

# ***NIBA, ADBC and iDigBio: Transforming the Landscape***

Greg Riccardi  
Florida State University  
iDigBio  
griccardi@fsu.edu



*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.*



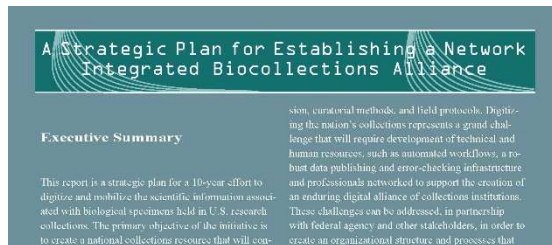
# From NIBA to ADBC to iDigBio

- Network Integrated Biocollections Alliance (NIBA)
  - Community vision of what is needed and what can be achieved
  - Two reports produced
- Advancing Digitization of Biodiversity Collections (ADBC)
  - NSF commitment to 10 years of funding
  - Hub for community invol
  - Many digitization projects (Thematic Collections Networks)
- iDigBio
  - The Hub for data and digitization

# Informatics for Biodiversity

- Informatics
  - Acquisition
  - Management
  - Usage
- Purpose of the talk is to explore the informatics issues for supporting usage of collections

# National Initiatives from NSF



**2012**

**Goal 1:** Establish an that will provide the mechanism required strategic plan.

**Goal 2:** Advance eng cyberinfrastructure around core c workflows, ef archives, innc policy, and ut

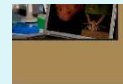
**2010**

**Vision Statement:** The Network Integrated Biocollections Alliance will develop an inclusive, vibrant partnership of U.S. biological collections that collectively will document the nation's biodiversity resources and create a dynamic electronic resource that will serve the country's needs in answering critical questions about the environment, human health, biosecurity, commerce, and the biological sciences.

**2007**

**Vision**

and support of a comprehensive cyberinfrastructure essential to 21st century advances in science and engineering research and education.



# NIBA strategic plan

- Plan for a 10-year effort to digitize and mobilize the scientific information associated with biological specimens held in U.S. research collections.
- Call to Action: we need
  - A central organization for integration of people and data
  - Thematic networks based on research areas
  - Regional and clade-based efforts
  - New tools and technologies
  - Training and outreach
  - Partner with other agencies, organizations

# NIBA Use Scenarios

- *A massive oil spill occurs off the coast of Louisiana.*
  - *With the national biological database completed, stakeholders have instantaneous access to all life forms that have been recorded from the region actually or potentially threatened by the spill.*
- *The Yosemite National Park mammal survey was combined with baseline data to evaluate the threat of climate change on living animal populations.*
  - *Baseline from the Joseph Grinnell expeditions of 1914-1920 in the Museum of Vertebrate Zoology at UC Berkeley*
  - *Over the intervening 90+ years temperatures for the park had risen by as much as 7° F and several key mammal species ranges had changed dramatically.*

