iDigBio

Integrated Digitized Biocollections

Welcome!
A few logistical details

Wiki: (https://www.idigbio.org/wiki/index.php/Paleo_Imaging_Workshop)
Twitter: #palpix
Adobe Connect (Kevin Love): https://idigbio.adobeconnect.com/paleo
Being broadcast and recorded
Be observant of remote audience; use microphone to make comments, ask questions
Chat box for remote participants
Collaborative Documents Folder: http://tinyurl.com/palimg
Efficiency: Starting on time, staying on track
Meals: Lunch provided; breakfast and evening on your own; no receipts required for meals
Attendees: Editable private list available via private url disseminated at the workshop.

This material is based upon work supported by the National Science Foundation under Cooperative Agreement EF-1115210. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Integrated Digitized Biocollections (iDigBio) An Introduction

Gil Nelson
Institute for Digital Information and Scientific Communication
Integrated Digitized Biocollections
Florida State University

Paleo Imaging Workshop
29 April-1 May, 2014

This material is based upon work supported by the National Science Foundation under Cooperative Agreement EF-1115210. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

The U.S. National Science Foundation estimates there may be as many as 1.8 billion biological and paleontological specimens stored in U. S. museums and academic institutions (perhaps as many as 3 billion worldwide). But, no one really knows!

In an effort to make these collections universally accessible to taxonomists, ecologists, researchers, and the general public, in 2011 NSF launched a $100 million, 10-year Advancing Digitization of Biodiversity Collections program and named Florida State University and University of Florida jointly as the national resource for digitization.

Advancing Digitization of Biodiversity Collections

Integrated Digitized Biocollections (iDigBio)
University of Florida
Florida State University
Florida Museum of Natural History

The goal is to digitize and make available via the Web at least 1 billion biological and paleontological records over the 10-year life of the project.

