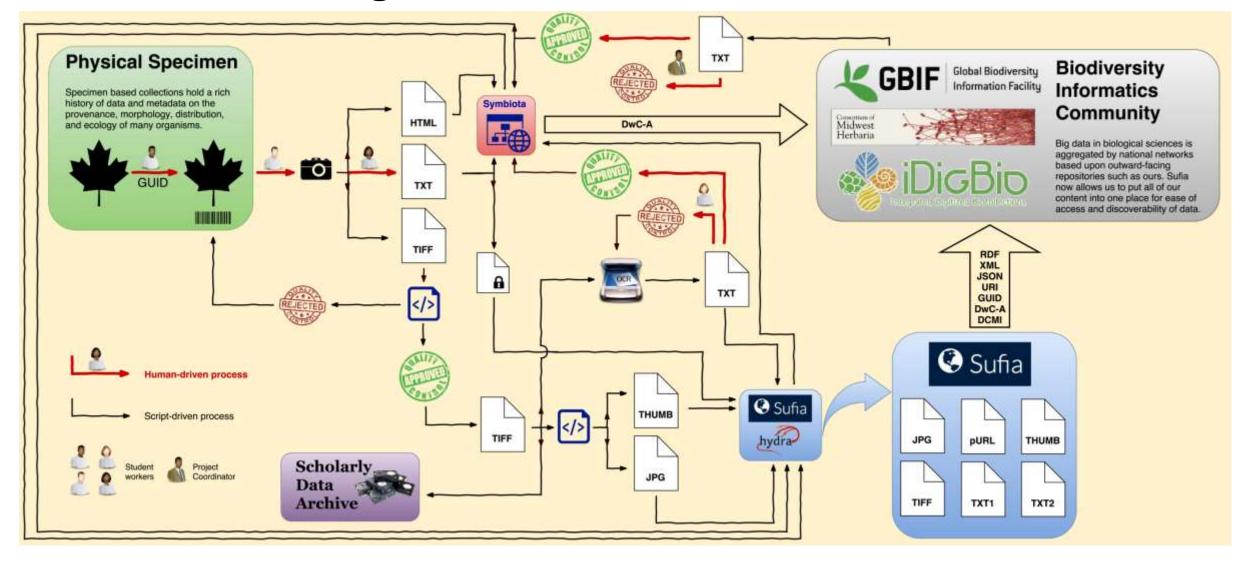




Scholarly.

Archive

### Herbarium Digitization Workflow



DCC because good research needs good data							Search		
Home	Digital curation	About us	News	Events	Resources	Training	Projects	Community	Tailored support

Home > Drupal > Resources > Metadata Standards > Darwin Core

#### Darwin Core 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.htmr				

#### **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



### Research Data Sharing without barriers







Biodiversity Information Standards



The Society For The Preservation of Natural History Collections

Making data and images of millions of biological specimens available on the web

60,409,876 Specimen Records

16,756,120 Media Records

732

Recordsets

Search the Portal



Why digitization matters

More about what we do and why

### Digital Content



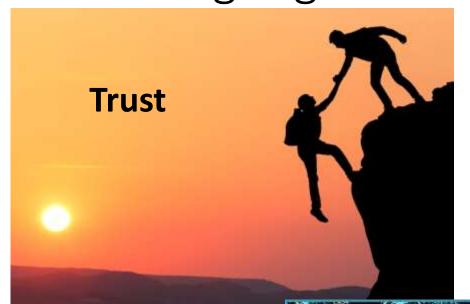
www.arkive.org/pipistrelle-bats/pipistrellus-pipistrellus-and-pipistrellus-pygmaeus/image-A18965.html



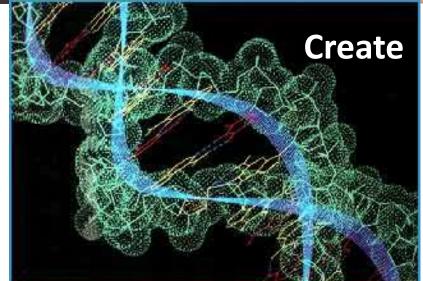


3D files are the product of many composite files and hours of work (by both computers and humans)

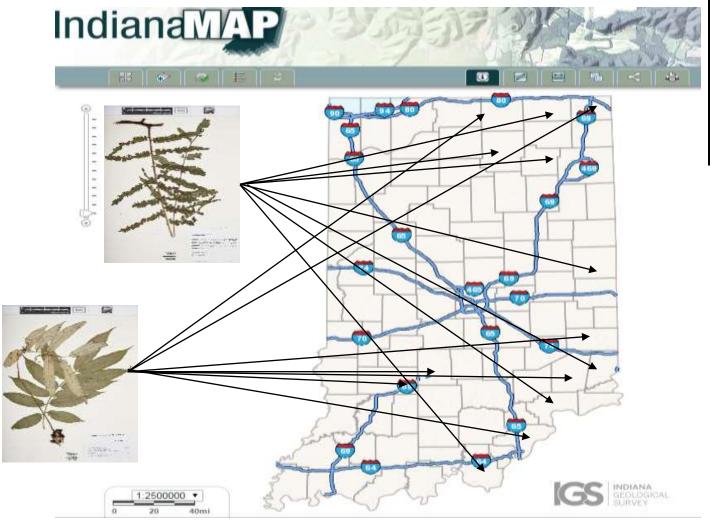
### Preserving Digital Assets

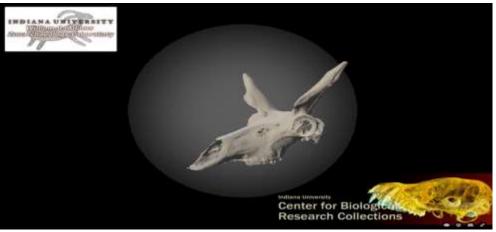






### On the horizon...





https://skfb.ly/FH9I



https://skfb.ly/FGRz

### UITS Research Technologies Cyberinfrastructure enabling research and creative activities

### Thank you!







- Jennifer Laherty
  - Librarian, Head of Sciences
  - jlaherty@Indiana.edu
- Gary Motz
  - Paleontologist & Digitization Coordinator
  - garymotz@Indiana.edu









#### INDIANA UNIVERSITY BLOOMINGTON

