Challenges and Efficiencies of organizing an ADBC TCN project on Southeast Freshwater Macrofauna

Hank Bart

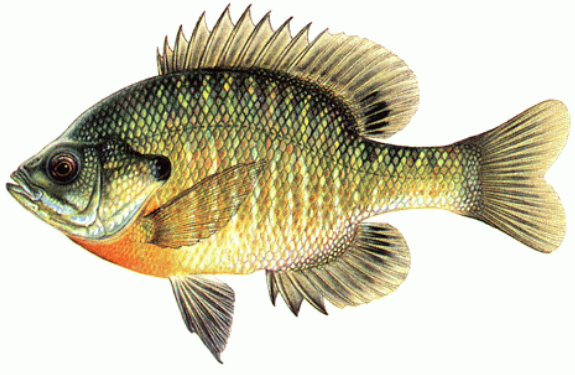
Tulane University Biodiversity Research Institute

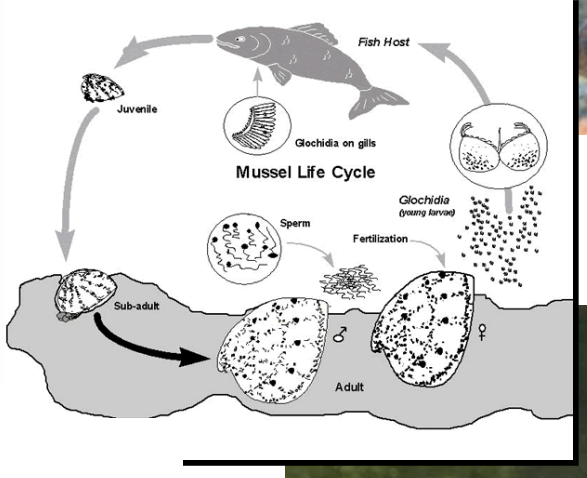
hbartjr@tulane.edu

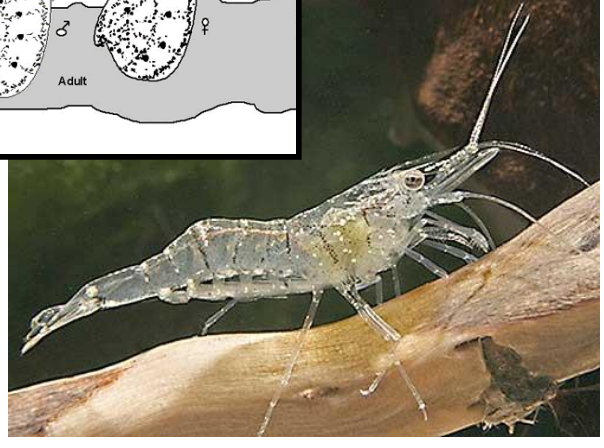
Proposed TCN Concept

- Organizing a TCN project on Freshwater Macrofauna of the Southeastern U.S.
- “Macrofauna” = retained on a 0.5mm sieve. (mollusks, decapod crustaceans, fishes).
- Research theme: impacts of river modification (channelization and impoundments) on diversity and complexity of macrofauna communities.
- TCN on aquatic insects led by Clemson (MOSAIC).

Diverse SE Macrofauna Communities







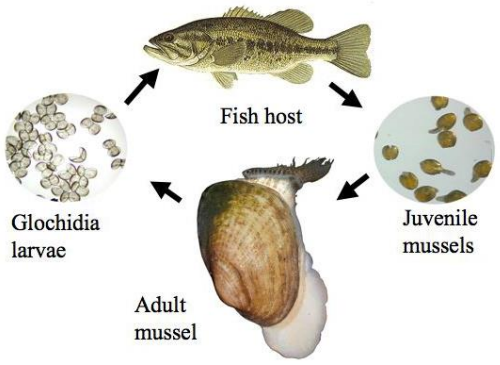
Aquatic Insects

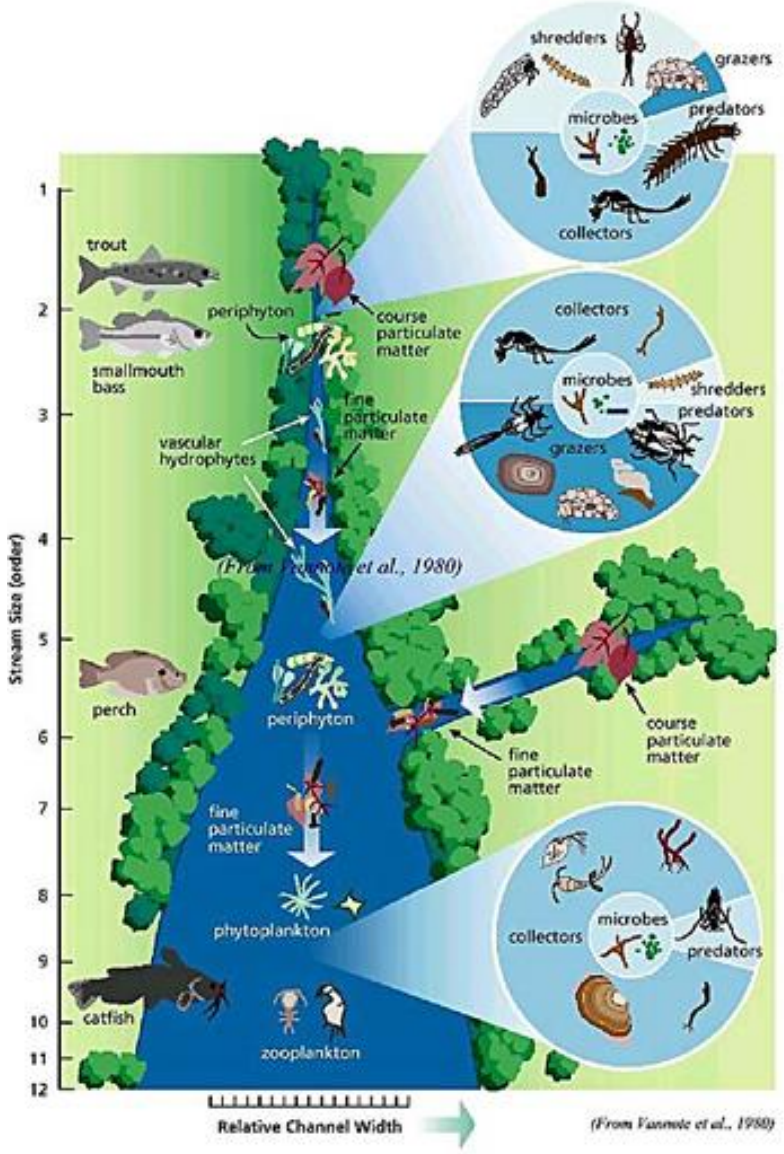
M.O.S.A.I.C.

- **Mobilizing the Organization of Southeastern Aquatic Insect Collections**: Freshwater Insects as Sentinels for Environmental Conditions.
- Aquatic insects are recognized throughout the world as a prognostic tool for the health of freshwater aquatic ecosystems
- Ephemeroptera, Plecoptera, and Trichoptera, (EPT, commonly known as mayflies, stoneflies, and caddisflies) exhibit a low threshold for pollution.
- Other aquatic insects groups (e.g., Diptera) exhibit high pollution thresholds.

