The Southeast US as a regional model for global studies

Zack Murrell

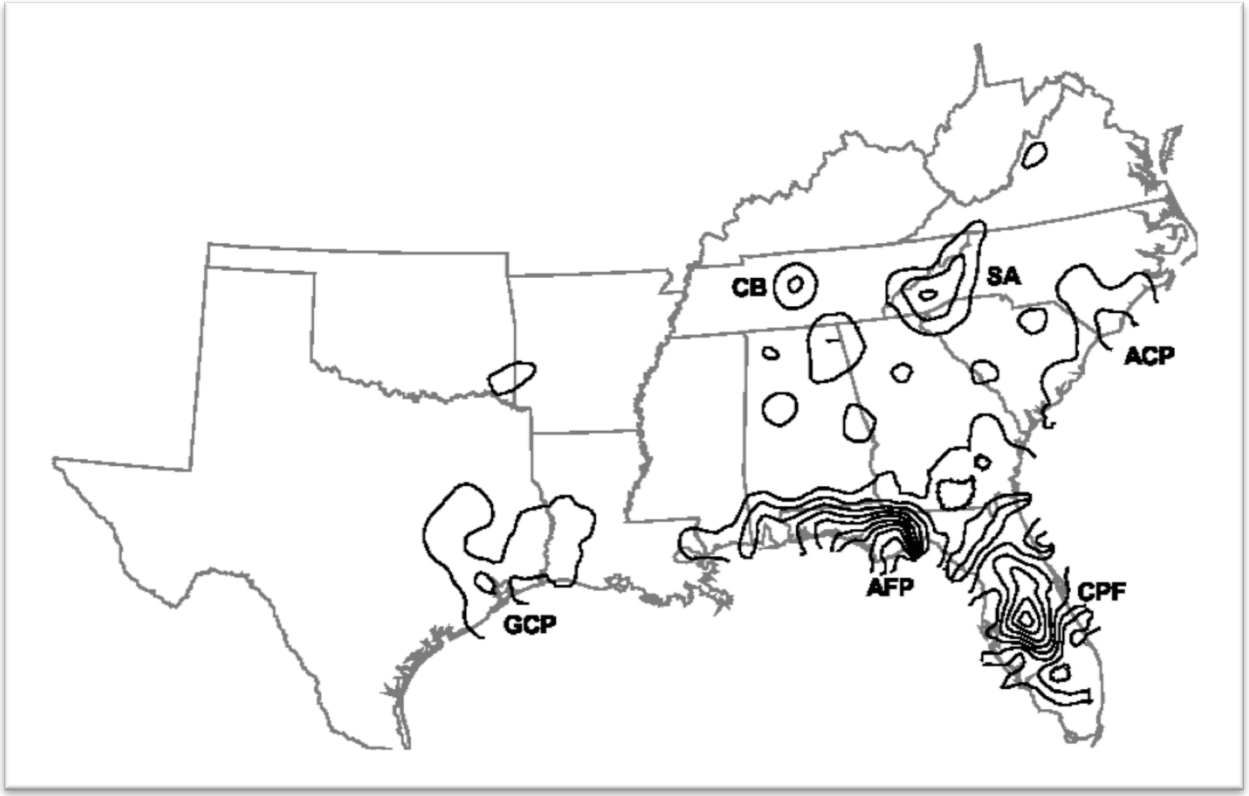
Appalachian State University

Using Biodiversity Specimen-Based Data to Study Global Change Workshop

Missouri Botanical Garden, St. Louis

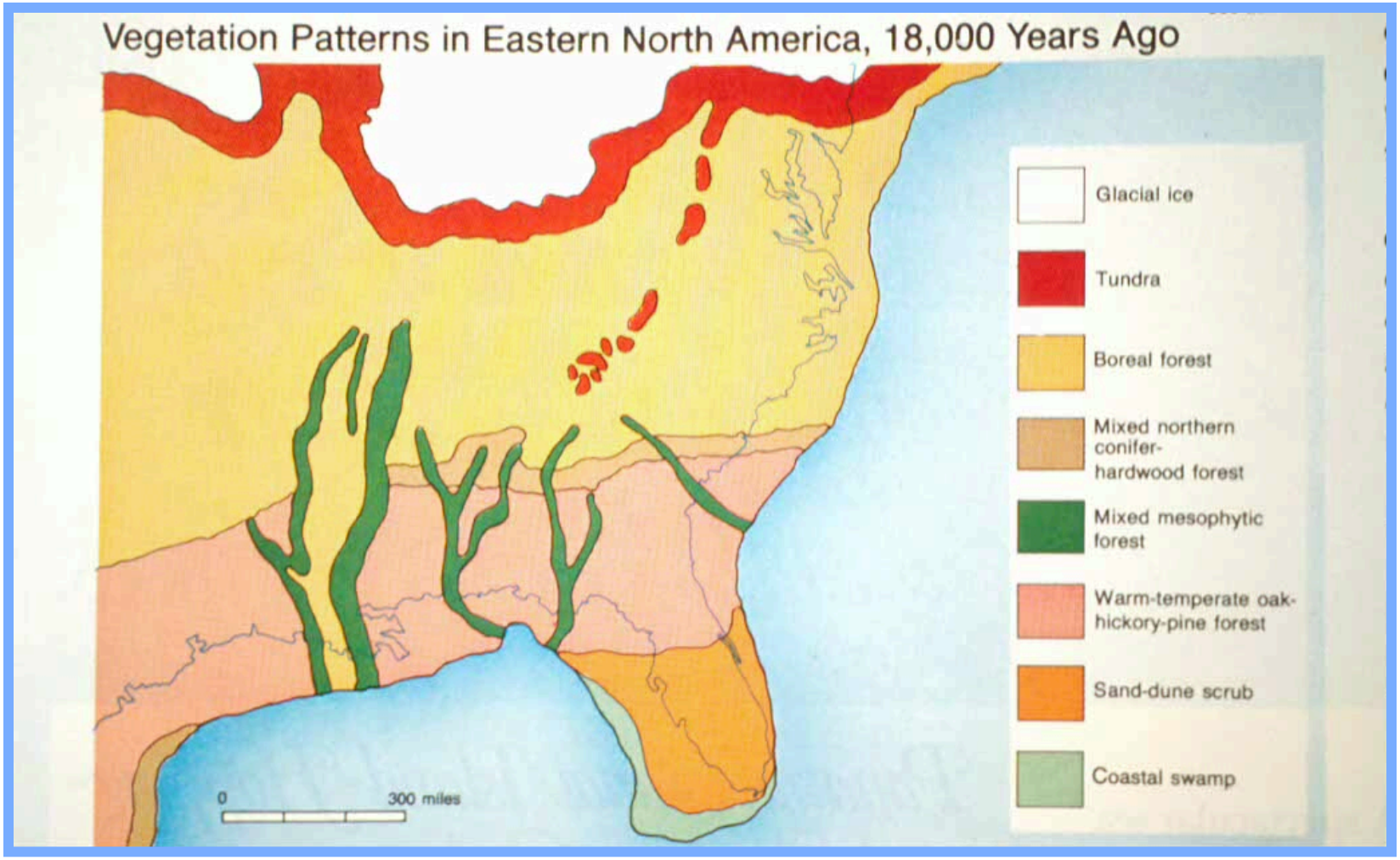
2-3 December 2015

Why: A biodiverse region with intense human population pressures

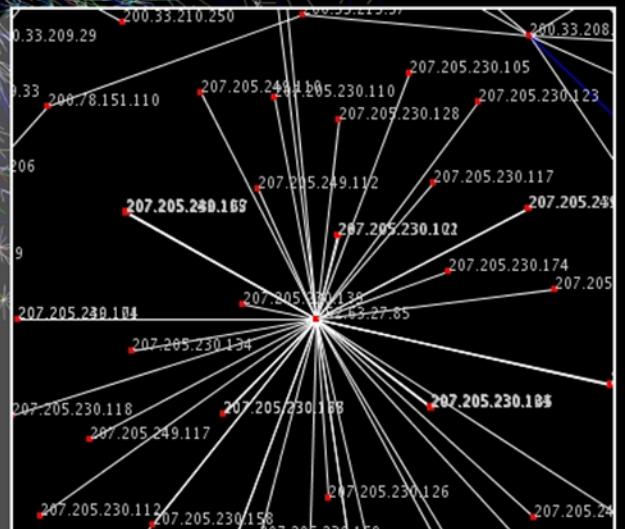


Hot spots of plant endemism in the southeast (Estill & Cruzan 2001)

Why: Rich tradition of historical biogeographic study & collecting ∴ potential to forecast future distributions



Electronic communication and storage now allows us to move information at breathtaking speeds across space and time, opening up new opportunities to aggregate and parse large parcels of information



"As access to information dramatically expands, so that people increasingly have access to almost all the information they might need at any time and in any place (and, surprisingly, at low or no cost), *the value of the cognitive skills still unreplicable by silicon becomes greater."*

But this value cannot be recognized without providing appropriate access for biologists.

L. Prusak, Where did knowledge management come from?, IBM Systems Journal Volume 40, Number 4, 2001 [“Knowledge Management”], http://www.research.ibm.com/journal/sj/404/prusak.html

Cyberinfrastructure

A combination of databases, network protocols and computational services that brings people, information and computational tools together to perform science in this information-driven world.

Lincoln D. Stein

www.nature.com/reviews/genetics

Cyberinfrastructure

- Data infrastructure and storage
- Computational infrastructure
- Communication infrastructure
  - Low-level (connectivity and bandwidth)
  - Syntactic (common formats to organize data)
  - Semantic (common terminology/ontology)
- Human infrastructure
  - Sharing and curation