iDigBio
Integrated Digitized Biocollections

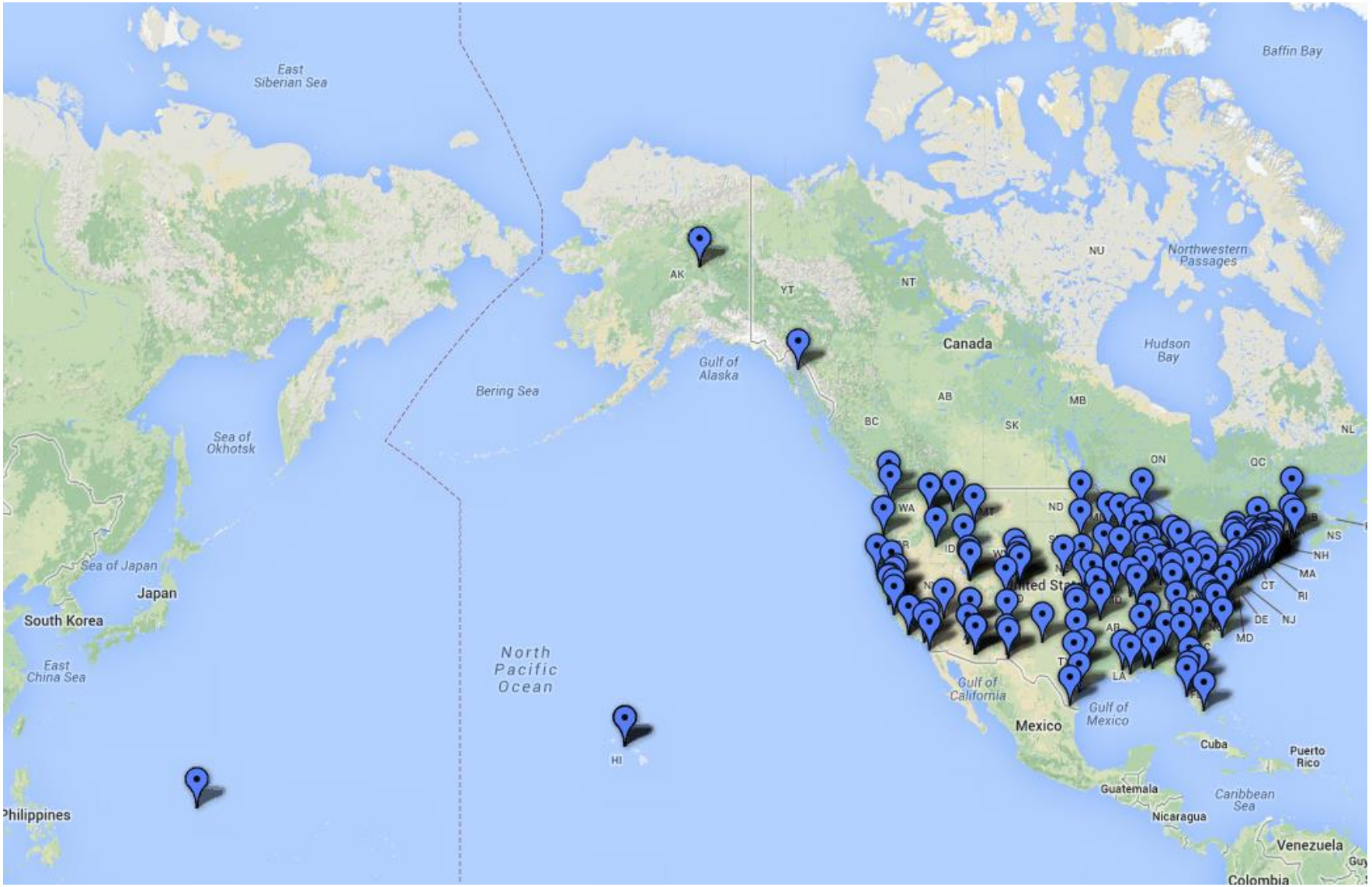
Mandate and Responsibility

- Provide/facilitate portal access to collections data
  - Make information available and discoverable
  - Label Data and images
- Enable digitization and research
  - Facilitate digitization workflows
  - Oversee implementation of standards and best practices for digitization
  - Allow for data discovery across organismal groups
- Be a client of digitization projects/networks
  - Actively seek partners and data sources
  - Respond to cyberinfrastructure needs
- Engage communities
  - Collections
  - Research
  - Citizen science and education
- Support ADBC goals
  - Access to information
  - Support for collections
  - Sustainability

Ten Thematic Collections Networks (TCNs) plus 2 Partner to Existing Networks (PENs)

- InvertNet: An Integrative Platform for Research on Environmental Change, Species Discovery and Identification (*Illinois Natural History Survey, University of Illinois*) http://invertnet.org
- Plants, Herbivores, and Parasitoids: A Model System for the Study of Tri-Trophic Associations (*American Museum of Natural History*) http://tcn.amnh.org
- North American Lichens and Bryophytes: Sensitive Indicators of Environmental Quality and Change (*University of Wisconsin – Madison*) http://symbiota.org/nalichens/index.php http://symbiota.org/bryophytes/index.php (plus 2 PENs)
- **Digitizing Fossils to Enable New Syntheses in Biogeography** - Creating a PALEONICHES-TCN (*University of Kansas*)
- The Macrofungi Collection Consortium: Unlocking a Biodiversity Resource for Understanding Biotic Interactions, Nutrient Cycling and Human Affairs (*New York Botanical Garden*)
- Mobilizing New England Vascular Plant Specimen Data to Track Environmental Change (*Yale University*)
- Southwest Collections of Anthropods Network (SCAN): A Model for Collections Digitization to Promote Taxonomic and Ecological Research (*Northern Arizona University*) http://hasbrouck.asu.edu/symbiota/portal/index.php
- **iDigPaleo: Fossil Insect Collaborative**: A Deep-Time Approach to Studying Diversification and Response to Environmental Change
- Developing a Centralized Digital Archive of Vouchered Animal Communication Signals
- The Macroalgal Herbarium Consortium: Accessing 150 Years of Specimen Data to Understand Changes in the Marine/Aquatic Environment

National Resource (iDigBio), Thematic Collection Networks (TCNs)



To date: 10 TCNs, 2 PENs, 160+ participating institutions, 49 states

Key Features of iDigBio

- Ingest **all contributed data** with emphasis on use of **GUIDs**, no restrictions
- Maintain **persistent datasets** and versioning, allowing new and edited records to be uploaded as needed while preserving existing records
- Ingest **textual** specimen records, plus associated still **images, video, audio, and other media** (or links to these resources as determined by the provider)
- Ingest linked documents and **associated literature**, including field notes, ledgers, monographs, related specimen collections, etc.
- Provide **virtual annotation** capabilities and track annotations back to the originating collection (collaborating with FilteredPush)
- Facilitate sharing and integration of data relevant to biodiversity research
- Provide computational services for biodiversity research

Information Dissemination

In March 2012, the iDigBio Steering Committee established a series of preparation-specific digitization training workshops focused on helping collections managers get started with and/or enhance local digitization programs, all to be held at host institutions.