Stream Macrofauna Associations

- Stream insects and other invertebrates play an important role in mobilizing energy from allochthonous fixed carbon sources.
- Form complex associations and interdependencies with stream fishes and other vertebrates.

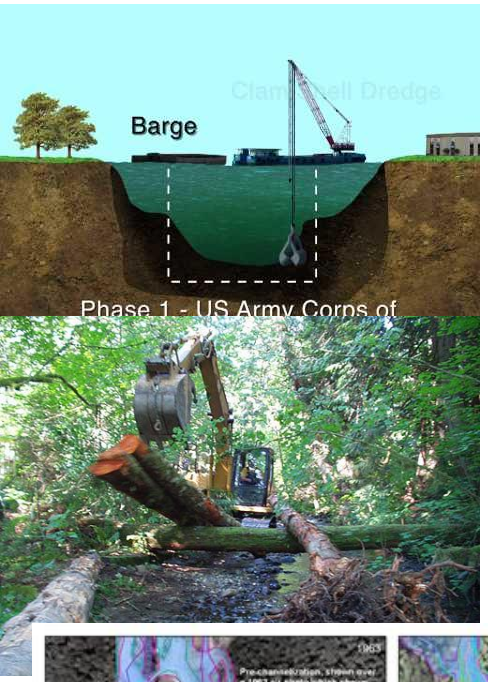


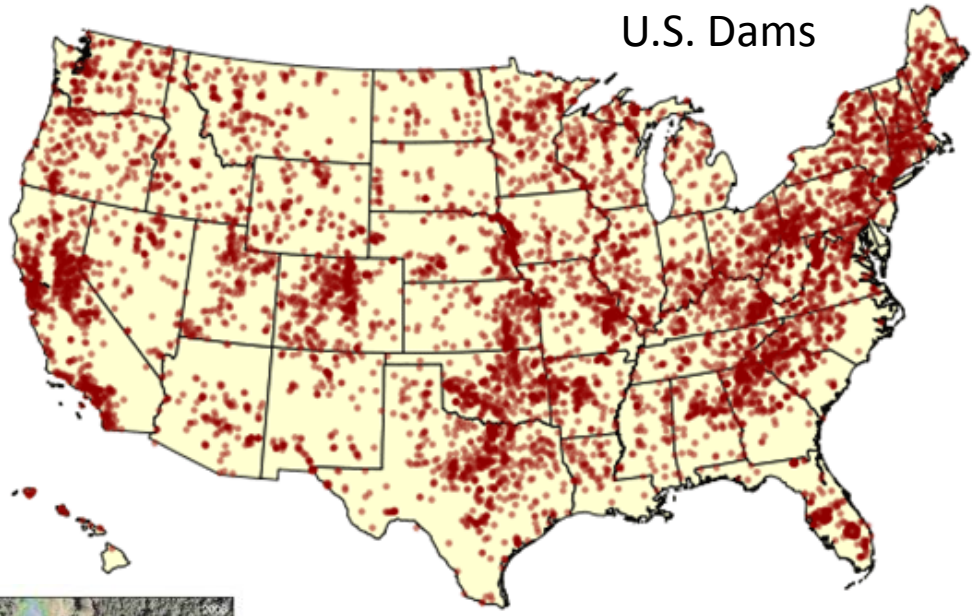


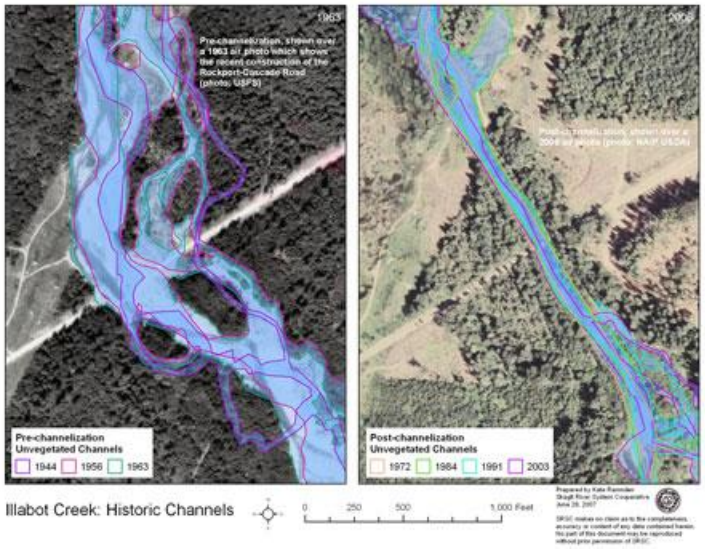
Southeastern U.S.

- Freshwaters of the southeastern U.S. are the richest biologically in North America
- Regrettably, the Southeast has led the nation in lost wetland acreage to activities like agriculture, flood control and urban sprawl.
- Streams of the southeast are also among the most modified in the U.S…

Stream Modification







Environmental Sentinels

- The use of living organisms as sentinels has advantages over traditional chemical and microbial water-quality analyses.
- Freshwater organisms, such as fish and macroinvertebrates, live almost continuously in the water and respond to all environmental stressors.
- Populations, species and communities of these organisms severely impacted by environmental disturbance.

Unified TCN Macrofauna Project

- **M**obilizing the **O**rganization of **S**outheastern **A**quatic **I**nvertebrate **C**ollections: Freshwater *Invertebrates* as Sentinels of Environmental Disturbance.
- 33 institutions from 14 states.
- Study area covers all or parts of 18 states (Texas, Oklahoma, Louisiana, Arkansas, Mississippi, Tennessee, Alabama, Georgia, North Carolina, South Carolina, Florida, Virginia, West Virginia, Kentucky, Kansas, Missouri, Illinois, Indiana).

Objectives

- Digitizing collections of freshwater invertebrates of the southeastern U.S.
- Integration of data across taxonomic groups to facilitate across-taxon studies of species associations, interdependencies, and correlated responses to instances of environmental disturbance and degradation.
- Take advantage of technological efficiencies of other biodiversity informatics projects.

Participating Collections

15 participating collections housing 582,000 mollusk and decapod crustacean specimens.

