Introduction to iDigBio, Survey of TCNs, PENs, RDCNs Integrated Digitized Biocollections

Joanna McCaffrey

(jmccaffrey@flmnh.ufl.edu)

(Florida Museum of Natural History, iDigBio)

20 June 2013 SPNHC, Rapid City, South Dakota



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.



Seven Thematic Collections Networks (TCNs) plus 2 Partner to Existing Networks (PENs)

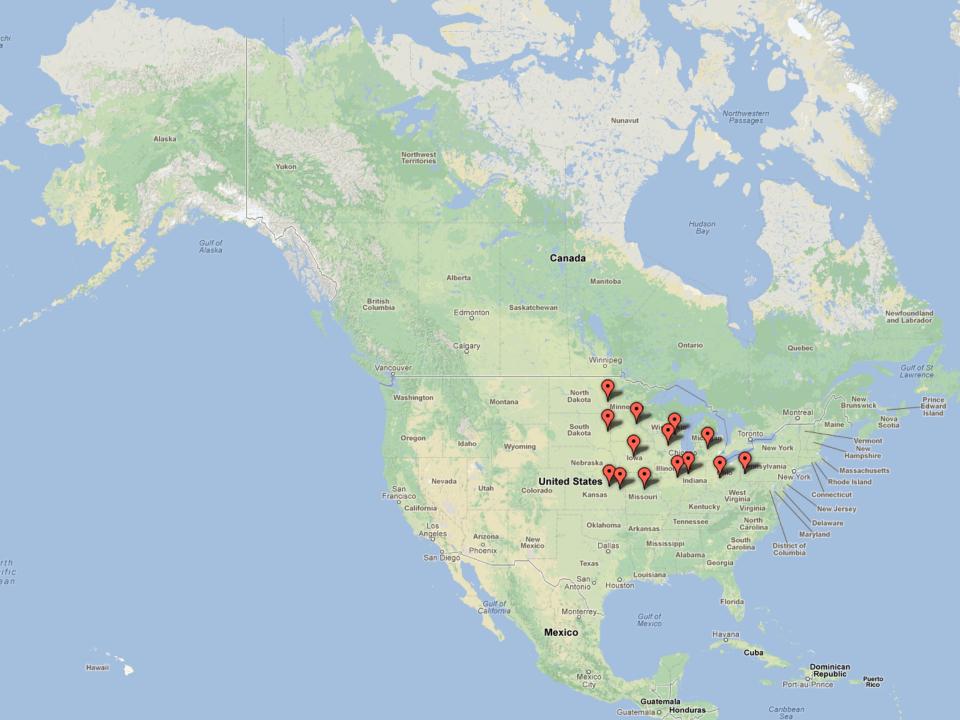
- InvertNet: An Integrative Platform for Research on Environmental Change, Species Discovery and Identification (*Illinois Natural History Survey, University of Illinois*) http://invertnet.org
- Plants, Herbivores, and Parasitoids: A Model System for the Study of Tri-Trophic Associations (American Museum of Natural History) http://tcn.amnh.org
- North American Lichens and Bryophytes: Sensitive Indicators of Environmental Quality and Change (*University of Wisconsin – Madison*) http://symbiota.org/bryophytes/index.php (plus 2 PENs)
- 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 Anthropods Network (SCAN): A Model for Collections Digitization to Promote Taxonomic and Ecological Research (Northern Arizona University) http://hasbrouck.asu.edu/symbiota/portal/index.php