## NIBA Use Scenarios 2

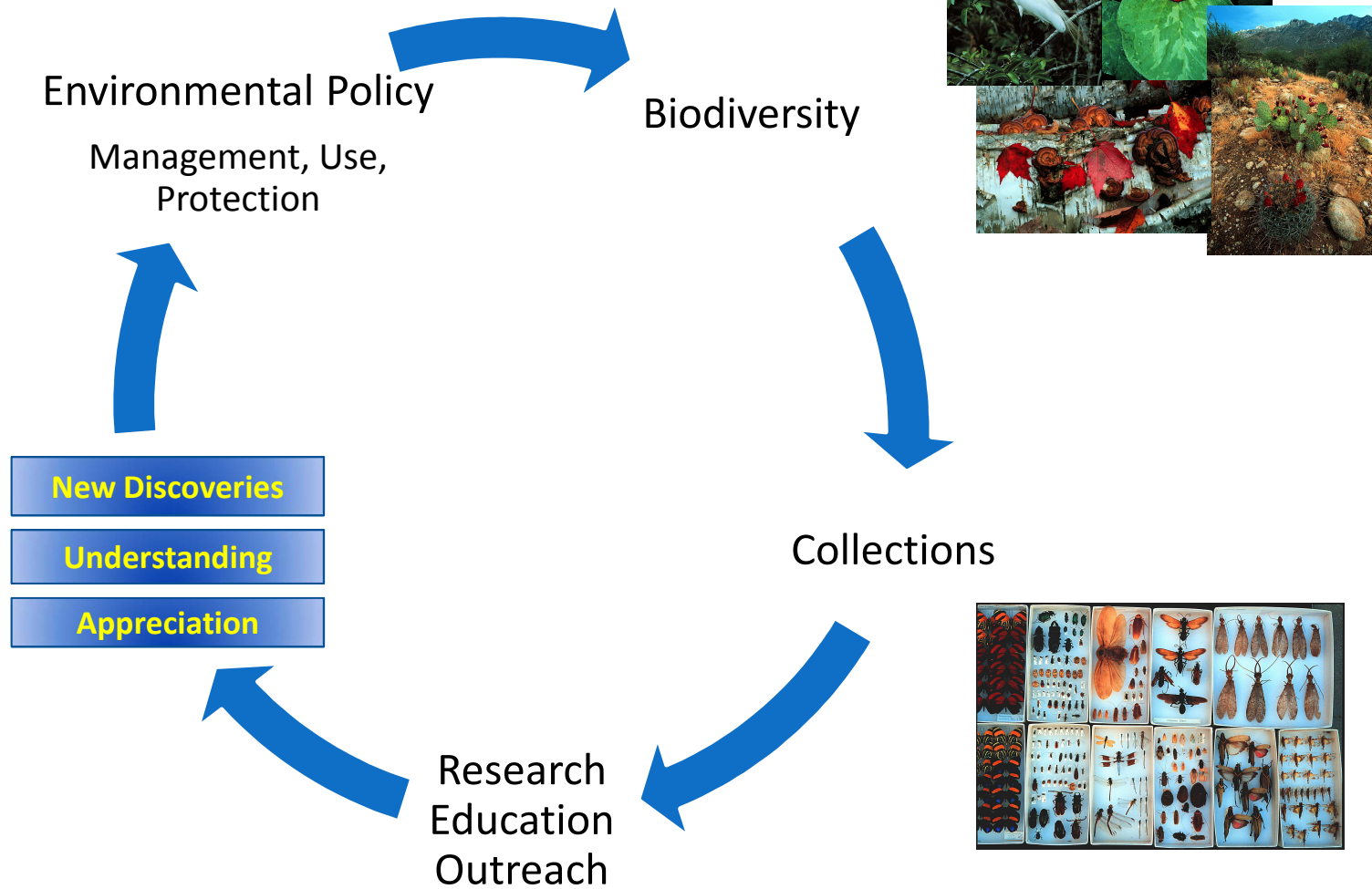
- *An airliner encounters a bird strike while taking off from Reagan National Airport*
  - *Analysis of feathers and comparison with collections showed migratory birds off their normal paths*
- *A deadly virus suddenly appears in the southwestern United States, infecting hundreds of people, at least half of them fatally.*
  - *Field biologists quickly determine that it is carried by several species of rodents.*
  - *Public health officials and zoologists were able to delineate the geographic area in which the virus is likely to occur, the habitats of the rodent species, associated species, and other parameters of their life cycle that will influence the spread of the virus.*
  - *The availability of digital records of the extensive mammal collections from the region, as well as ongoing field research with the species involved (Yates et al., 2002).*

# The Scope of Collections Digitization

- Collections contain a cross-section of the world's biodiversity including fossils, invertebrates, vertebrates, protists, fungi, plants, and human cultures
- Collections digitization is defined broadly to include
  - transcription into electronic format of various types of data associated with specimens,
  - capture of digital images of specimens, and
  - georeferencing of specimen-collection localities
- Prior to this initiative, there has been no nationwide coordination of the effort to digitize and electronically mobilize data from biological research collections