- DROID (Developing Robust Object->Image->Data, May 2012)
- Herbarium digitization (Valdosta State, September 2012)
- Fluid-preserved collections digitization (U. Kansas, March 2013)
- Dried insect collections digitization (Field Museum, April 2013)
- Collections Digitization (West Virginia, ASB, April 2013)
- Imaging fluid-preserved invertebrates (U. Michigan, September 2013)
- Georeferencing Train-the-Trainers (iDigBio, Gainesville, August 2103)
- Paleontology digitization (Yale Peabody Museum, September 2013)
- Small Herbarium Digitization (Florida State University, December 2013)
- Digitization in the South Pacific (Honolulu, March 2014)
- Paleoimaging (Austin, TX, April 2014)
- Small Herbarium Digitization (Boise, Botany 2014, July 2014)
- Leveraging Digitization Knowledge Across Multiple Domains (Santa Barbara, October 2014)

Workshops

In March 2012, the Steering Committee established a series of preparation-specific digitization training workshops focused on helping collections managers get started with and/or enhance local digitization programs, all to be held at host institutions.

- **DROID (Developing Robust Object->Image->Data, May 2012)**
- **Herbarium digitization (Valdosta State, September 2012)**
- **Augmenting OCR Hackathon (Ft. Worth, February 2103)**
- **Fluid-preserved collections digitization (U. Kansas, March 2013)**
- **Dried insect collections digitization (Field Museum, April 2013)**
- **Collections Digitization (West Virginia, ASB, April 2013)**
- **Georeferencing Train-the-Trainers (iDigBio, Gainesville, August 2103)**
- **Imaging fluid-preserved invertebrates (U. Michigan, September 2013)**
- **Paleontology digitization (Yale Peabody Museum, September 2013)**
- **Small Herbarium Digitization (Florida State University, December 2013)**
- **CitScribe Hackathon (iDigBio, Gainesville, December 2013)**
- **Broadening Biodiversity in the Biodiversity Sciences (Atlanta, January, 2014)**
- **Education and Outreach (iDigBio, Gainesville, January 2014)**
- **Original Source Materials Digitization (Yale Peabody Museum, March 2014)**
- **Digitization in the South Pacific (Honolulu, March 2014)**
- **Recruiting and Retaining Small Collections in Digitization (Mt. Pleasant, MI, April 2014)**
- **Paleo Imaging Workshop (Astin, TX, April 2014)**

IDigBio Working Groups

Contents [hide]