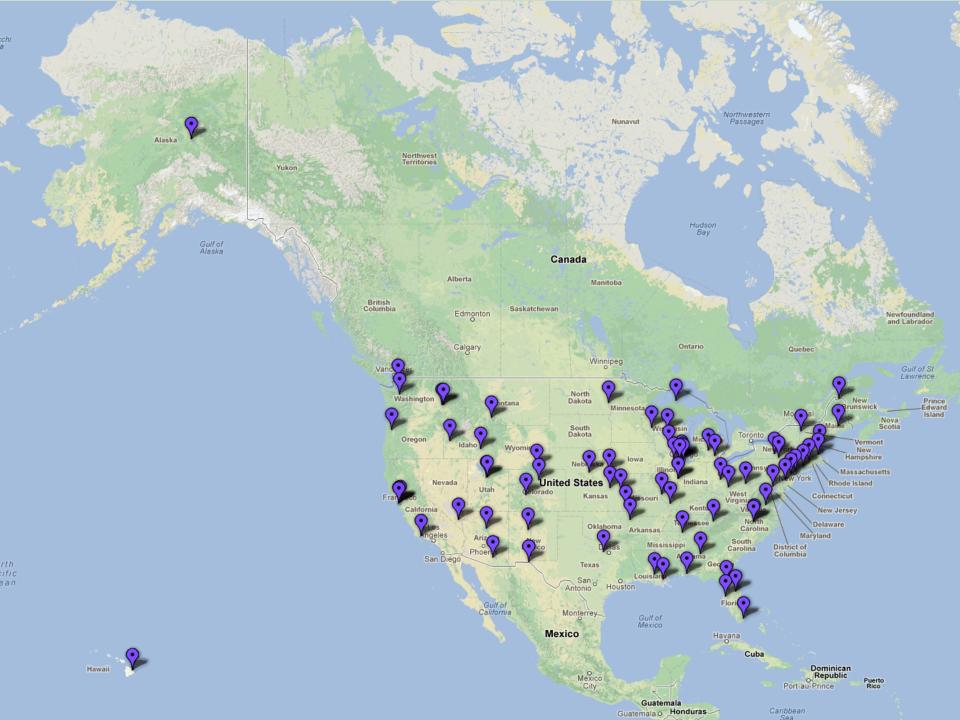
National Resource (iDigBio), Thematic Collection Networks (TCNs)