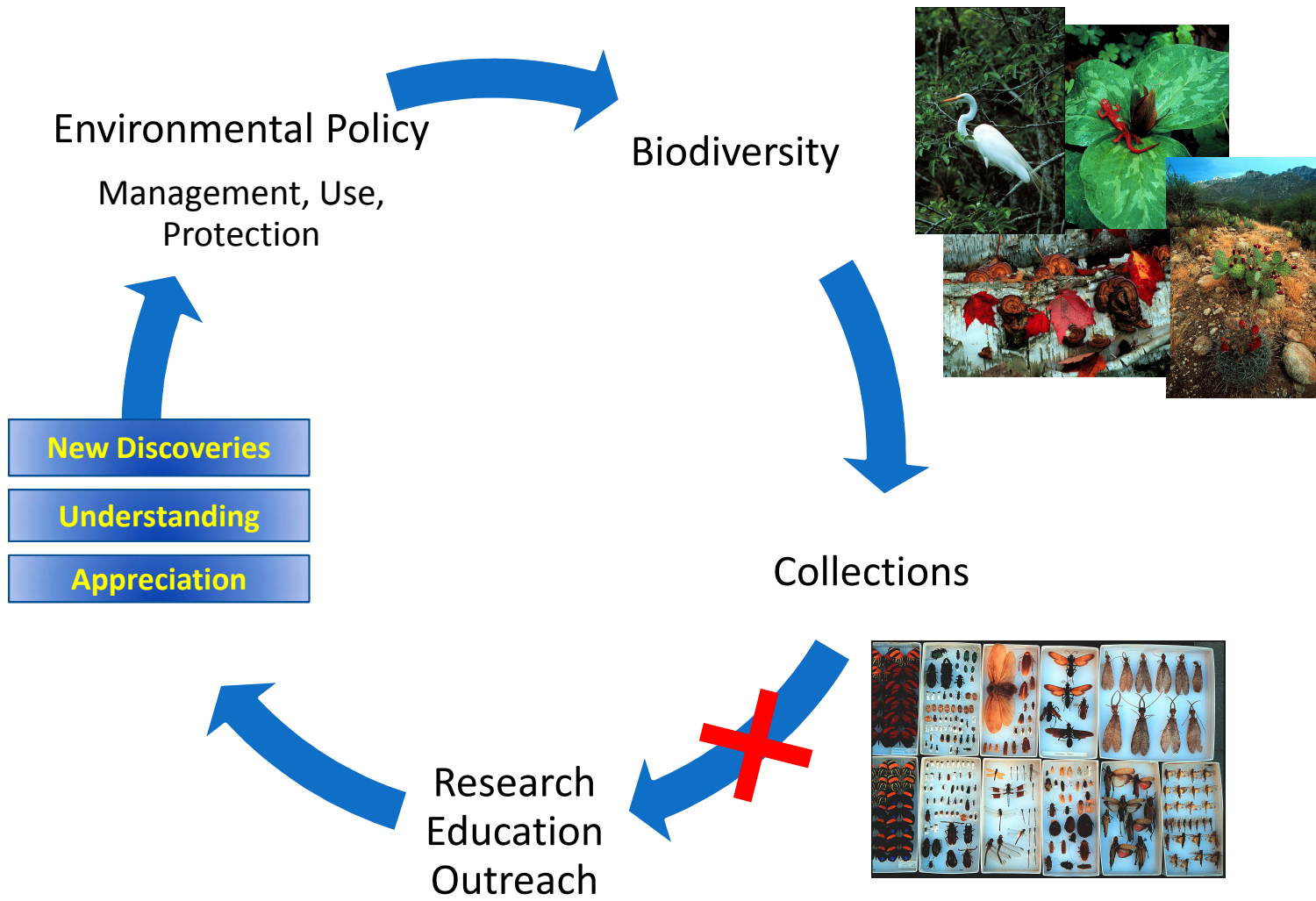


# The Vision of NIBA



# The Reality

The data in biodiversity collections are inaccessible to most potential users



## U.S. National Science Foundation

### Advancing Digitization of Biodiversity Collections Program (ADBC)

The goal of **ADBC** is to *remove this inaccessibility* through **digitization**: putting information online so that researchers, educators, students, natural resource managers, environmentalists, and policymakers have access.