Lincoln D. Stein

www.nature.com/reviews/genetics

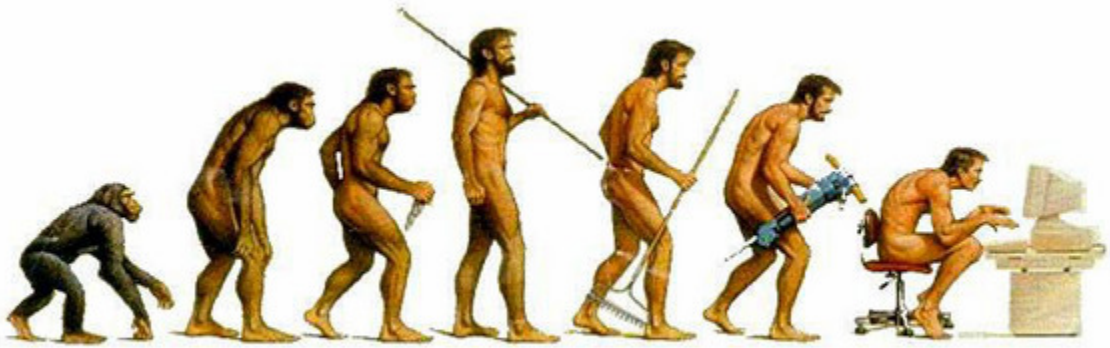
We are drowning in information, while starving for wisdom. The world henceforth will be run by **synthesizers**, people able to put together the right information at the right time, think critically about it, and make important choices wisely.

E.O. Wilson

Consilience: The Unity of Knowledge (1998)

Selfish and contentious people will not cohere, and without coherence nothing can be effected.

Charles Darwin Descent of Man. 1871



Principles of the Conservation Commons

1. Promote Free and Open Access

2. Mutual Benefit (Reciprocity)

3. Rights and Responsibilities
   - ❖ Attribution
   - ❖ Integrity of Original Work

http://www.conservationcommons.net/

The Key to the Cabinets: Building and Sustaining a Research Database for a Global Biodiversity Hotspot

A new NSF-Advancing Digitization in Biological Collections (ADBC) collaborative project in the southeast

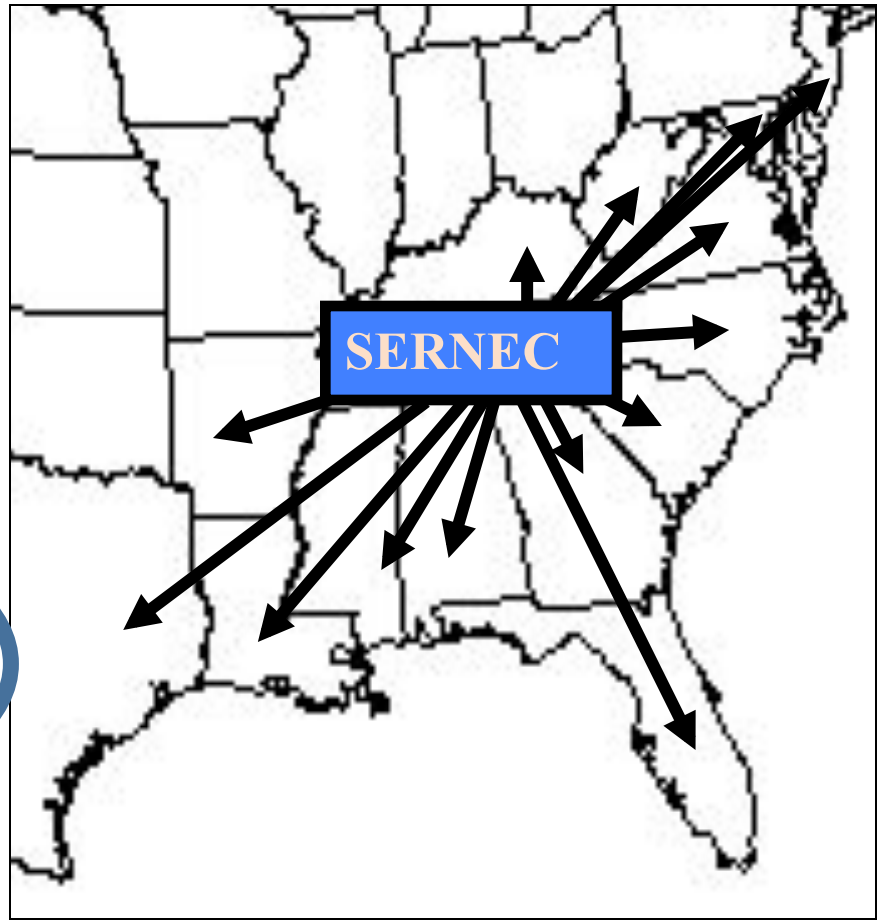
NSF ADBC #1410069 (plus collaborative awards (1410077 1410081 1410086 1410087 1410092 1410094 1410098 1410143 1410200 1410288 1410439 1410445)

Appalachian
STATE UNIVERSITY
BOONE, NORTH CAROLINA 28608

SERNEC

Southeast Regional Network of Expertise and Collections

- 2005 NSF Research Coordination Network (RCN) awarded to build the 'human infrastructure' to begin the '2020 Vision effort'
- 5-year project
- 150 herbaria participated
- Used existing infrastructure of the Association of Southeastern Biologists

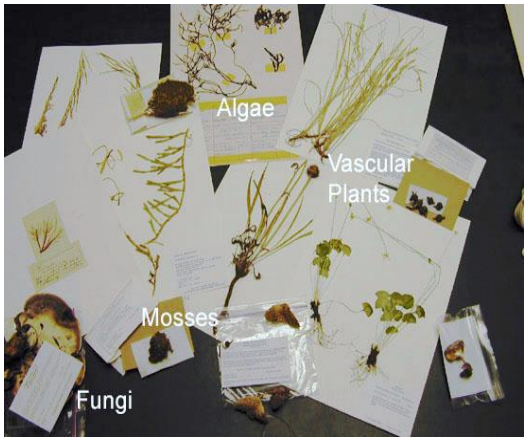


What is a herbarium?

