

# iDigBio - Integrated Digitized Biocollections

Mobilizing natural history collections for the understanding of island biodiversity

Shelley James, Deborah Paul, Matthew Collins & Gil Nelson iDigBio sjames@flmnh.ufl.edu











iDigBio is funded by a grant from the National Science Foundation's Advancing Digitization of Biodiversity Collections Program (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.



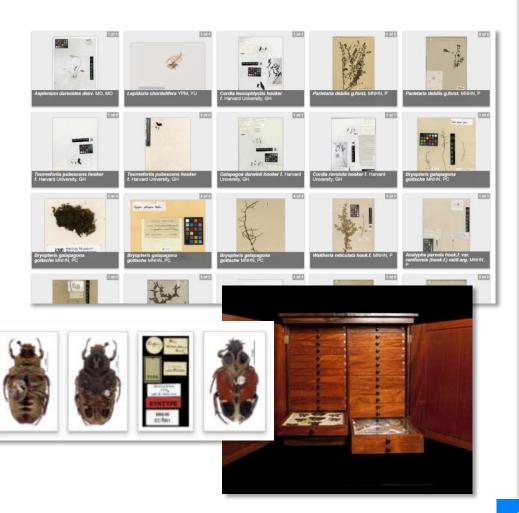
# Biological collections and Island Biology From before Darwin & Wallace to the present and into the future

Historical baseline data
Trait and habitat information
Taxon distributions
New species discovery

• • •









### Estimates suggest that there are between:

### 500 million and 1 billion

biological and paleobiological specimens in the United States;

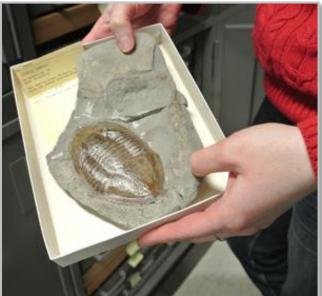
### 3-4 billion worldwide?

No one really knows for sure! A wealth of biodiversity information.











# Integrated Digitized Biocollections: who we are

- Central coordinating center for the U.S. national effort to digitize natural history collections
  - University of Florida & Florida State University
- Part of NSF's Advancing Digitization of Biodiversity Collections (ADBC) program
  - 10 year, \$100 million effort
- National network of over 250 institutions in 50 U.S. states
   and 1 territory
- Digitize biological and paleontological (non-federal) public collections in U.S. institutions











### What do we do?

- Enable digitization of biodiversity collections data
  - Develop efficient & effective standards & workflows
  - Workforce education & training
- Provide portal access to biodiversity data in a cloud computing environment
  - Respond to cyberinfrastructure needs
  - Enable access & discoverability
- Facilitate use of biodiversity data to address key environmental and economic challenges
  - Researchers, educators, general public, policy-makers, ...
- Plan for long-term sustainability of the national digitization network & effort
  - Expand participation: partners, data sources, public, ...
  - Proliferate and broaden uses of biodiversity data





# More about ADBC – the US funding mechanism

- Thematic Collections Networks (TCNs) have a research focus, driving digitization efforts
- Collaborative, community building
- Converting analog voucher specimen data to digital format, to include transcription of text data (labels, catalogs, field notes, etc.) and capturing specimen images, video, sounds.