Institution	Types of invertebrate collections	Specimens/Lots	SE %	% Databased	% Georeferenced	Internet Accessible?	Active curator
Appalachian State University	Mollusks, decapods	700 lots	90% or more	100%	15%	No	Soon
Auburn University	Mollusks, decapods	10,000 specimens	90% or more	100%	?	Yes	**Active**
Arkansas State University	Mollusks, decapods	6,000 specimens	90% or more	70%	50%	No	**Adjunct**
Delaware Museum of Natural History	Mollusks	30,000 specimens	50%	90%	0%	No	Active
Eastern Kentucky University	Mollusks, decapods	5,000 specimens	90% or more	?	?	No	Active
Florida Museum of Natural History	Mollusks	40,000 lots	90% or more	100%	100%	Yes	Active
Georgia Museum of Natural History	Mollusks, decapods	51,000 specimens	90% or more	50%	0%	No	Active
Georgia State University	Decapods	1500 lots	90% or more	?	?	No	Active
Illinois Natural History Survey	Mollusks	144,000 specimens	33%	100%	90%	Some	Active
Mississippi Museum of Natural Science	Mollusks, decapods	~140,000 specimens	90% or more	60%	60%	No	**Retiring**
North Carolina Museum of Natural Sciences	Mollusks, decapods	50,000 specimens	70% or more	100%	50%	Yes	**Mixed**
Ohio State University	Mollusks	740,000 specimens	33%	100%	95%	Yes	Active
University of Alabama	Mollusks, decapods	9,000 specimens	90% or more	100%	85%	No	**Searching**
Univ. of Tennessee (Frank McClung Museum)	Mollusks	50,000 specimens	85%	?	?	No	**Adjunct**
Texas Natural Science Center	Mollusks, decapods	12,000 specimens	70%	70%	90%	No	Acting

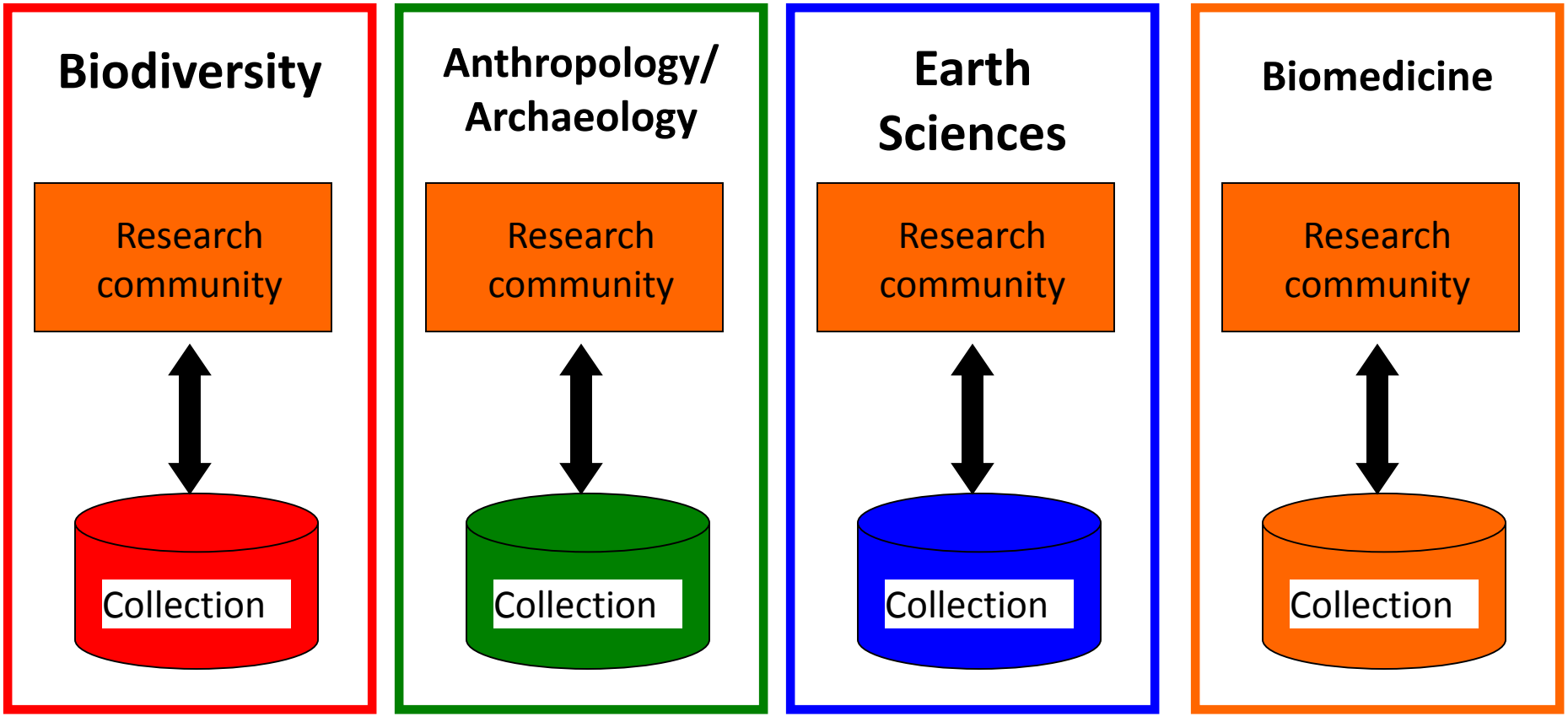
MOSAIC Institutions, Taxa Covered and Number of Specimens: (E,O,P,H,M,N,C,T,L,D = Ephemeroptera, Odonata, Plecoptera, aquatic Hemiptera, Megaloptera, aquatic Neuroptera, aquatic Coleoptera, Trichoptera, aquatic Lepidoptera, and aquatic Diptera, respectively)

Institution	Taxa Covered	Number of Specimens
Alabama Dept. of Environmental Management	E, O, P, H, M, N, C, T, L, D	937
Alabama Museum of Natural History (University of Alabama)	E, O, P, H, M, C, T, D	9,463
Clemson University	E, O, P, H, M, N, C, T, L, D	498,155
Florida A&M University	E, O, P, H, M, C, T, D	34,650
Florida Dept. of Environmental Protection (Includes SBIO Database)	E, O, P, H, M, N, C, T, L, D	4,000
Florida State Collection of Arthropods	E, O, P, C, T	215,880
Georgia Department of Natural Resources, Environmental Protection Division	E, O, P, H, M, N, C, T, L, D	5,962
Illinois Natural History Survey	E, O, P, H, M, C, T	157,710
International Odonata Research Institute	O	187,000
Kentucky Department of Environmental Protection	E, O, P, T, D	537
Mississippi Department of Environmental Quality	E, O, P, H, M, N, C, T, L, D	668
Mississippi State University (MEC)	E, O, P, H, M, N, C, T, D	42,094
North Carolina Dept. of Environmental and Natural Resources, Division of Water Quality	E, O, P, H, M, N, C, T, L, D	1,250,998
North Carolina State University	E, O, P, H, M, N, C, T, L, D	64,713
South Carolina Department of Health & Environmental Control (SCDHEC)	E, O, P, H, M, N, C, T, L, D	473,939
Tennessee Department of Environment & Conservation	E, O, P, H, M, N, C, T, L, D	5,832
University of Georgia	E, O, P, H, M, N, C, T, D	103,032
University of Kentucky	E, P, T, D	3,850
University of Tennessee (Includes Etnier Collection)	O, C, T, D	97,056
US National Museum of Natural History (Database**)	E, O, P, H, M, N, C, T, L, D	133,576
TOTAL		**3,099,220**