Organized collections of carefully preserved and documented plant specimens

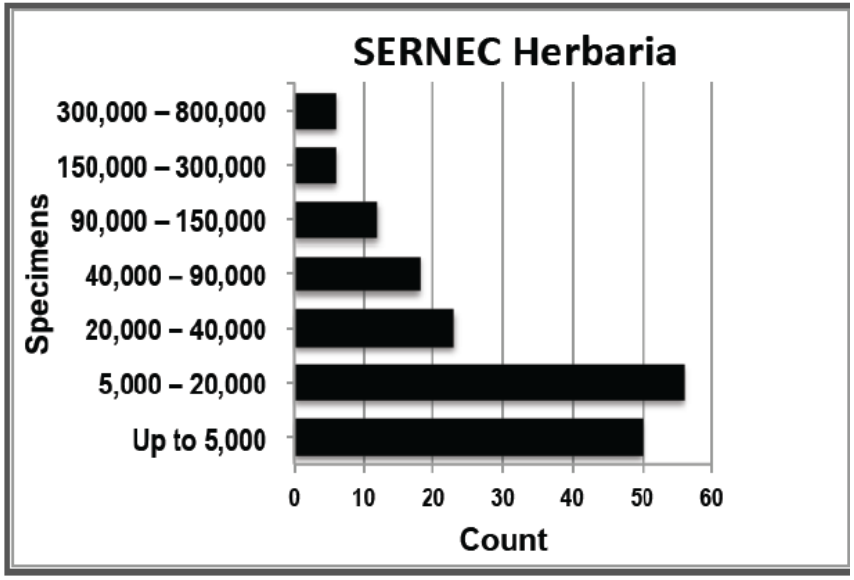
-plus-

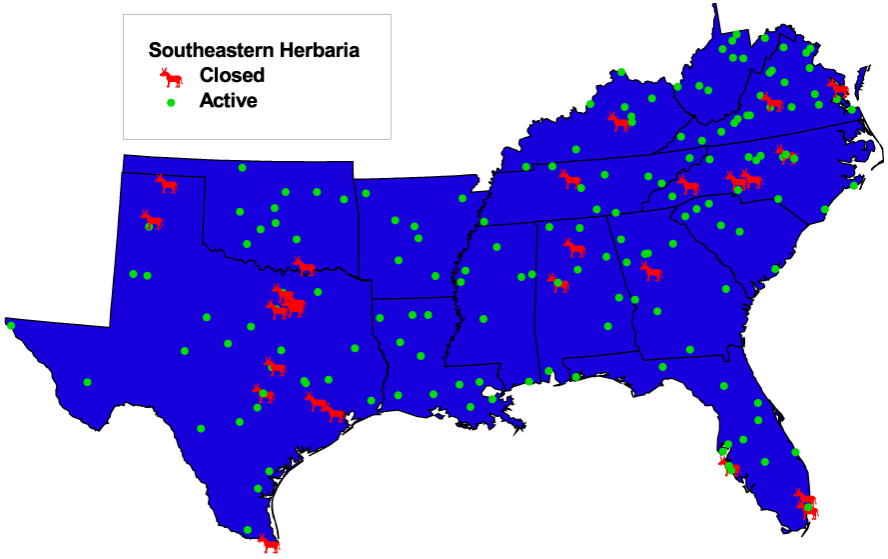
Associated resources such as curators, taxonomic specialists, collections managers, students, floras, monographs, microscopes, and curation supplies.



State of Herbaria in the Southeast USA

- 233 collections vary in size from 3,000 to 750,000 specimens.
- Curators are taxonomists, ecologists, paleobotanists, population biologists.
- Some institutions with significant budgets, IT support and curatorial staff.
- Some institutions are single person operations without budgets or IT support.
- ESTIMATED 15 MILLION SPECIMENS FROM THE SOUTHEAST.

