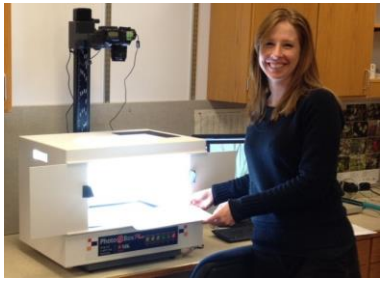
iDigBio
Integrated Digitized Biocollections

Overview, progress, and future



Larry M. Page

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.

NSF

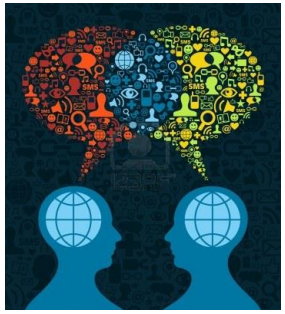
iDigBio
Integrated Digitized Biocollections

UF
UNIVERSITY of FLORIDA

Coordinating Center for Advancing Digitization of Biodiversity Collections (ADBC)

FLORIDA STATE UNIVERSITY
VIRES ARTES MORES
1851

ADBC: *$100 million over 10 years from U.S. National Science Foundation*



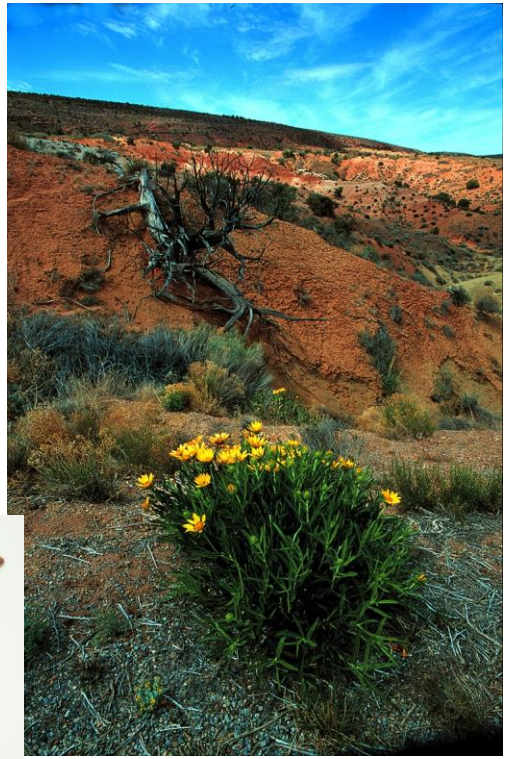
University of Florida

Larry Page: *Project Director*
Jose Fortes*: Cyberinfrastructure*
Pam Soltis*: Research opportunities*
Bruce MacFadden: *Outreach activities*

Florida State University

Greg Riccardi*: Digitization activities*

Biodiversity Collections



The single largest source of information on biological diversity (outside nature)

1,300 – 1,500 collections in U.S.A.

1 billion specimens in USA

3 billion specimens globally

Biodiversity Collections

- **Morphology/anatomy**
- **Taxonomy**
- Every species name is based on 1 or more specimens
- **What, when & where**
- Distributions – **maps (spatial)**
- Changes over time (**temporal**; climate-change)
- **DNA --> phylogenetics**
- **Fossils**: All that we know about prehistoric life
- **Associated data**
- Syntopic species
- Parasites
- Diet
- Reproductive cycles/phenology
- Development/ontogeny

Biodiversity

Collections: specimens, DNA, ancillary information

1 billion specimens

Benefits

Environmental Policy
Management, Use, Protection

- New Discoveries
- Understanding
- Appreciation

Research
Education
Outreach

PROBLEM:

The data in biodiversity collections are inaccessible to most potential users

Biodiversity

Collections: specimens, DNA, ancillary information

1 billion specimens

Benefits

Environmental Policy
Management, Use, Protection

- New Discoveries
- Understanding
- Appreciation

Research
Education
Outreach

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.

U.S. National Science Foundation

Advancing Digitization of Biodiversity Collections Program (ADBC)

The ADBC Program was developed in response to the **Network Integrated Biocollections Alliance (NIBA) Strategic Plan** for a sustained effort to digitize the nation's biodiversity collections

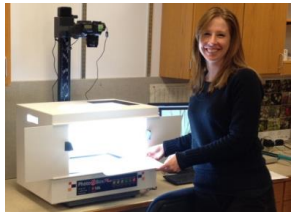
$100 million over 10 years
non-federal collections

Biodiversity

Collections: specimens, DNA, ancillary information

ADBC: Digitization

Databases
Georeferencing
Images



Benefits

Environmental Policy

Management, Use,
Protection

New Discoveries

Understanding

Appreciation

Research

Education

Outreach

Advancing Digitization of Biodiversity Collections

Funds

'Thematic Collections Networks' or TCNs – groups of institutions that digitize data organized around a research question