Issues

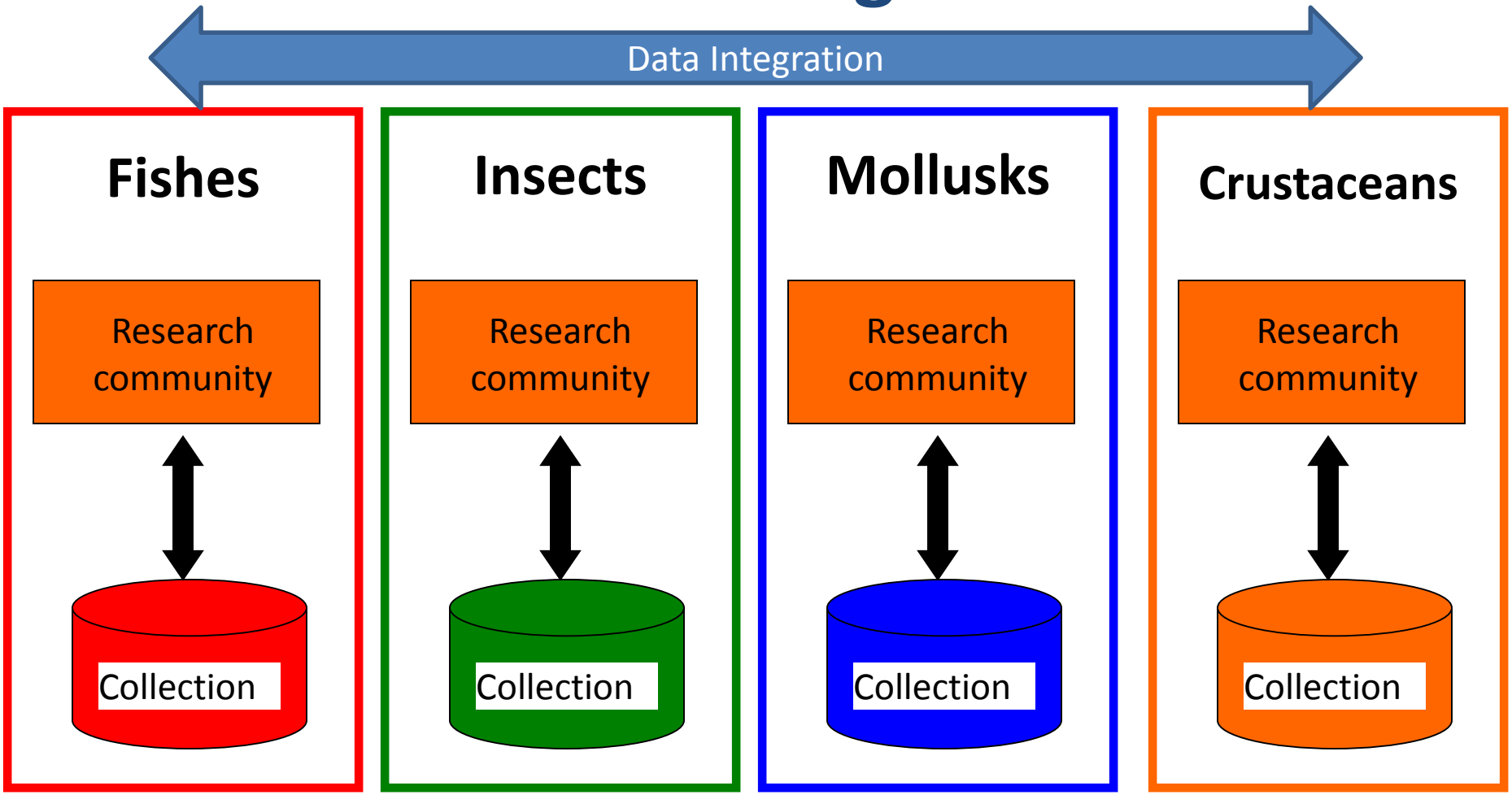
- Mixes of dry shells and fluid-preserved soft tissues; specimens and lots; jars and vials; sometimes ledgers, sometimes not.
- Maintaining healthy southeastern collections
  - Restoring curators to orphaned collections
  - Increasing institutional appreciation/support for collections.

‘Stove-piped’ Collections and Research Programs



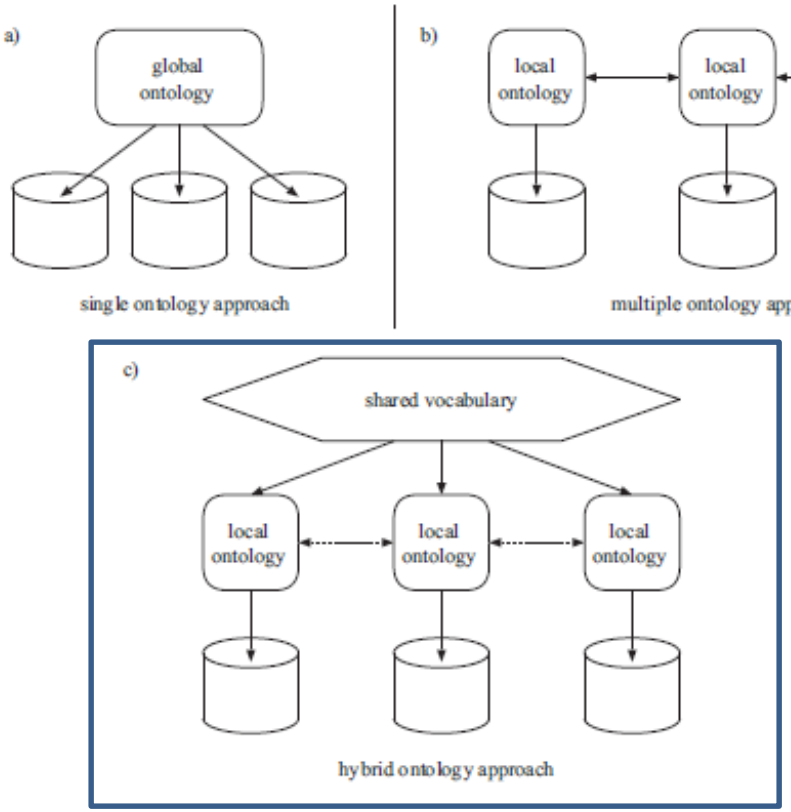
CONSORTIUM FOR THE BARCODE OF LIFE

‘Stove-piped’ Collections and Research Programs



CONSORTIUM FOR THE BARCODE OF LIFE

Ontology-Based Data Integration



Semantics of each source described by its own ontology, but source ontologies built upon one global shared vocabulary. Greatest flexibility for future expansion.

Figure 1: The three possible ways for using ontologies for content explication

Building Source Ontologies

- Files built in RDF and/or OWL and shared over the web.
- Ontologies would be built for describing and integrating taxonomy, natural history objects, ecological associations (e.g., functional feeding groups) and environmental associations.
- Would integrate the ontologies with databasing platforms (e.g., Specify, Symbiota).

Efficiencies

- Fish collections largely digitized, databased, networked (www.fishnet2.net).

- Project will take advantage of technology/experience gained from digitizing/networking/collaboratively georeferencing U.S. fish collections.