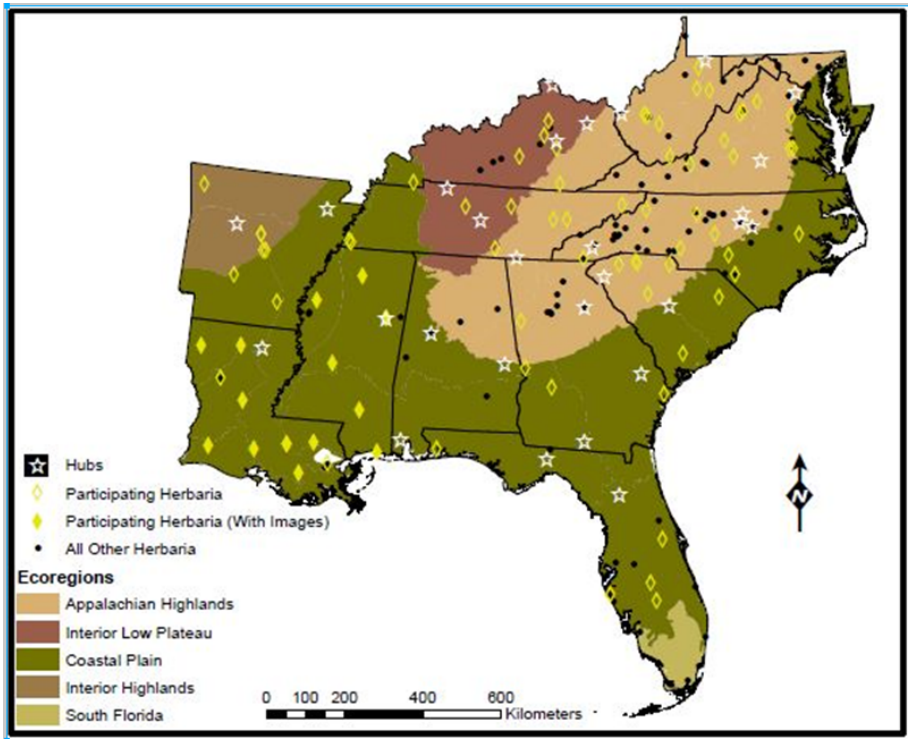


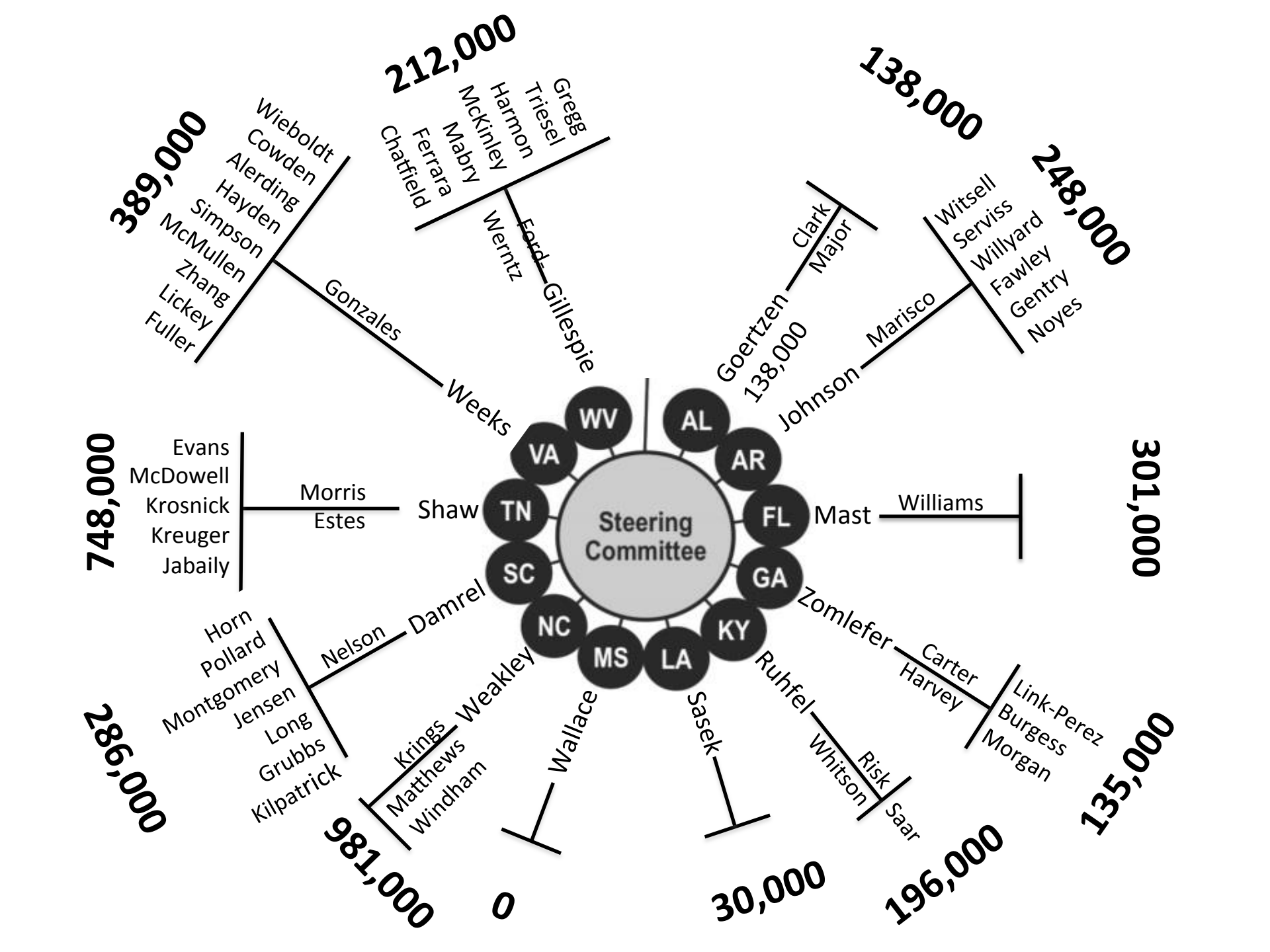


Data derived from Index Herbariorum (http://www.nybg.org/bsci/ih/ih.html) .

Who is involved in the ADBC project?