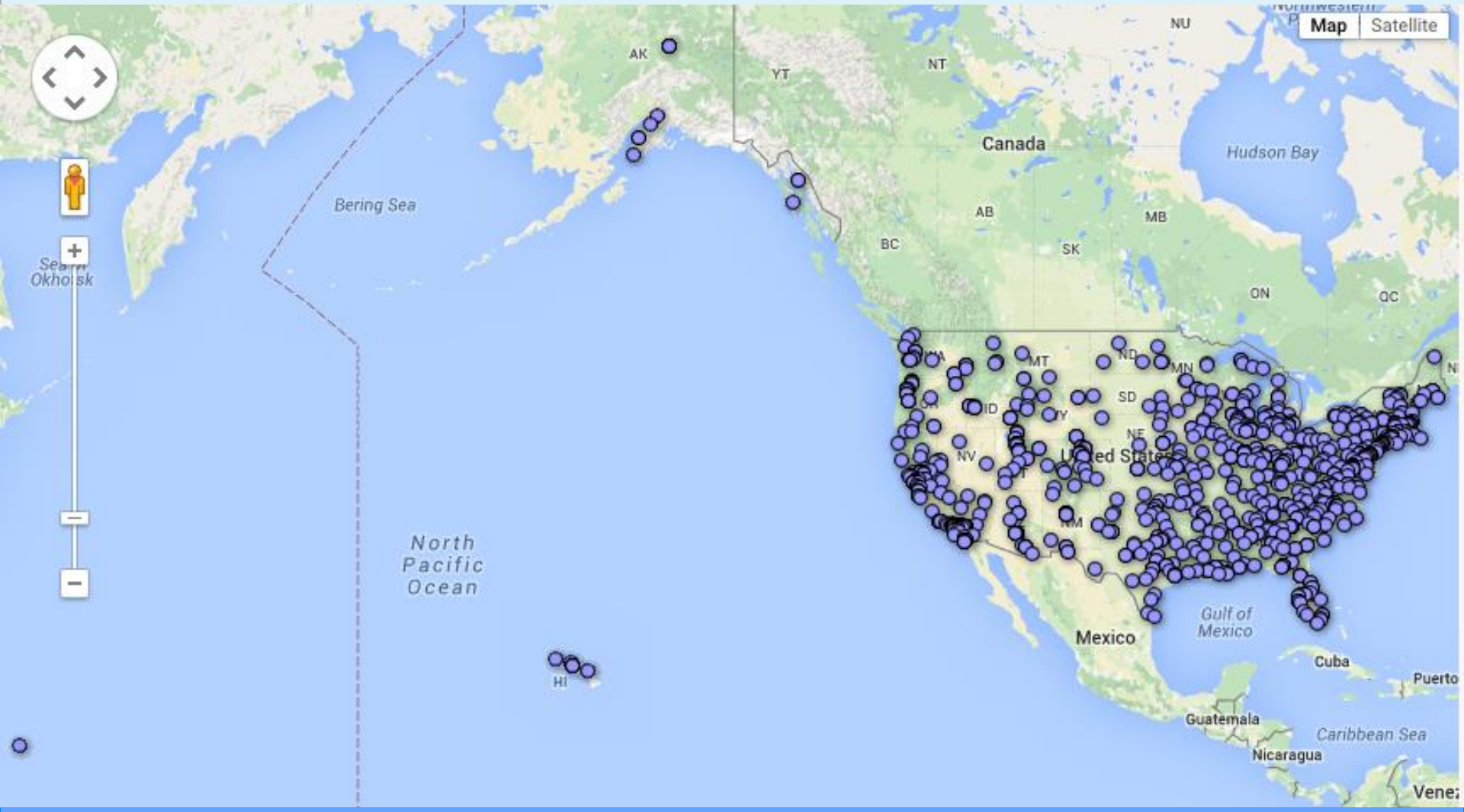
(climate change, invasive species, agricultural pests, etc.)