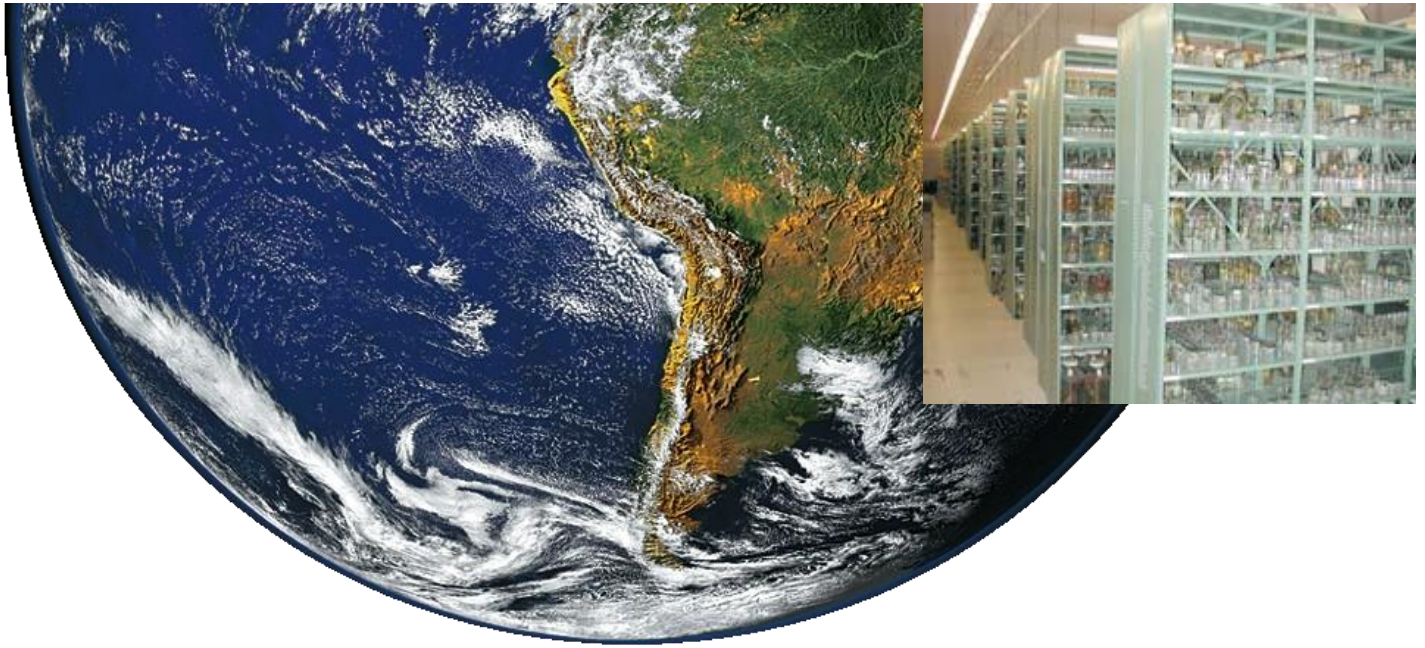
Georeferencing Natural History Collections Data: The GEOLocate Project

Nelson E. Rios

TUBRI

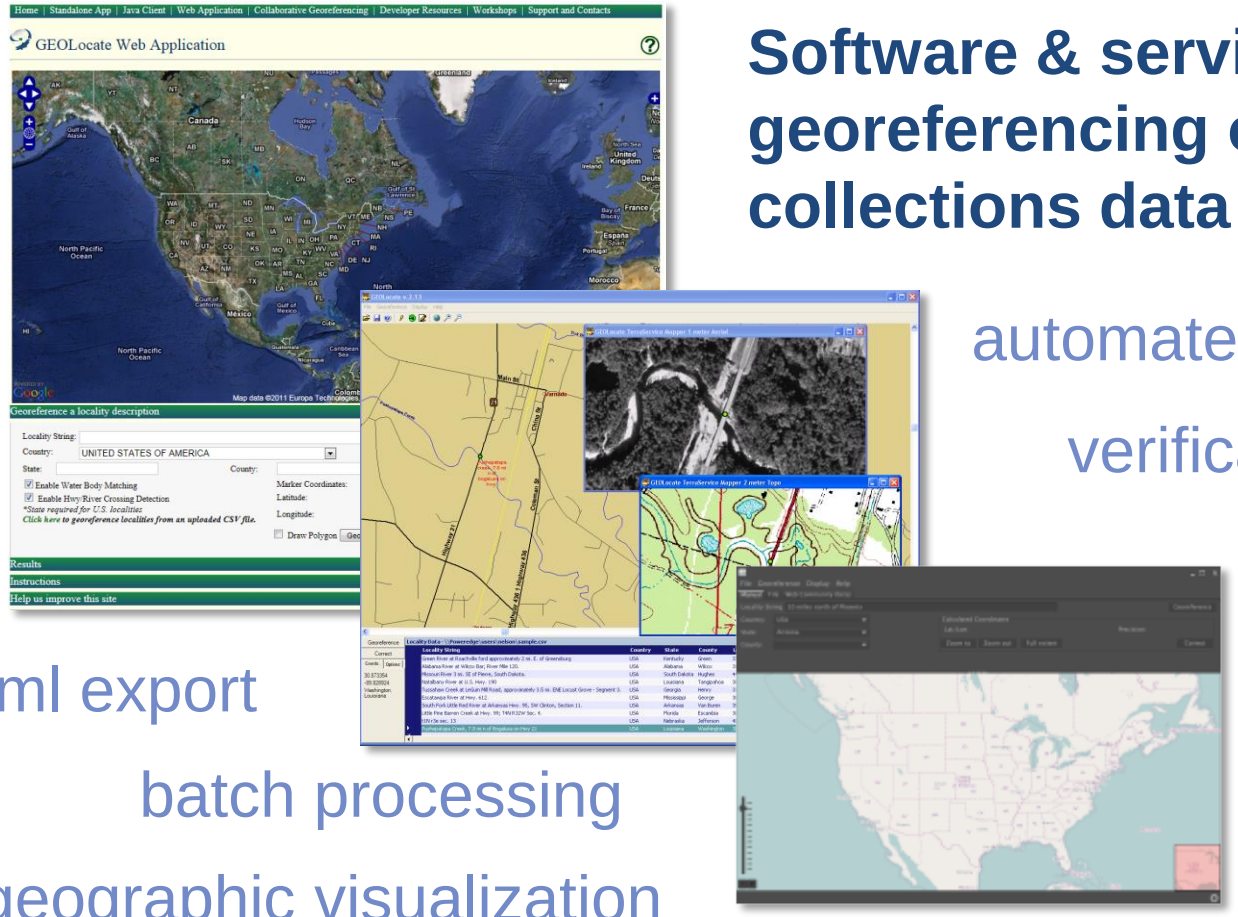
What is Georeferencing

3 billion specimens with ½ billion collecting events



Traditional Methods

GEOLocate



Software & services for georeferencing of natural history collections data

automated georeferencing

verification & correction

multi-lingual

interoperability
soap & rest api

training

uncertainty determination

kml export

batch processing

geographic visualization
google, bing, openstreet, wms

collaborative georeferencing

Typical GEOLocate Workflow

1. **Data Entry & Preparation**
2. **Automated Processing**
3. **Manual Verification**

pushepatapa creek, trib. to pearl river, 7.8 miles north of bogalusa at hwy 21; Washington; LA; USA