- 93 herbaria (plus 12)
- IT (Symbiota, GEOLocate, SPECIFY, iPLANT-TACC)
- Citizen scientists
  - Notes From Nature
  - Adler Museum
- Education experts
- Herbarium Affiliates
- State Natural Heritage Programs
- iDigBio



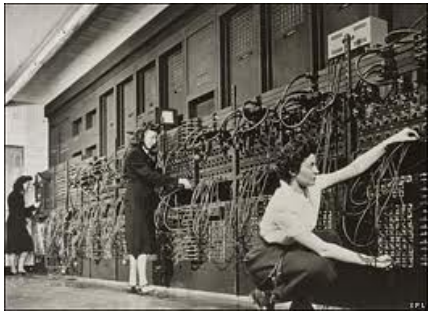

Steering Committee
WV
AL
VA
AR
TN
FL
SC
GA
NC
KY
MS
LA
212,000
Gregg
Triesel
Harmon
McKinley
Mabry
Ferrara
Chatfield
Ford
Werntz
Gillespie
389,000
Wieboldt
Cowden
Alerding
Hayden
Simpson
McMullen
Zhang
Lickey
Fuller
Gonzales
Weeks
138,000
Clark
Major
Goertzen
138,000
248,000
Witsell
Serviss
Willyard
Fawley
Gentry
Noyes
Marisco
Johnson
748,000
Evans
McDowell
Krosnick
Kreuger
Jabaily
Morris
Estes
Shaw
301,000
Williams
Mast
286,000
Horn
Pollard
Montgomery
Jensen
Long
Grubbs
Kilpatrick
Nelson
Damrel
Zomlefer
135,000
Carter
Harvey
Link-Perez
Burgess
Morgan
981,000
Krings
Matthews
Windham
Weakley
0
Wallace
30,000
Sasek
196,000
Ruhfel
Whitson
Risk
Saar

Data Management Workflow

Specimen Image Capture
"Herbarium Digitarium"

TACC

Texas Advanced Computing Center



Archival Storage

4,700,000 SPECIMENS
100+ INSTITUTIONS
12 STATES

Web Derivatives

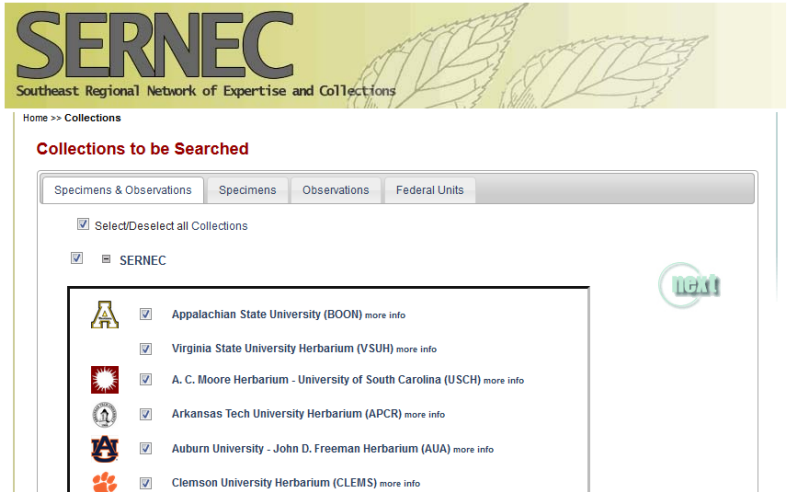
Citizen Scientists

Metadata to Portal

GEOLocate

Batch Georeferenced Data to Portal

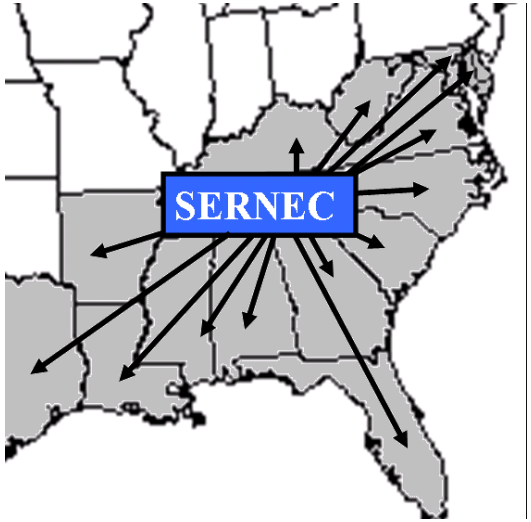
Repatriation of Metadata to Local Institution Symbiota and Specify