**\$100 million over 10 years to digitize specimen-based data in non-federal U.S. collections**

# The ADBC Plan



# ADBC was born!

- Focused on:
  - Central coordinating resource for digitization projects (HUB)
  - Digitization based on research challenges
  - TCNs = Thematic Collections Networks
- Requires:
  - Innovative approaches to digitization
  - Inclusivity and prioritization
  - Training and outreach

# Requirements for the HUB

- coordinate the digitization efforts
- enable links with existing digitization projects
- assume a leadership and coordination role
- document the use of digitized collections data
- play a key role in the development of a national resource
- serve as a central site for integrating data and techniques
- develop a coordination network
- identify gaps and priorities for digitization efforts
- design and maintain a living resource that tracks research outcomes
- form the coordinating scientific team
- link to similar efforts in the federal government
- assure that results are disseminated to the scientific community
- coordinate activities with the thematic collections networks
- in a timely manner and regular schedule
- facilitate implementation of digitization and data management
- link to other countries
- establish a community-wide plan for storage, management, and access
- enable ongoing communication between collections
- assess gaps in digitization of collections
- plan for the long-term sustainability of the national resource
- demonstrate experience and understanding of digitization
- link to existing networks of collections
- establish cohesion and interconnectivity among digitization efforts

# Requirements for TCNs

perform fundamental collections digitization

engage in training activities

develop appropriate technology and standards to produce an interoperable network

conduct the digitization effort at a number of collections

focus on digitization and mobilization of collections data and images in existing collections

integration across types of collections is expected for broad research themes

share infrastructure among the collections involved in the project

identify deliverable goals and metrics for assessment

identify specific needs for community support

reach out to other collections for inclusion in the digitization effort

fill gaps in the effort to provide online access to specimen data for all existing biological collections