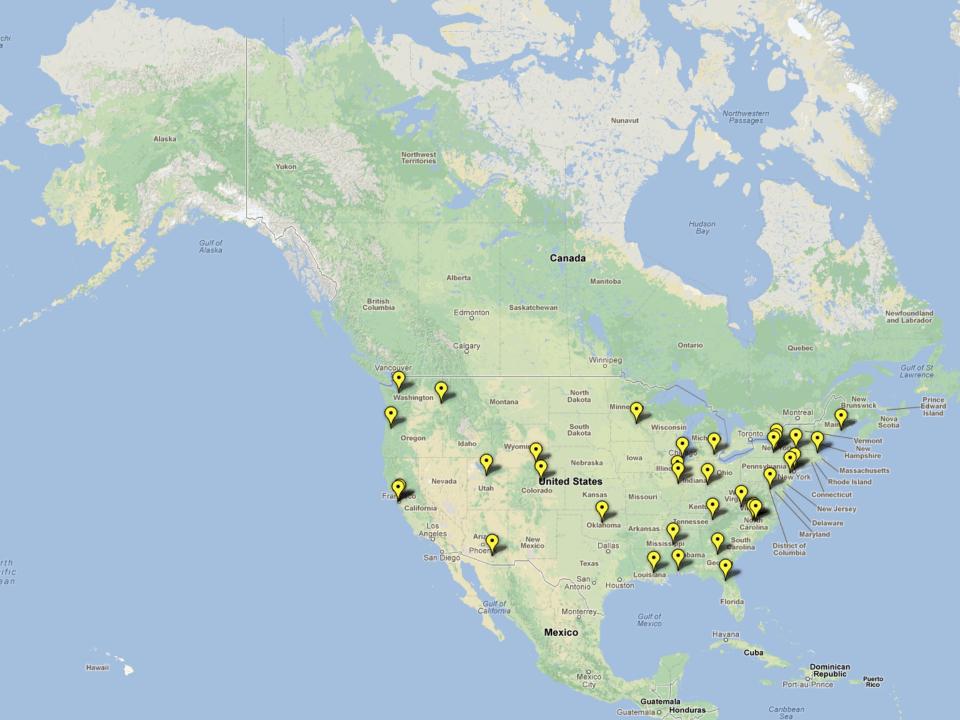


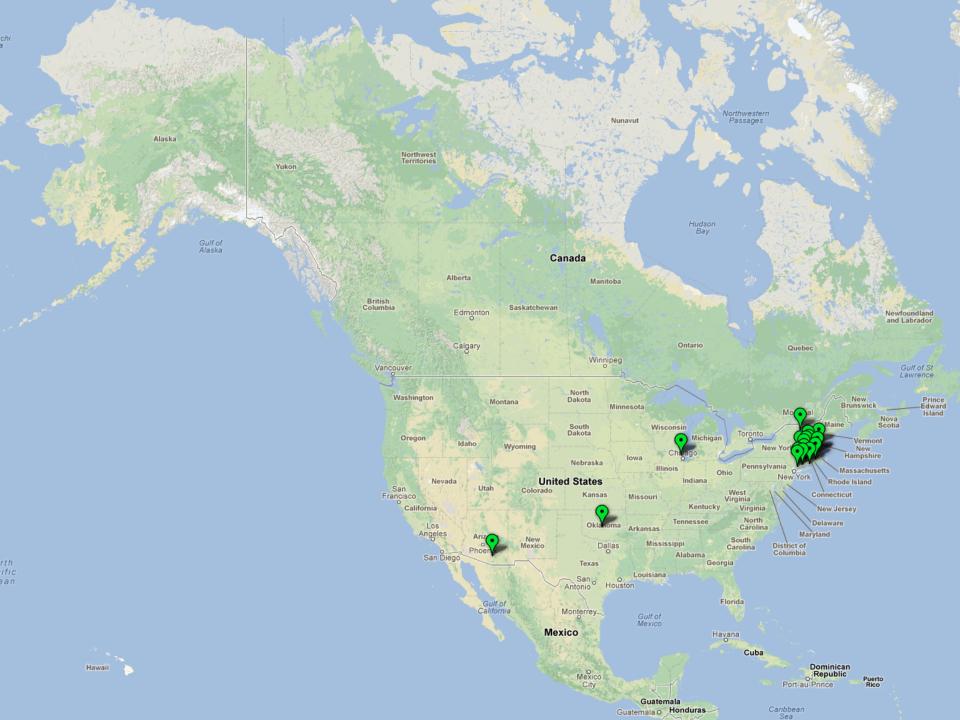