MOBILIZATION AND COMMUNICATION

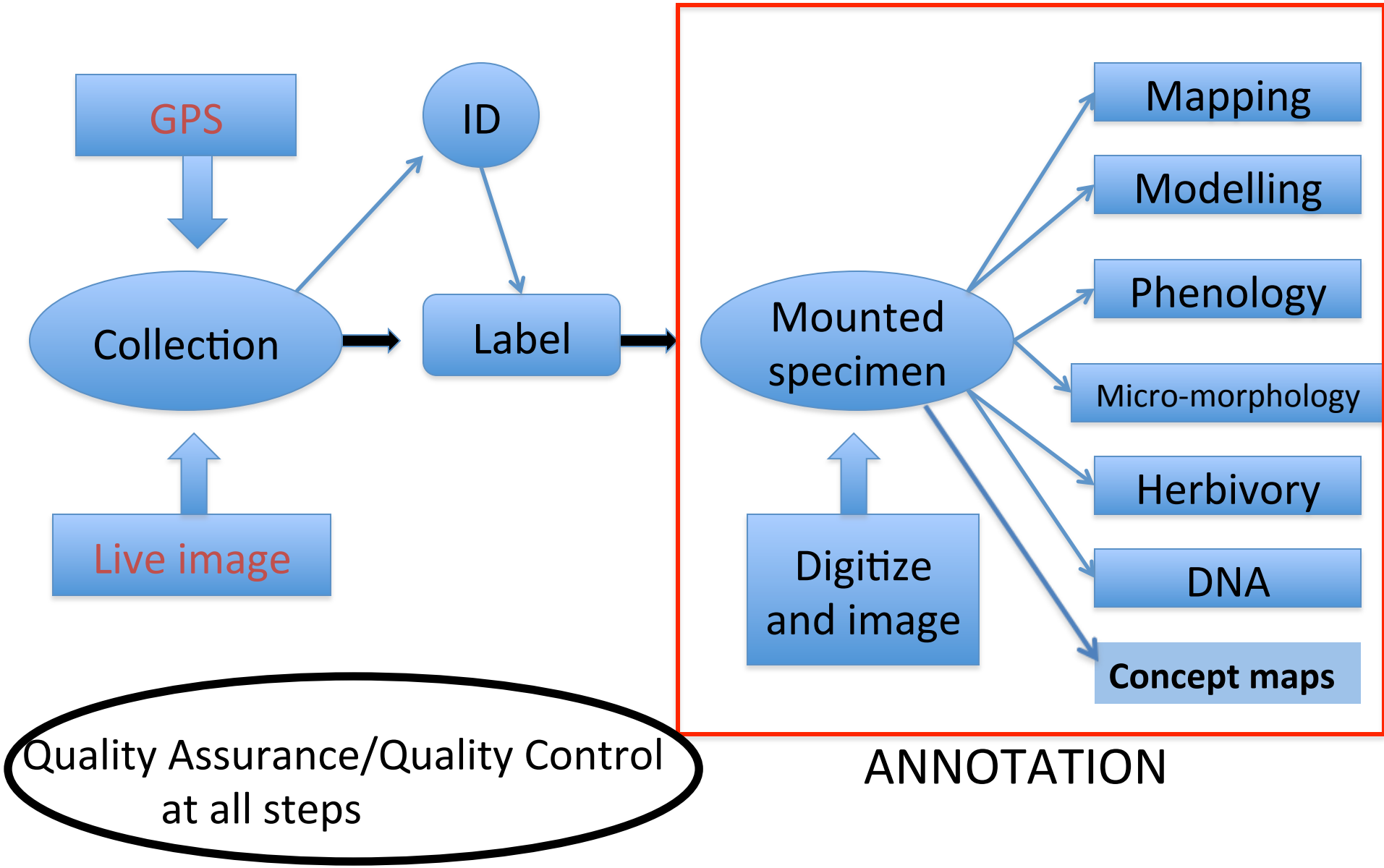
WORKING GROUPS:

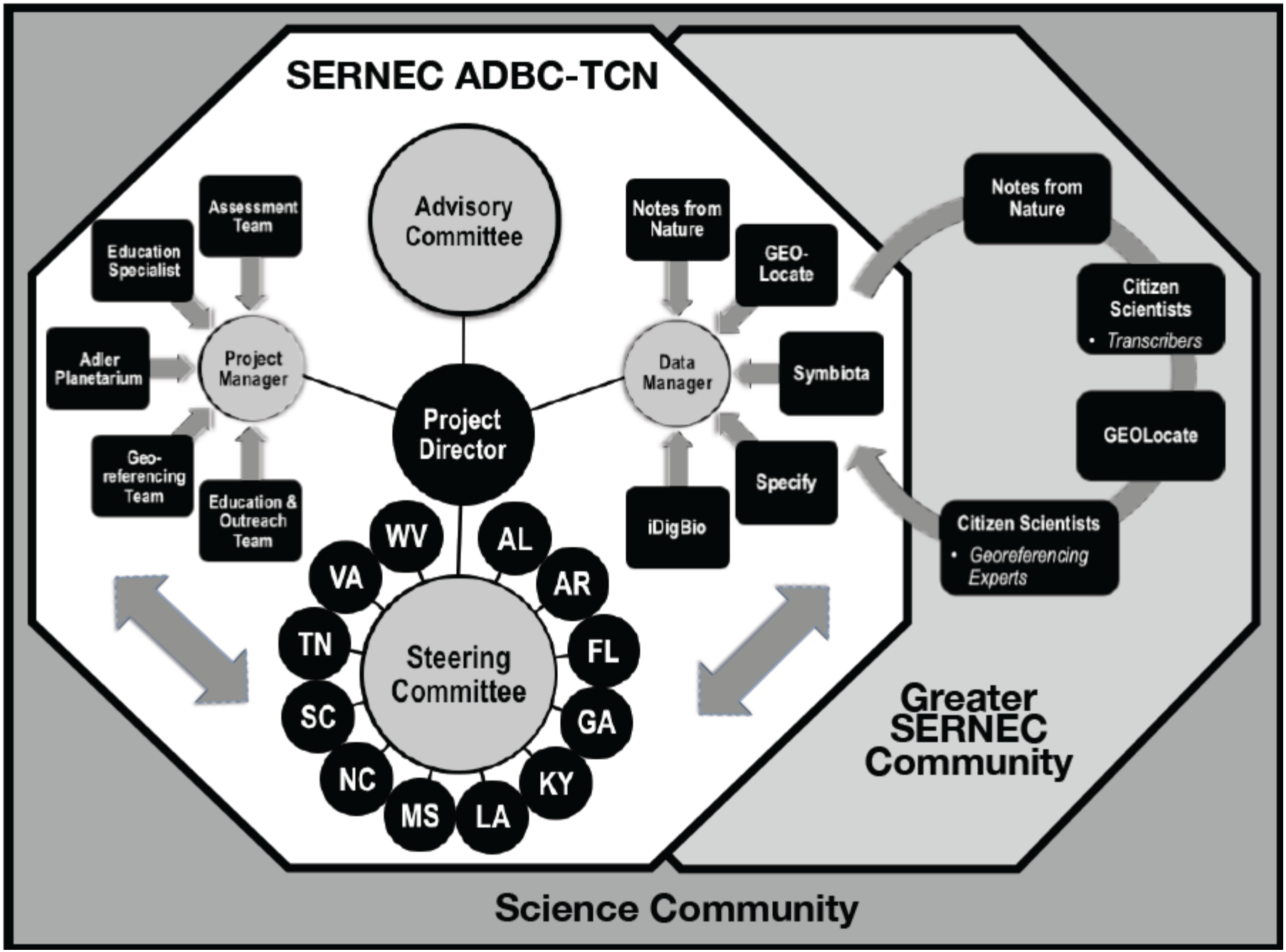
1. IT and infrastructure
2. Libraries and literature
3. Images (specimen and live)
4. Education and outreach
5. Taxonomic concept mapping
6. State working groups