↓

Georeferencing Algorithm

↓

Visualize, verify & adjust output coordinates & uncertainties

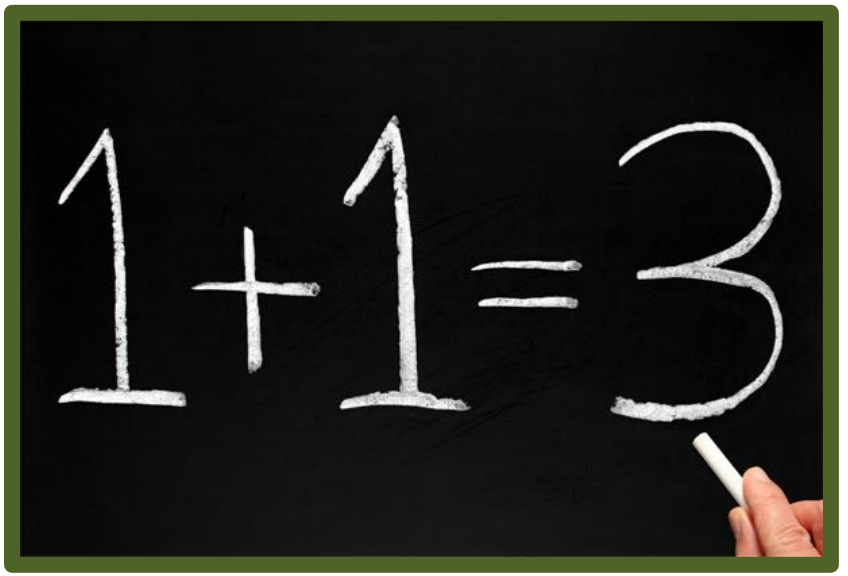
latitude: **30.88797**
longitude: **-89.83601**

uncertainty radius: **48m**

uncertainty polygon:
30.88823,-89.83641,
30.88815,-89.83634,
30.88808,-89.83622…

Collaborative Georeferencing (CoGe)

- Increased output by taking advantages similarities across collections.
- Distribution of workloads to appropriate expertise.



FishNet 2

Collaborative Georeferencing: Sharing Data

Homochitto River, 4.9 mi. E Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1960	11
Homochitto River 5 mi. E of Union Church, MS Hwy. 550.	USA	Mississippi	Lincoln	1968	11
Homochitto River at Route 550 bridge, 5.0 mi. E of Union Church.	USA	Mississippi	Lincoln	1970	10
Homochitto River at MS Hwy. 550 bridge 5.0 mi. E of Union Church.	USA	Mississippi	Lincoln	1971	2
Homochitto River at Hwy. 550 bridge 5.0 mi. E of Union Church.	USA	Mississippi	Lincoln	1971	
Homochitto River 5 mi. E of Union Church.	USA	Mississippi	Lincoln	1971	
Homochitto River 5 mi. E of Union Church.	USA	Mississippi	Lincoln	1971	
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	
Homochitto River 4.7 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1970	
Homochitto River 4.7 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	
Homochitto River 4.7 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1973	
Homochitto River 5 mi. SE of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1978	
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1971	8
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1971	6
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	8
Homochitto River 5 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	7
Homochitto River 4.7 mi. E of Union Church, Hwy. 550.	USA	Mississippi	Lincoln	1972	6

Corrected 22 collecting events, approx. 200 specimen records

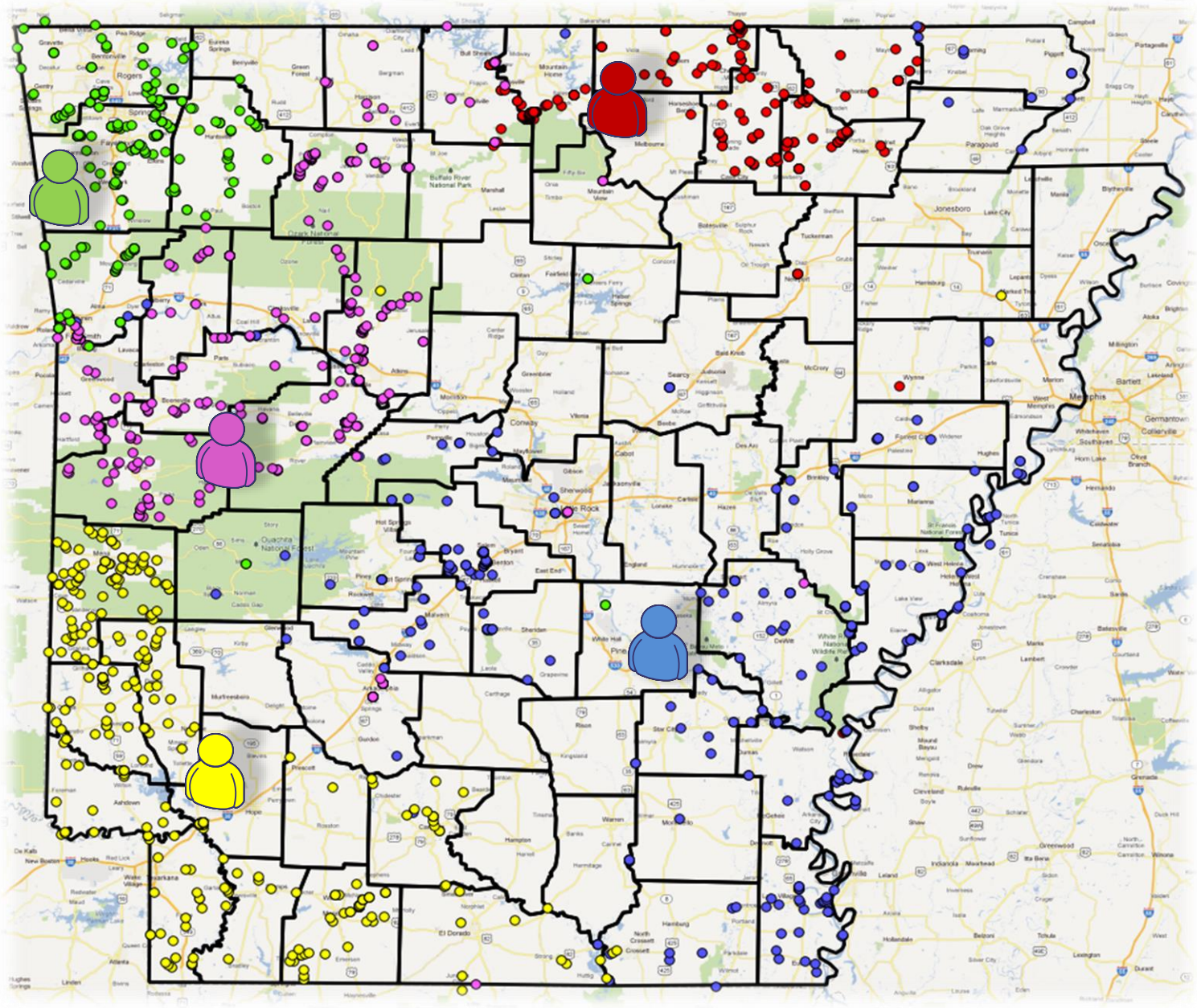
Collaborative Georeferencing Performance

- 2100 randomly selected collecting events from the Tulane University fish collection were imported and georeferenced using the collaborative georeferencing framework
- 33% were duplicates
- 30% more related by similarity index