# 15 Thematic Collections Networks (TCNs) 15 Partners to Existing Networks (PENs)

- **Invert**Net: An Integrative Platform for Research on Environmental Change, Species Discovery and Identification
- Plants, Herbivores, and Parasitoids: A Model System for the Study of Tri-Trophic Associations
- North American Lichens and Bryophytes: Sensitive Indicators of Environmental Quality and Change
- Digitizing **Fossils** to Enable New Syntheses in Biogeography Creating a PALEONICHES-TCN
- The **Macrofungi** Collection Consortium: Unlocking a Biodiversity Resource for Understanding Biotic Interactions, Nutrient Cycling and Human Affairs
- Mobilizing New England Vascular Plant Specimen Data to Track Environmental Change
- Southwest Collections of **Arthropods** Network (SCAN): A Model for Collections Digitization to Promote Taxonomic and Ecological Research
- 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
- Collaborative: Documenting the Occurrence through Space & Time of Aquatic **Non-indigenous Fish, Mollusks, Algae, & Plants**Threatening North America's Great Lakes
- Collaborative Research: The Key to the Cabinets: Building and Sustaining a Research Database for a Global Biodiversity Hotspot (**plants**)
- InvertEBase: reaching back to see the future: species-rich **invertebrate faunas** document causes and consequences of biodiversity shifts
- The **Microfungi** Collections Consortium: A Networked Approach to Digitizing Small Fungi with Large Impacts on the Function and Health of Ecosystems (MiCC)
- Documenting Fossil Marine Invertebrate Communities of the Eastern Pacific Faunal Responses to Environmental Change over the last 66 million years (PCMIF)
   Island Biology TCN?



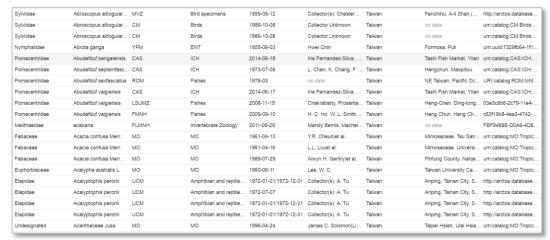
# iDigBio: HUB and data aggregator

# Our scope:

- Specimen data Darwin Core
- Media Audubon Core
- Global; all taxa

















Biodiversity Information

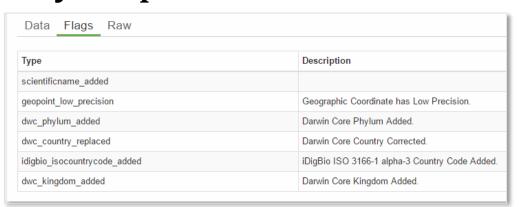
Standards



# Why collections publish/share data with iDigBio

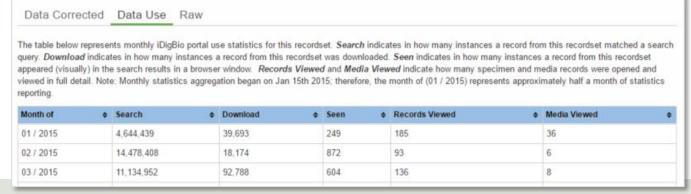
- Discovery and use
- Data quality, improvement

Specimen Record



Attribution, credit, collection value metrics

Recordset





### What are the *minimum* data fields needed?

### **Darwin Core**



- recordID
- **occurrenceID** (unique!)
- scientificName
- eventDate
- recordedBy
- Locality information
- catalogNumber
- institutionID
- collectionID
- Geological Context
- (Other DwC data fields & extensions)

