University of Florida
Florida State University

www.iDigBio.org
Advancing Digitization of Biodiversity Collections

• Facilitate use of biodiversity data to address environmental and economic challenges
  ▪ Researchers
  ▪ Educators
  ▪ General public
  ▪ Policy-makers

• Enable digitization of biodiversity collections data
  ▪ Develop efficient and effective digitization standards and workflows
  ▪ Respond to cyberinfrastructure needs

• Provide portal access to biodiversity data in a cloud-computing environment

• Plan for long-term sustainability of the national digitization effort
  ▪ Expand participation: partners and data sources
Seven Thematic Collections Networks (TCNs)

- **InvertNet**: An Integrative Platform for Research on Environmental Change, Species Discovery and Identification (*Illinois Natural History Survey, University of Illinois*) [http://invertnet.org](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](http://tcn.amnh.org)


- **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*) [http://hasbrouck.asu.edu/symbiota/portal/index.php](http://hasbrouck.asu.edu/symbiota/portal/index.php)
National Resource (iDigBio), Thematic Collection Networks (TCNs), and Collaborators

7 TCNs with 130 participating institutions
Building the iDigBio Cloud

- Cloud-based strategy
  - Providing useful services/APIs (programmatic and web-based Application Programming Interface)
  - Federated scalable object storage and information processing
  - Digitization-oriented virtual appliances
  - Reliance on standards, proven solutions and sustainable software

- Continuous consultation with stakeholders
  - Surveys, working groups, workshops, person-to-person

![Diagram of iDigBio System Architecture]
What Makes iDigBio Unique?

- Ingest all contributed data with emphasis on GUIDs, not only a restricted set of selected data elements
- Maintain persistent datasets and versioning, allowing new and edited records to be uploaded as needed
- Ingest textual specimen records, associated still images, video, audio, and other media
- 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
- Facilitate sharing and integration of data relevant to biodiversity research
- Provide computational services for biodiversity research
Recent and Ongoing Activities

- Assessment of common and effective practices (paper in ZooKeys)
- Minimum information for scientific collections working group
- Collaborative georeferencing pilot project at Godfrey Herbarium
- Digitization workflows working groups
- Public Participation in Digitization of Biodiversity Specimens workshop
- Georeferencing working group & train-the-trainers workshop
- OCR/natural language processing working group
- Linked data workshop
- Series of digitization training workshops
- Call for appliances
- Call for working groups
- Cyberinfrastructure working group
- Specimen data portal v0 implementation
- Server hosting

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 NSF.