integrate with other ongoing digitization activities

be broad and inclusive of all collections related to the theme

# Ten Thematic Collections Networks (TCNs)

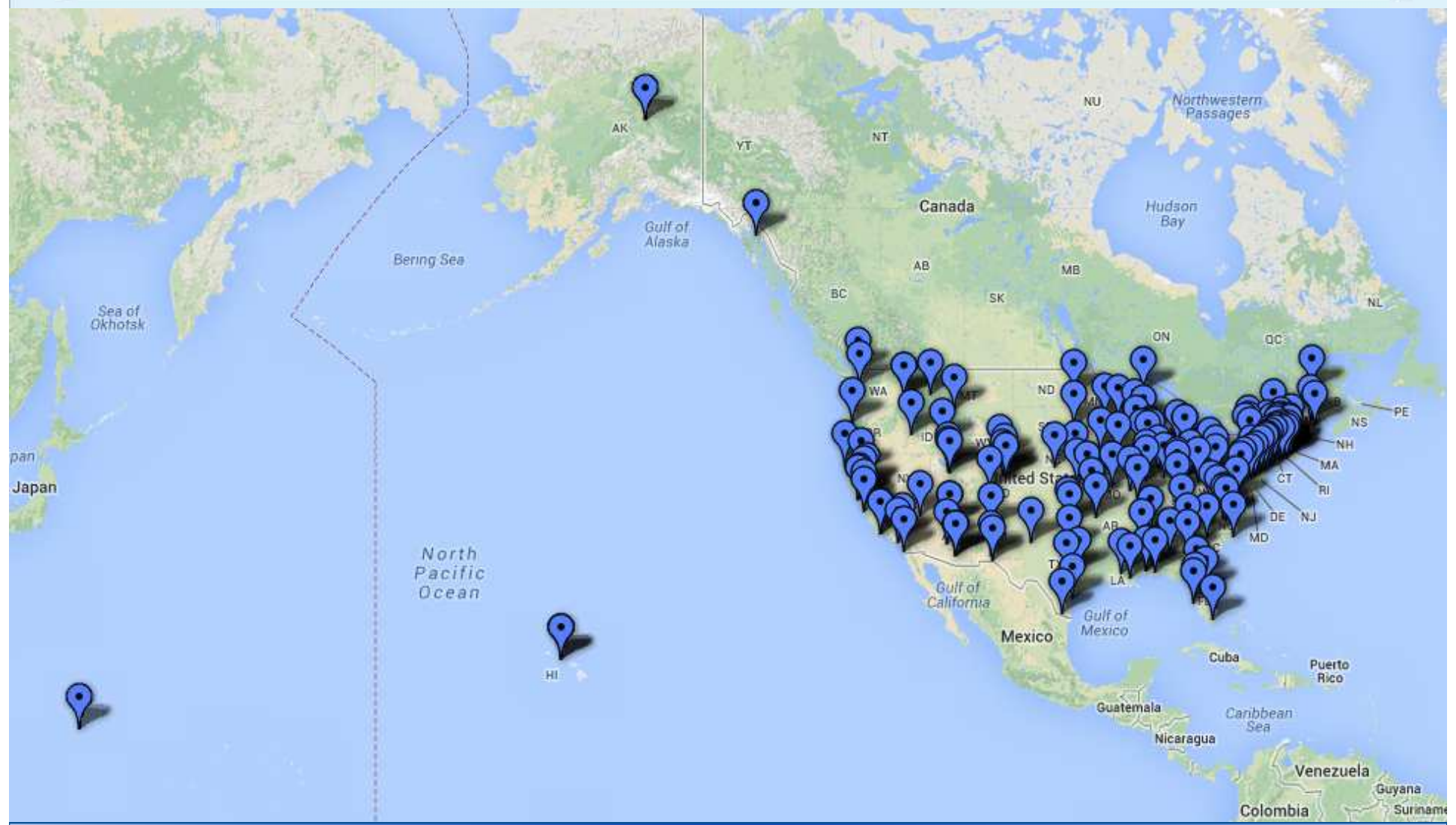
- **InvertNet:** An Integrative Platform for Research on Environmental Change, Species Discovery and Identification (Illinois Natural History Survey, University of Illinois)
- **Plants, Herbivores, and Parasitoids:** A Model System for the Study of Tri-Trophic Associations (American Museum of Natural History)
- **North American Lichens and Bryophytes:** Sensitive Indicators of Environmental Quality and Change (University of Wisconsin – Madison)
- **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 Arthropods Network (SCAN):** A Model for Collections Digitization to Promote Taxonomic and Ecological Research (Northern Arizona University)
- **The Macroalgal Herbarium Consortium:** Accessing 150 Years of Specimen Data to Understand Changes in the Marine/Aquatic Environment (University of New Hampshire)
- **Developing a Centralized Digital Archive of Vouchered Animal Communication Signals** (Cornell University)
- **Fossil Insect Collaborative:** A Deep-Time Approach to Studying Diversification and Response to Environmental Change (University of Colorado at Boulder)





National Resource for Advancing  
Digitization of Biological Collections

# NATIONAL HUB, THEMATIC COLLECTION NETWORKS, AND COLLABORATORS



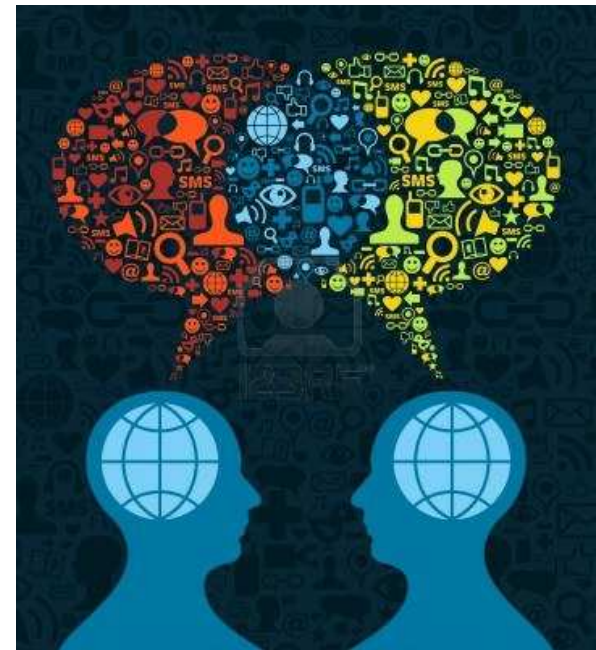
**Ten TCNS and collaborating institutions: 152 institutions in 50 states**