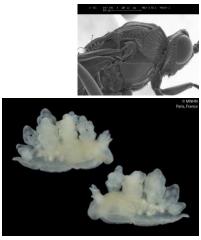
### Audubon Core @\_\_\_\_\_



- recordId
- occurrenceID of specimen
- URL
- Camera EXIF
- photographer

### Metadata

- Institution
- Collection
- Contact & info
- Description
- **URL**

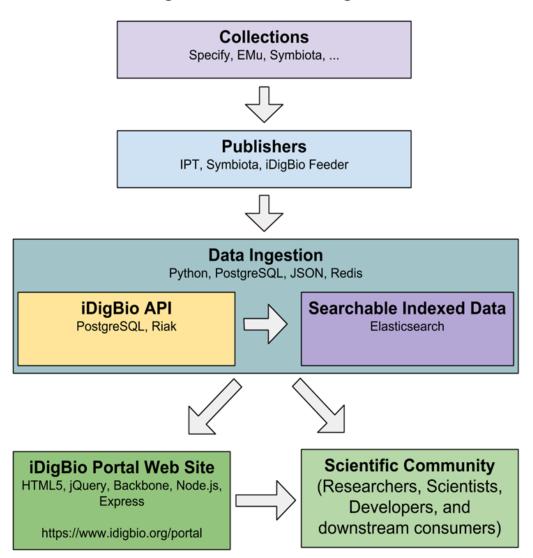








### iDigBio Data Flow Diagram





## The face of iDigBio: idigbio.org



About iDigBio

Research

**Technical Information** 

Education



Search

ly account Log out

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

64,015,275

Specimen Records

14,321,696

Media Records

786

Recordsets

Search the Portal



Why digitization matters

More about what we do and why



#### Digitization

Learn, share and develop best practices



Documentation on data ingestion



#### **Working Groups**

Join in, contribute, be part of the community



#### **Proposals**

New tool and workshop ideas

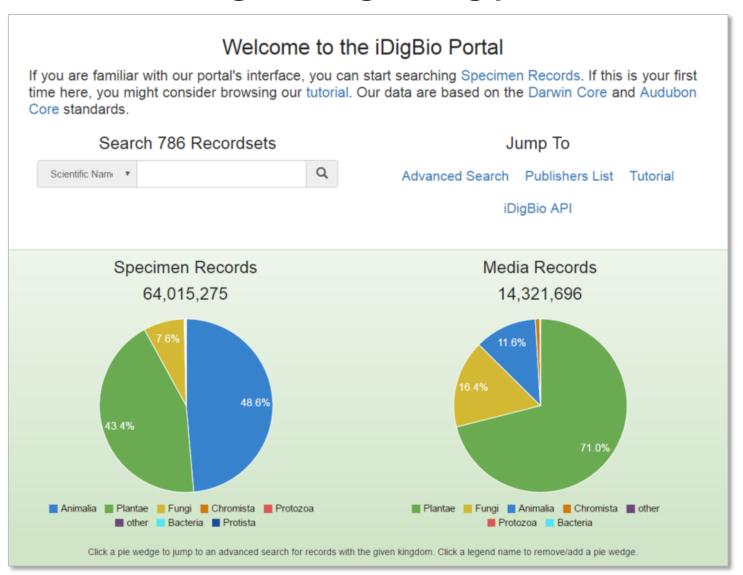


#### **Citizen Scientists**

How can you help biological collections?



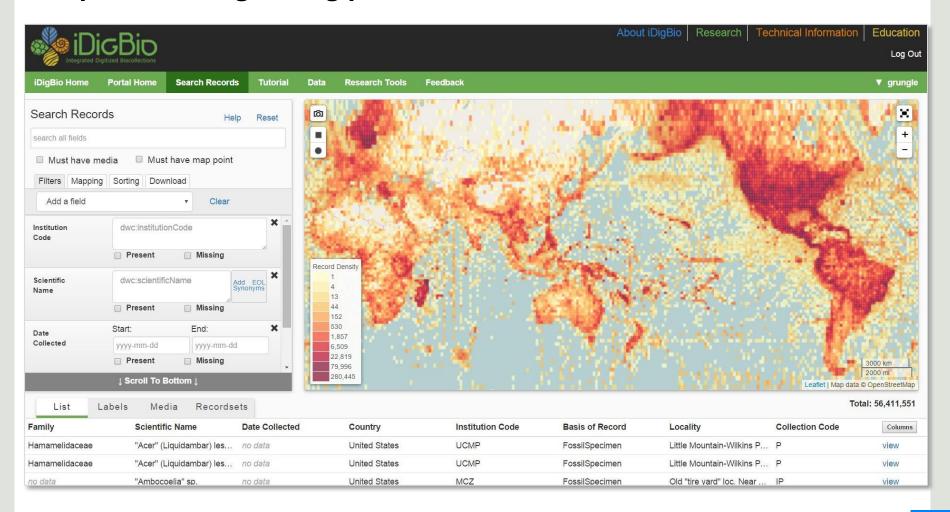
### The data within iDigBio: idigbio.org/portal