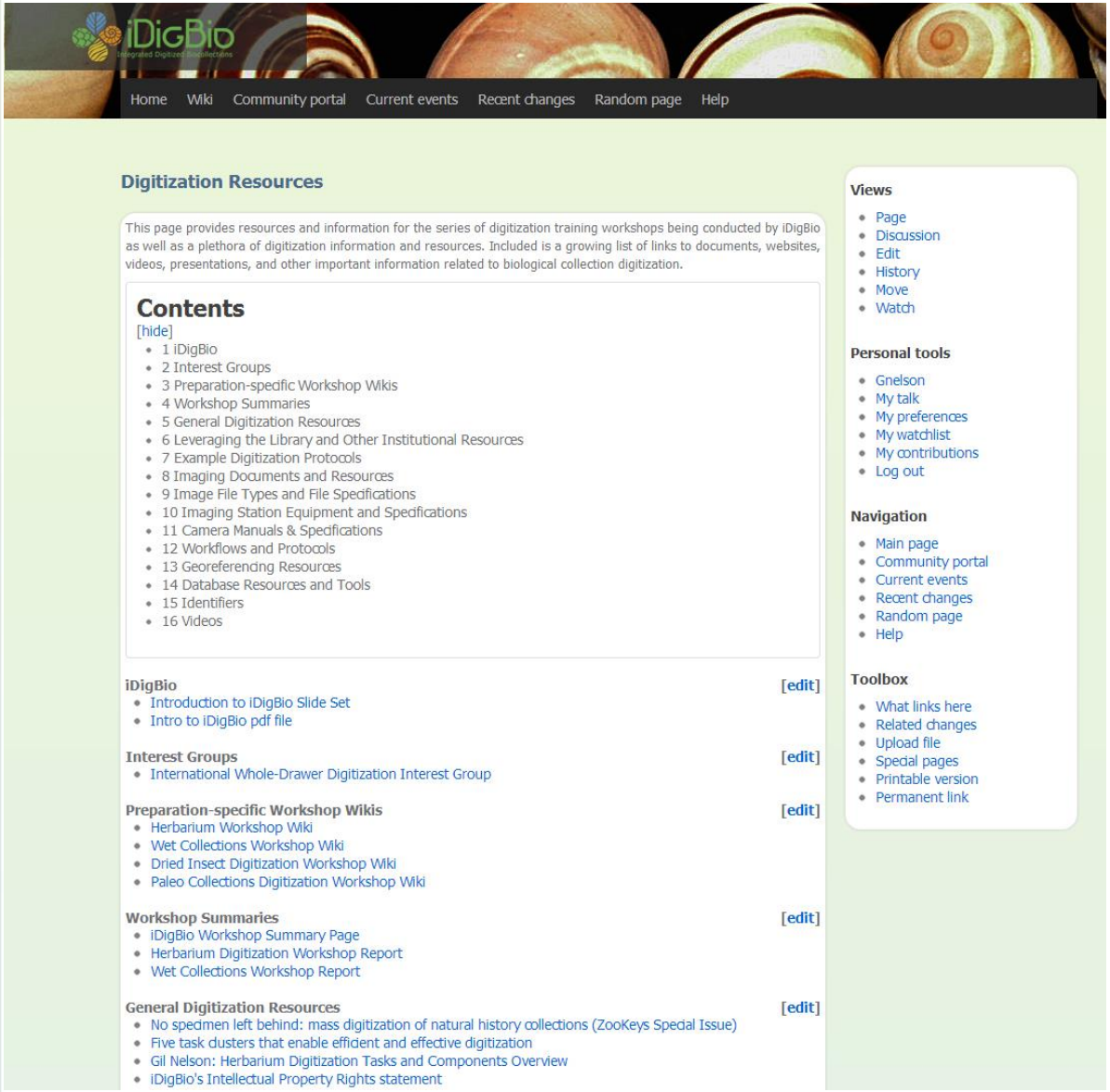
1 Overview
2 Active Working Groups
- 2.1 Augmenting OCR (aOCR)
- 2.2 Biodiversity Informatics Management (BIM) Working Group
- 2.3 Cyberinfrastructure (CYWG)
- 2.4 Developing Robust Object to Image to Data (DROID1)
- 2.5 Developing Robust Object to Image to Data (DROID2)
- 2.6 Developing Robust Object to Image to Data (DROID3): 3D Objects and Things in Spirits
- 2.7 Education & Outreach (E&O)
- 2.8 Georeferencing Working Group (GWG)
- 2.9 International Whole-Drawer Digitization Interest Group (WDD)
- 2.10 Minimum Information Standards, Authority Files, & Semantics (MISC)
- 2.11 NANSH Working Group (NANSH)
- 2.12 Paleo Digitization Working Group (PaleoDigi)
- 2.13 Paleontology (Paleo)
- 2.14 Public Participation in Digitization (CitSci)
- 2.15 Strategic Communication Interest Group
- 2.16 Website Content Providers Editorial Board and Interest Group

3 Inactive Working Groups
- 3.1 Authority Files
- 3.2 Intellectual Property Policy

Overview[edit]

iDigBio supports a number of Working Groups and Interest Groups. Several working groups are focused on the delivery of short-term objectives, while other standing Working Groups are tasked with ongoing research, development, and improvement activities. This page provides an overview of both current (active) and disbanded (inactive) Working Groups and Interest Groups.

The section "Overlap with Other Working Groups" should be used to list subject areas that may duplicate some effort from another Working Group. When Working Group activities stray into these overlapping subject areas then collaboration between working groups is warranted for those tasks.



Developed a community-oriented digitization resources wiki in support of our workshops and to serve digitization-related information across all preparation types.

Established a digitization list serv to promote workshop follow-up as well as community discussion and sharing.


iDigBio
Integrated Digitized Biocollections