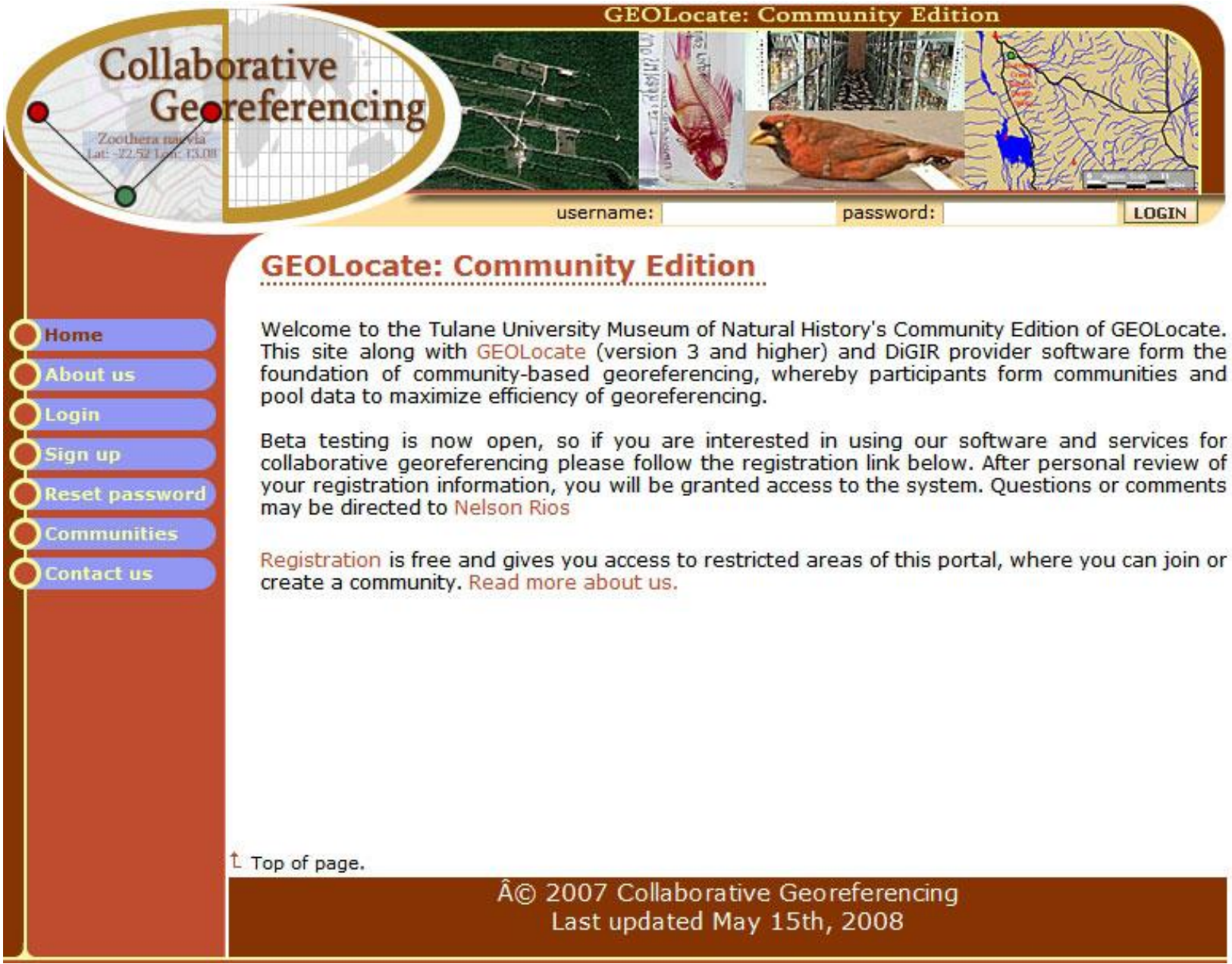
2100 → 782

★ 63% reduction in effort overall!!!

Distributing Workloads


Rogers
Jonesboro
Memphis
Bartlett
Germantown
Collierville
Arkadelphia
Hot Springs
Little Rock
North Little Rock
Conway
Searcy
Batesville
Paragould
Blytheville
Clarksdale
Greenwood
Grenada
Tupelo
Texarkana
El Dorado
Pine Bluff
Stuttgart
Brinkley
Helena
West Helena
Marianna
Forrest City
Wynne
McCrory
Cabot
Jacksonville
Sherwood
Benton
Malvern
Hope
Nashville
Prescott
Magnolia
Camden
Warren
Hampton
Monticello
McGehee
Dumas
Rison
Star City
Pine Bluff
White Hall
DeWitt
Lake View
Marks
Lambert
Crowder
Courtland
Batesville
Clarksdale
Shelby
Mound Bayou
Cleveland
Ruleville
Drew
Indianola
Inverness
Leland
Greenville
Belzoni
Tchula
Lexington
Durant
Holly Grove
Des Arc
Hazen
Carlisle
Lonoke
Austin
Gibson
Bald Knob
Judsonia
Higginson
Griffithville
McRae
Beebe
Searcy
Kensett
Harrisburg
Lepanto
Tyronza
Marion
West Memphis
Whitehaven
Southaven
Horn Lake
Olive Branch
Millington
Arlington
Brighton
Atoka
Covington
Brownsville
Somerville
Oakland
Collierville
Germantown
Bartlett
Memphis
Hughes
Forrest City
Wheatley
Parkin
Earle
Crawfordsville
Turrell
Jennette
Lepanto
Tyronza
Harrisburg
Weiner
Trumann
Caraway
Jonesboro
Lake City
Bay
Monette
Manila
Leachville
Osceola
Wilson
Burdette
Dyess
Keiser
Joiner
Black Oak
Lepanto
Marked Tree
Tyronza
Earle
Parkin
Wheatley
Hughes
Forrest City
Palestine
Wynne
Cherry Valley
Widener
Colt
Hulbert
Lexa
Haynes
Elaine
Helena
West Helena
Lake View
Barton
Almyra

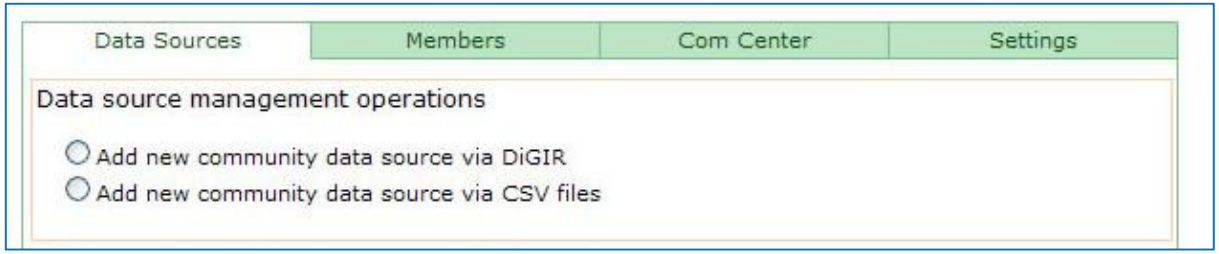
Collaborative Georeferencing Management Portal