## iDigBio portal interface

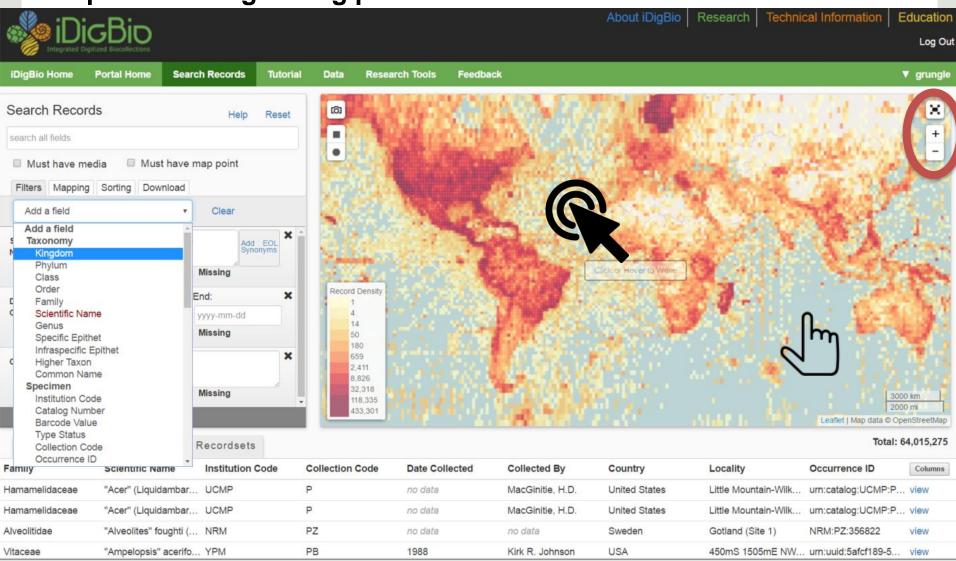
https://www.idigbio.org/portal/search





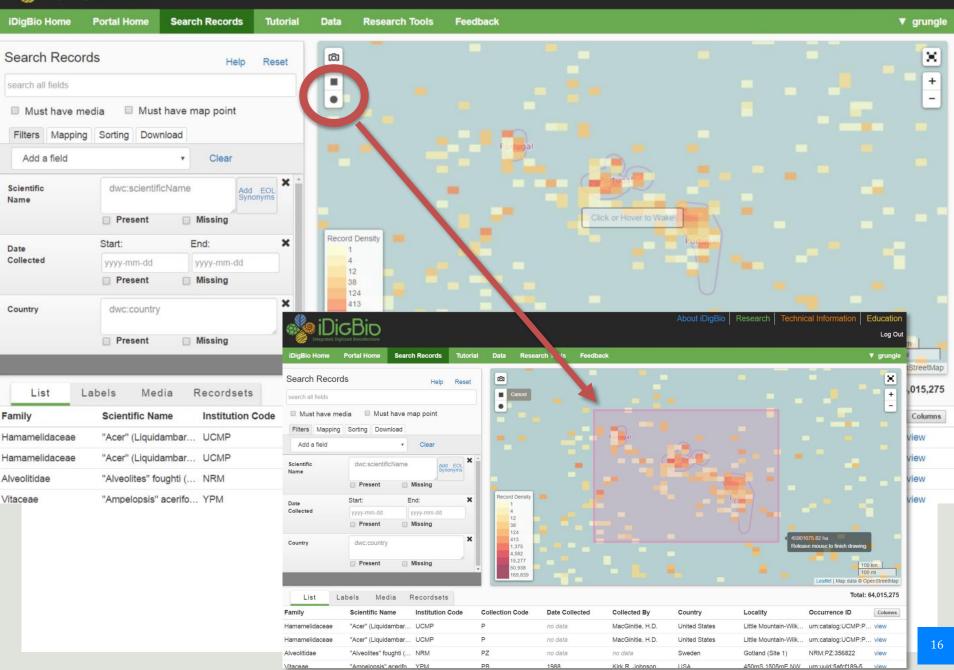
# iDigBio portal interface

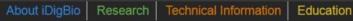
https://www.idigbio.org/portal/search



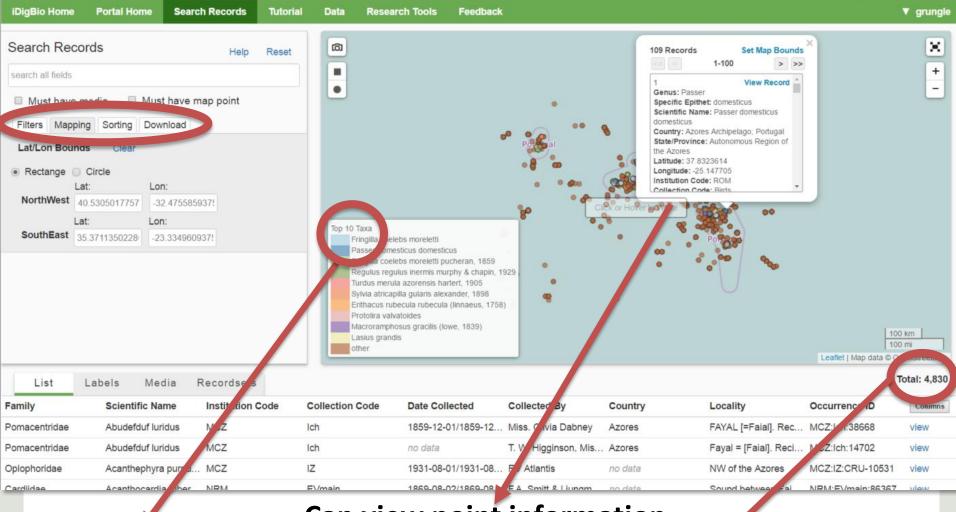
Log Out











Top 10 taxa

Can view point information

Number records with geopoint (5,680 using filters)