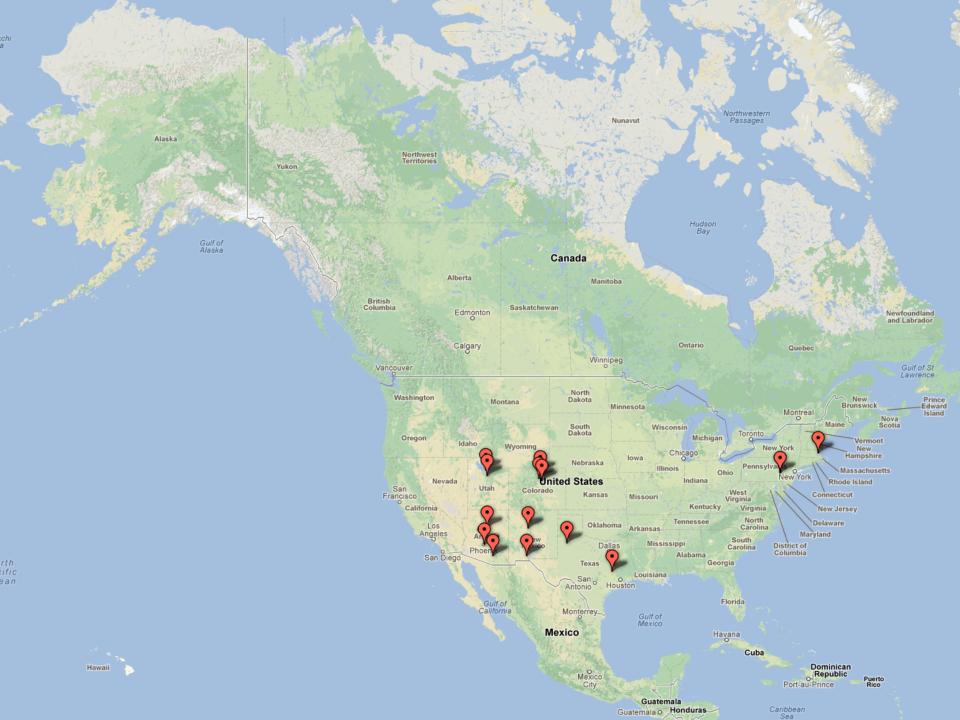
To date: 7 TCNs, 2 PENs, 143+ participating institutions, 49 states – more coming