"IT TAKES A VILLAGE"

Steps to increasing the value of natural history collections




SERNEC ADBC-TCN
Advisory Committee
Assessment Team
Education Specialist
Adler Planetarium
Project Manager
Geo-referencing Team
Education & Outreach Team
Project Director
Steering Committee
WV
AL
VA
AR
TN
FL
SC
GA
NC
KY
MS
LA
Notes from Nature
GEO-Locate
Data Manager
Symbiota
Specify
iDigBio
Notes from Nature
Citizen Scientists
• Transcribers
GEOLocate
Citizen Scientists
• Georeferencing Experts
Greater SERNEC Community
Science Community

Governance structure based on the Society of Herbarium Curators constitution

State consortia

Regional chapters of SHC

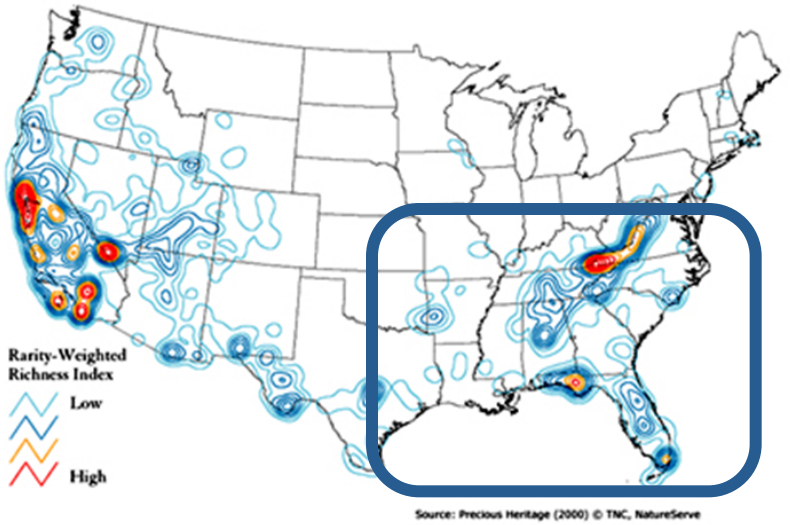
Decision-making
Consensus building
Conflict resolution
Knowledge transfer
Longevity
Sustainability

What are the ADBC project goals?

Use our robust human infrastructure

1) improve protocols at every stage of specimen digitization.

2) interface with citizen scientists to develop a strong collaborative community that will sustain high transcription and georeferencing rates through a multi-year effort.



What works?

1. Use existing infrastructures (Association of Southeastern Biologists)
2. Provide a reward system of making the group's aggregated data available for their own manipulation
3. Transformational leadership?
4. Communication???

What can biologists do?

1. Build our communities (include amateurs and affiliates).
2. Embrace and promulgate new technologies.
3. Take part in the development of community best practices (among data creators, users, editors and publishers).
4. Expand the community to include computer, social and library scientists.
5. Generate a paradigm shift in how productivity is evaluated by administrators.
6. Utilize more effective means of communication.

Snapshot versus long-term shift in best practices. We need to train biologists in traditional studies of natural history, while at the same time encouraging the use of new technologies.

SERNEC: SouthEast Regional Network of Expertise and Collections

National Science Foundation
Research Coordination Network
RCN Award # 0542320

National Science Foundation
Advancing Digitization of
Biological Collections
Award # 1410069