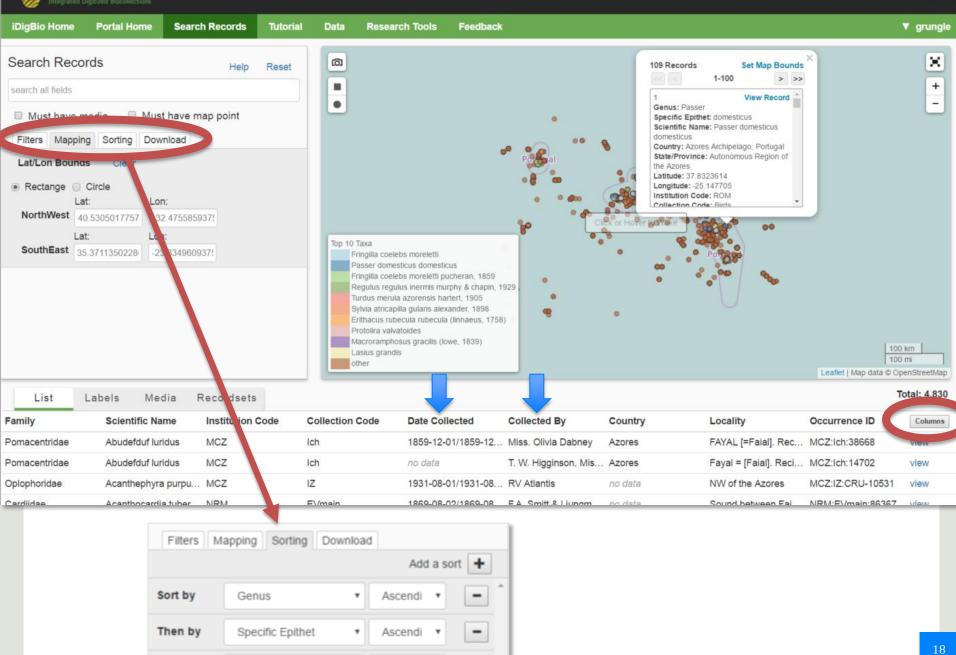
Log Out



Then by

Date Collected

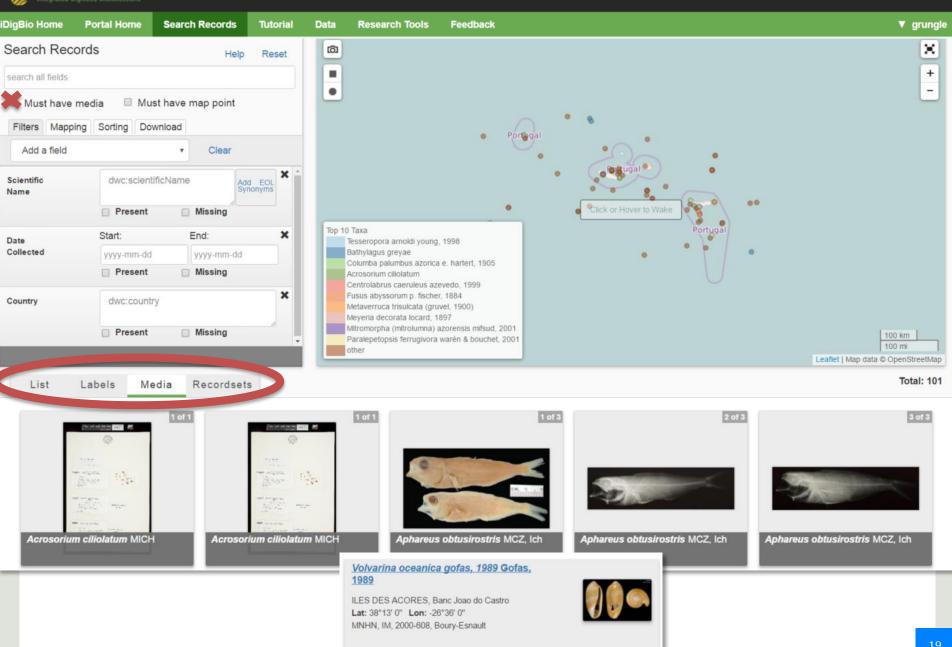
Ascendi



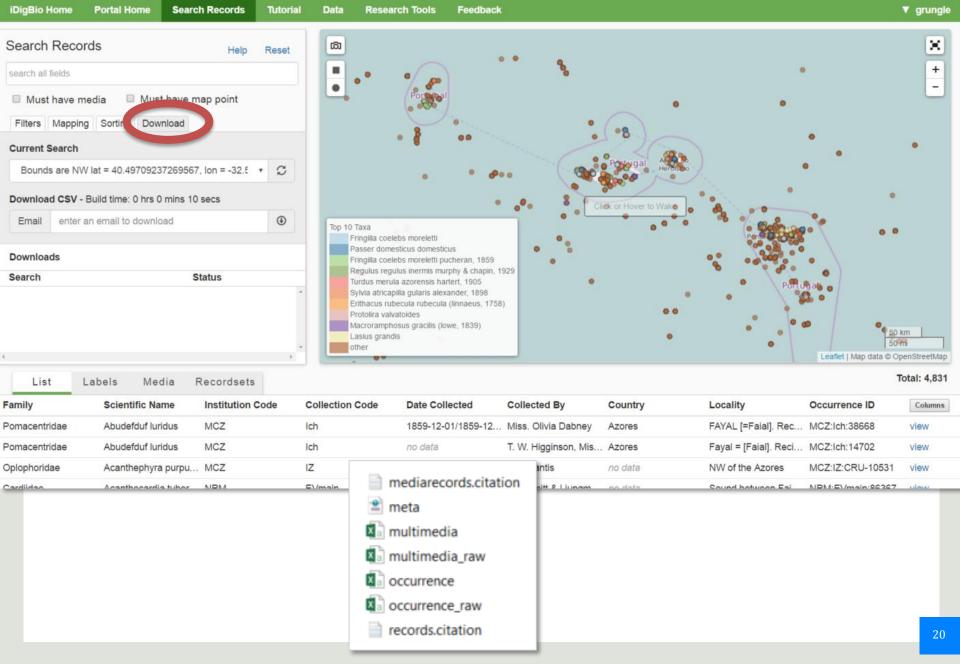
1971-10-07

Log Out









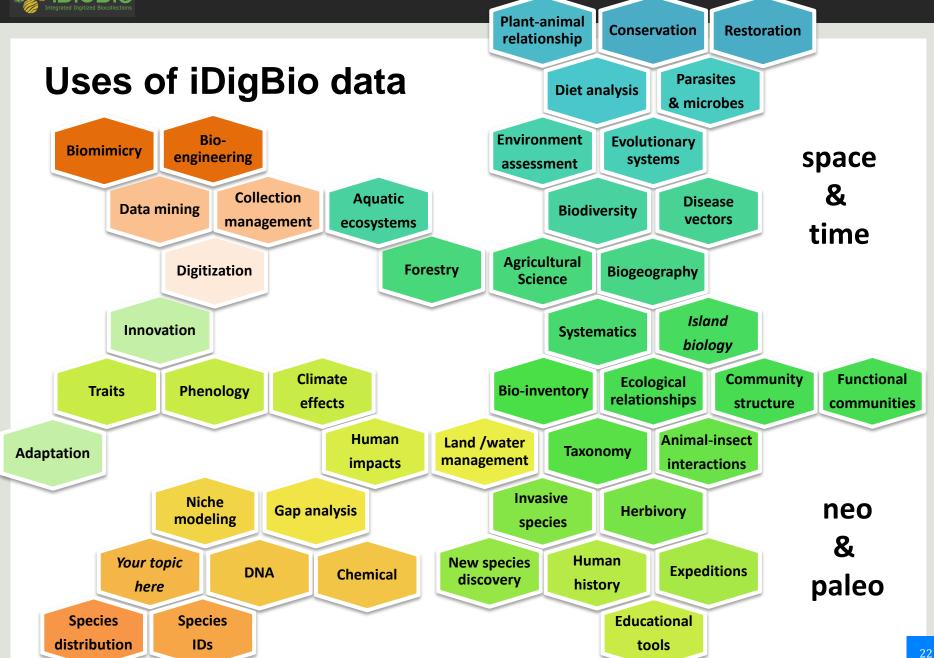


### >3.6 million voucher specimen records for islands and archipelagos



Imagine if more regional collection recordsets were included!







# Enabling research using specimen data

iDigBio Search API

ridigbio: Interface to the iDigBio Data API



Species distribution tools



Integrated workflows with phylogenies







Effechecka Gouda Fresh data

Open Source cluster computer framework



Enhancing
Paleontological and
Neontological Data API



## iDigBio Collaborates!

















Information Standards





































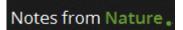














# Data services – what do you need?

- Tools for using data for research
- Linking to genetic resources or isotopic data
- Base maps climate layers etc.
- Trait data as a surrogate for functionality
- Mining for habitat data
- Biological community structure
- Enhancing & enriching data (e.g. literature, field notes)
- Tools for measuring characteristics from images

Let's collaborate!





# Training & Outreach! Promoting research use, data improvement, future scientists, and humanity

- iDigBio has accommodated >2,600 participants from 500 unique institutions in >75 workshops, about 20 a year
- Island community needs? Overcoming georeferencing challenges Field to aggregator workshops Digitization workflow

### **Curriculum development, training**







**Citizen Science** 

www.idigbio.org

Research Spotlight **Biodiversity Spotlight** 

Webinar vimeo com/idigbio

Contribute to the community via www.idigbio.org/ wiki



### www.iDigBio.org



Advancing Digitization of Biological Collections

Shelley James, sjames@flmnh.ufl.edu