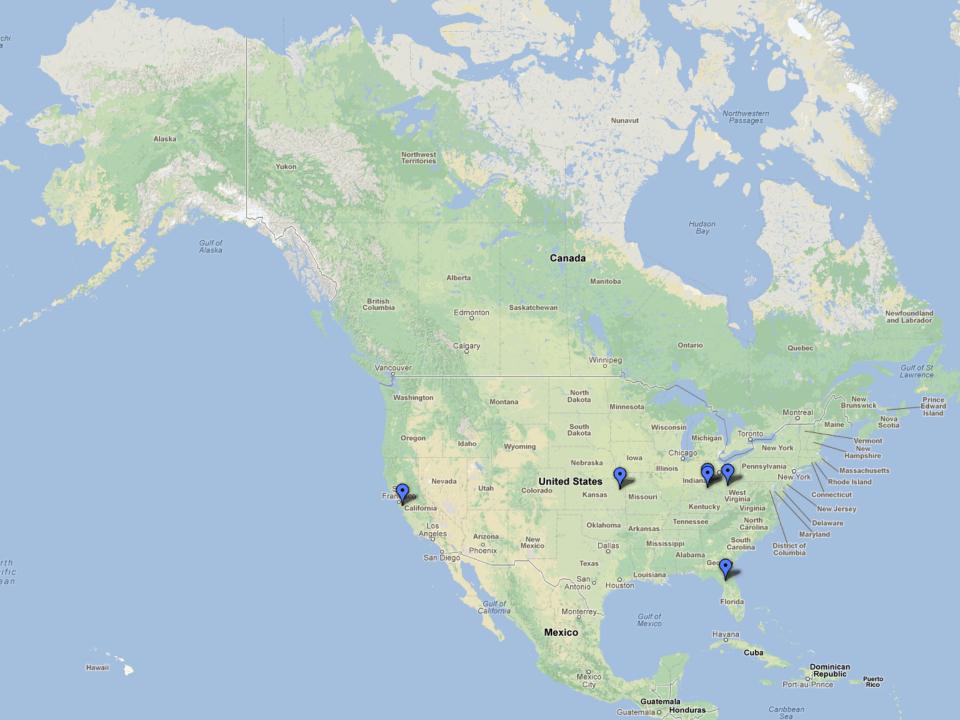


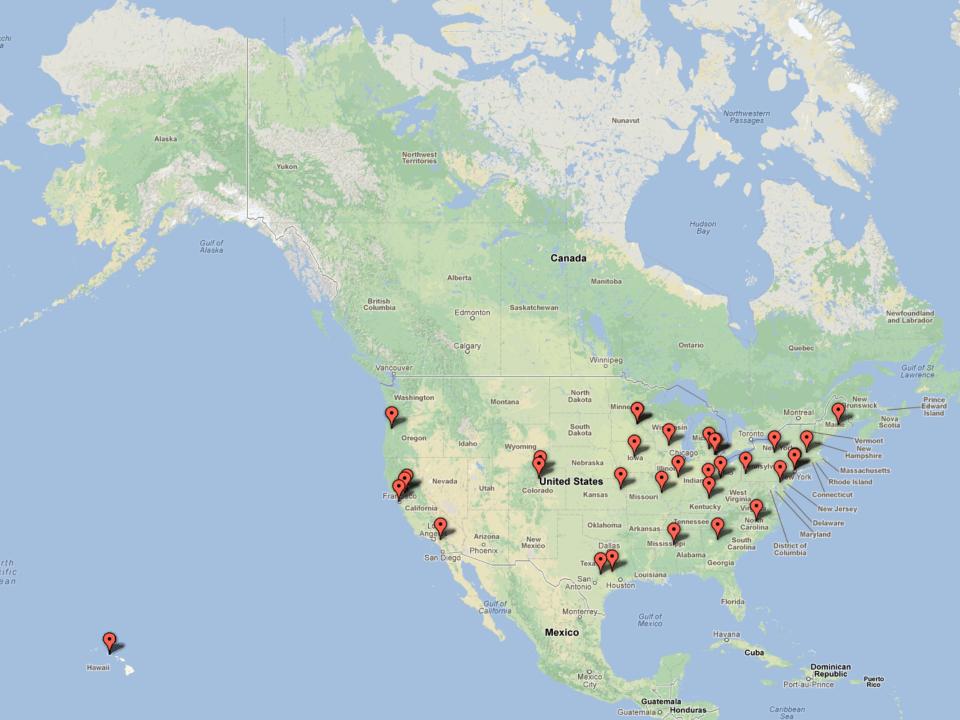












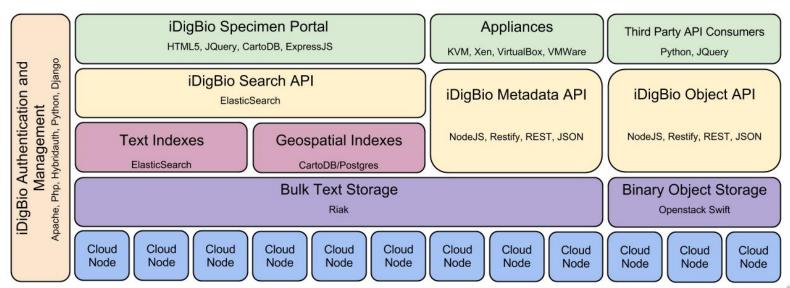
R (?) Digitization Collection Networks (RDCNs)

- When the thematic concept does not fit your proposal
- Administered by CSBR
- Here are some current ones -
 - Consortium of California Herbaria
 - Consortium of Pacific Northwest Herbaria
 - Mississippi Herbaria Consortium: Magnolia grandiFLORA
 - Georeferencing US Fish collections



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





What Makes iDigBio Unique?

- Ingest all contributed data with emphasis on use of 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, plus 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 digitization practices (paper in ZooKeys)
- Working groups
 - Minimum information for scientific collections working group (MISC)
 - Digitization workflows working groups
 - Biodiversity Informatics Manager working group
- Workshops year 2:
 - > 150 institutions, 9 workshops, 3 symposia
 - 368 sponsored participants
 - Video archives on Vimeo, live streaming for remote participation
 - New model this year: train the trainer
 - Series of digitization training workshops (herbaria, wet collections, entomology, paleontology, fluid-preserved invertebrate imaging)
- Server hosting: 8 virtual machines, TCN support
- Specimen data portal and website improvements, and more on the way
- Call for appliances, frequent opinion surveys















Advancing Digitization of Biodiversity Collections

- Facilitate use of biodiversity data to address environmental and economic challenges
 - Researchers
 - Educators
 - General public, citizen scientists
 - Policy-makers



- Enable digitization of non-federal US 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
- Develop research, education, and outreach collaborations
- Plan for long-term sustainability of the national digitization effort

