Digitization of Biodiversity Collections in the U.S.A.

Greg Riccardi
SiBBr Workshop on Digitization
25 November 2014
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.
Where is iDigBio?
Biodiversity Informatics

The single largest source of information on biological diversity.

*1,500 natural history collections*

1 billion specimens in USA

3 billion specimens globally
Step 1: Make a plan and get funding

• Commission a report to demonstrate need
  – Network Integrated Biocollections Alliance
    • NIBA strategic plan
  – Called for establishing a national program to fund digitization

• Fund a program at the National Science Foundation
  – Advancing Digitization of Biodiversity Collections (ADBC)
Advancing Digitization of Biodiversity Collections (ADBC)

- The goals of ADBC are
  - to remove this inaccessibility through digitization
  - to put 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 U.S. collections
Step 2: Call for proposals

• Thematic Collections Networks
  – Digitize specimen information to enable research
  – Allow community to decide priorities

• Home Uniting Biocollections (HUB)
  – Enable digitization
  – Provide access to data
  – Build communities of practice
  – Reach out to communities
Thirteen Thematic Collections Networks

- **InvertNet**: 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 Anthropods 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
- **Vouchered Animal Communication Signals**
- **The Macroalgal Herbarium Consortium**: Accessing 150 Years of Specimen Data to Understand Changes in the Marine/Aquatic Environment
- **Aquatic Non-indigenous Fish, Mollusks, Algae, and Plants** Threatening North America's Great Lakes
- **InvertEBase**: Reaching Back to See the Future: Species-rich Invertebrate Faunas Document Causes and Consequences of Biodiversity Shifts
- **The Key to the Cabinets**: Building and Sustaining a Research Database for a Global Biodiversity Hotspot
Advancing Digitization of Biodiversity Collections (ADBC)

To date: 10 TCNs, 152 unique institutions, ~276 projects, 50 states

Developing networks, enhancing communication, facilitating collaboration
Where to go from here?

• How to sustain activities?
  – Digitization projects?
  – Support for digitization improvements?
• How to sustain data infrastructure?
  – Data persistence?
  – Data quality?
  – Data portal?
• How to sustain commitment
  – Governmental?
  – Community?