13 (15) 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*)
- **Great Lakes Invasives**: Documenting the Occurrence through Space and Time of Aquatic Non-indigenous Fish, Mollusks, Algae, and Plants Threatening North America's Great Lakes (*University of Wisconsin Madison*)
- **InvertEBase**: Reaching Back to See the Future: Species-rich Invertebrate Faunas Document Causes and Consequences of Biodiversity Shifts (*Field Museum of Natural History*)
- **The Key to the Cabinets**: Building and Sustaining a Research Database for a Global Biodiversity Hotspot (*Appalachian State University*)

iDigBio
Integrated Digitized Biocollections
National Resource for Advancing Digitization of Biological Collections

NATIONAL HUB, THEMATIC COLLECTION NETWORKS, AND COLLABORATORS



13 TCNS: 200 institutions in 50 states

Advancing Digitization of Biodiversity Collections

Facilitate use of biodiversity data to address environmental and economic challenges

- **Researchers**
- **Educators**
- **General public, citizen scientists**
- **Policy-makers**

- **Develop digitization standards and workflows**
- **Respond to cyberinfrastructure needs**
- **Develop research and outreach collaborations**
- **Plan for long-term sustainability of the national digitization effort**

Advancing Digitization of Biodiversity Collections

Facilitate use of biodiversity data to address environmental and economic challenges

- **Researchers**
- **Educators**
- **General public, citizen scientists**
- **Policy-makers**

- **Develop digitization standards and workflows** **WORKSHOPS, GROUPS**
- **Respond to cyberinfrastructure needs**
- **Develop research and outreach collaborations**
- **Plan for long-term sustainability of the national digitization effort**

Advancing Digitization of Biodiversity Collections

Facilitate use of biodiversity data to address environmental and economic challenges

- **Researchers**
- **Educators**
- **General public, citizen scientists**
- **Policy-makers**

- **Develop digitization standards and workflows**
- **Respond to cyberinfrastructure needs**
- **Develop research and outreach collaborations**
- **Plan for long-term sustainability of the national digitization effort**

Advancing Digitization of Biodiversity Collections

Facilitate use of biodiversity data to address environmental and economic challenges

- **Researchers**
- **Educators**
- **General public, citizen scientists**
- **Policy-makers**

- **Develop digitization standards and workflows**
- **Respond to cyberinfrastructure needs** **SPECIMEN PORTAL at iDigBio**
- **Develop research and outreach collaborations**
- **Plan for long-term sustainability of the national digitization effort**

Advancing Digitization of Biodiversity Collections

Facilitate use of biodiversity data to address environmental and economic challenges

- **Researchers**
- **Educators**
- **General public, citizen scientists**
- **Policy-makers**

- **Develop digitization standards and workflows**
- **Respond to cyberinfrastructure needs**
- **Develop research and outreach collaborations**
- **Plan for long-term sustainability of the national digitization effort**

Advancing Digitization of Biodiversity Collections

Facilitate use of biodiversity data to address environmental and economic challenges

Year 4…….

- **Efficient digitization standards and workflows have been developed**
- **Data are being digitized at 200 institutions funded by NSF/ADBC**
- **398 recordsets**
- **Searchable on the iDigBio portal: 27 million specimen-based records for 47 million specimens and 4.3 million images**
- **Enhanced communication in the NH Collections Community**

Advancing Digitization of Biodiversity Collections

Facilitate use of biodiversity data to address environmental and economic challenges

- ***Where do we go from here … in our 5th - 10th years?***

- **Emphasis on digitization will continue (this is what distinguishes…)**
- **Increased emphasis on:**
  1. **Data use: research and outreach collaborations**
  2. **Data management (attribution, corrections, annotations)**
  3. **Sustainability of the national digitization effort (RCN)**

Advancing Digitization of Biodiversity Collections

Facilitate use of biodiversity data to address environmental and economic challenges

- ***Where do we go from here … in our 5th - 10th years?***

- **A proposal for renewal of support for iDigBio….years for 6-10**
- **More TCNs will be funded**
- **Other digitization efforts in U.S.A.**

Advancing Digitization of Biodiversity Collections

Facilitate use of biodiversity data to address environmental and economic challenges

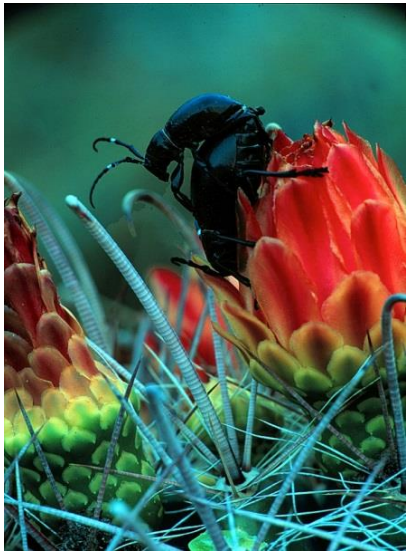
- **Where do we go from here … in our 5th - 10th years?**

  International collaborations

Opportunities

- **New research and education initiatives**

- Millions (billions?) of records – huge source of biodiversity data
- Opportunities for previously intractable large-scale research requiring large amounts of spatial and temporal data



→ Big-science questions related to human health, climate change, agriculture, species discovery, species extinctions, rates of evolutionary change, and ecosystem services


BIODIVERSITY
Thank you
Photos: M. Jeffords & G. Paulay