# iDigBio Accomplishments

- Community building
- Digitization
- Enabling data export and usage
- IT and Data Portal

## iDigBio: Workshops and Working Groups, Establishing Connections, and Cyberinfrastructure

- 26 workshops and symposia (about 1/month)
  - >510 (unique) participants from 260 institutions
- Representation at other events
- Established relations with other collections organizations and biodiversity initiatives
  - CollectionsWeb
  - iPlant
  - NESCent
  - NEON
  - AIBS
  - SPNHC
  - NSC Alliance
  - GBIF: USGS/BISON
  - EOL
  - ---nets



14,036,195  
Specimen Records



Making  
biological  
web

### Start Searching Specimen Records

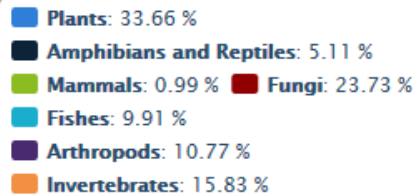
Press ESC to close auto-complete suggestions.

Search

### iDigBio Specimen Portal

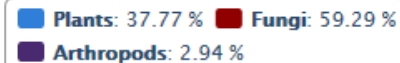
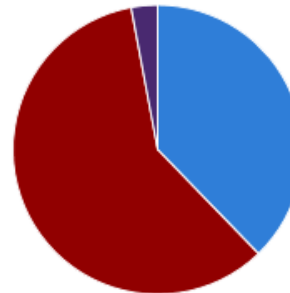
If you are familiar with our portal's interface, you can start searching [Specimen Records](#). If this is your first time here, you might consider browsing [our tutorial](#). Our data are based on the [Darwin Core](#) and [Audubon Core](#) standards.

Specimen Records by Collection Type



Highcharts.com

Media Records by Collection Type



Highcharts.com

14,036,195  
Specimen Records

2,290,983  
Media Records

235  
Recordsets



#### Digitization

Learn, share and  
develop best  
practices

#### Researchers

Learn about research

# Search by scientific name

## → Search Records

Full Text Search

only records with images [Hide Advanced Search](#)

## Current Results

Query: Scientific Name = percina ni

Records: 282

Approx. Download Time: 0hrs 0mins 15secs

Email:



## Advanced Search

<p>Scientific Name</p> <p>percina nigrofasciata</p> <p><input type="checkbox"/> Present <input type="checkbox"/> Missing</p>	<p>Genus</p> <p></p> <p><input type="checkbox"/> Present <input type="checkbox"/> Missing</p>	<p>Country</p> <p></p> <p><input type="checkbox"/> Present <input type="checkbox"/> Missing</p>	<p>State/Province</p> <p></p> <p><input type="checkbox"/> Present <input type="checkbox"/> Missing</p>
--	---	---	--

please select  Scientific Name  Ascending   [Tips & Hints](#)

Search Matched 282 Records

Family	Scientific Name	Genus	Country	State/Province
Percidae	Percina nigrofasciata	Percina	United States	Georgia
	Percina nigrofasciata		United States	South Carolina
	Percina nigrofasciata		United States	Alabama
	Percina nigrofasciata		United States	Louisiana
Percidae	Percina nigrofasciata	Percina	United States	Florida
	Percina nigrofasciata		United States	Alabama
Percidae	Percina nigrofasciata	Percina	United States	Alabama
Percidae	Percina nigrofasciata	Percina	United States	South Carolina

# Status of ADBC

- Communities have been established
  - Major digitization processes supported by working groups
  - TCNs are up and running in collaboration
  - Smaller collections are involved
  - Education and outreach
- Information is agile
  - Moves from Collections to iDigBio and other aggregators to users

# What's next?

- Important informatics issues still in process
  - Attribution for data use
  - Data quality assurance
  - Feedback on data quality
  - How can ADBC information management be sustainable
- Lots of room for improvement in
  - Digitization processes
  - Support for collection informatics, especially for small collections

# Thank you!



**Specimens are the  
scientific documentation  
of the Earth's biological  
diversity**



*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.*





# We Dig People!



# Problem with the vision

- The data in biodiversity collections are inaccessible to most potential users