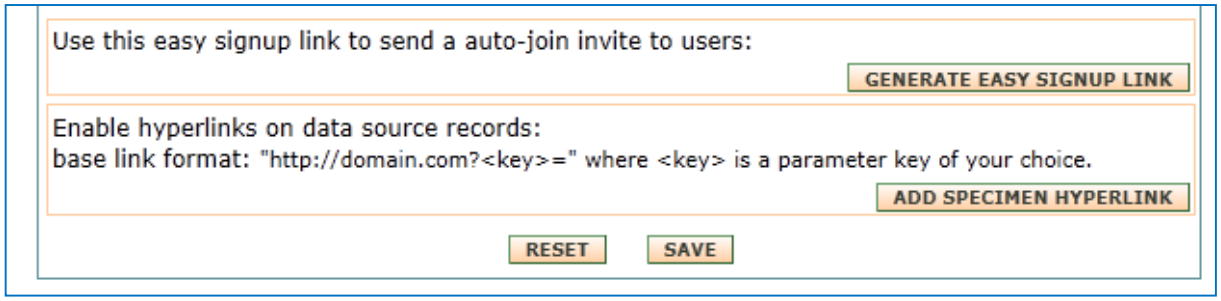


Georeferencing Communities

Create Communities

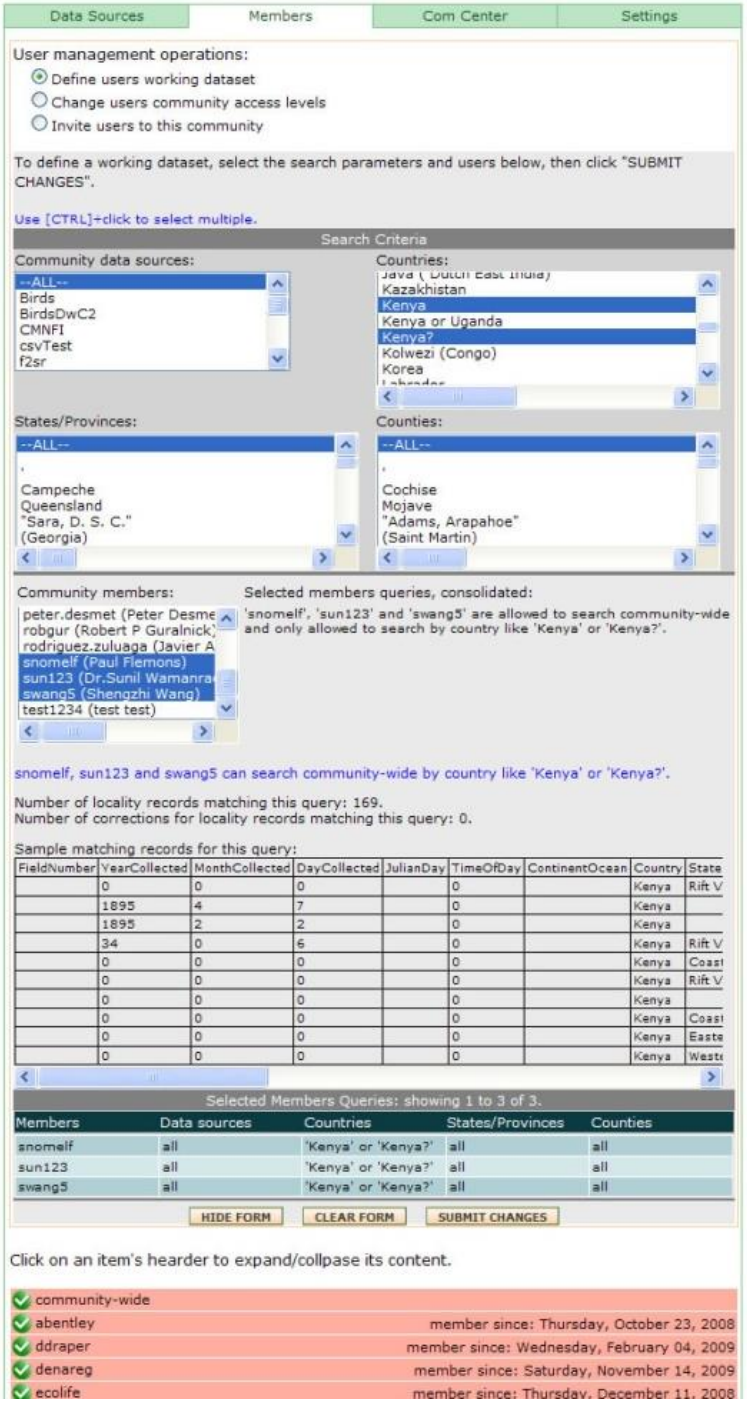
Data Sources





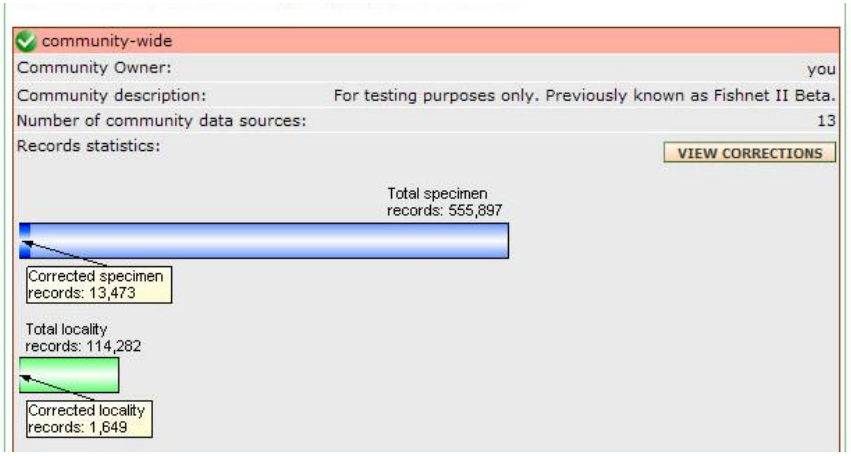
Add New Users & Link Out

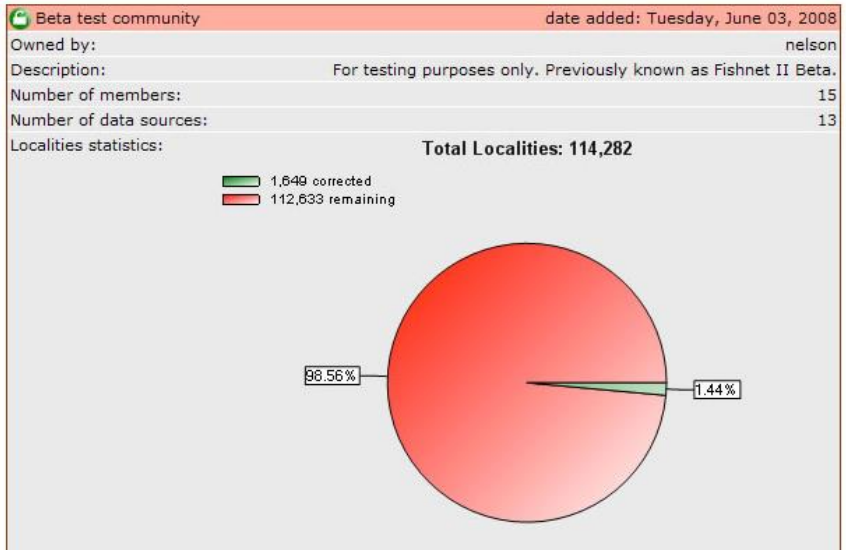
User Management: Task Assignment

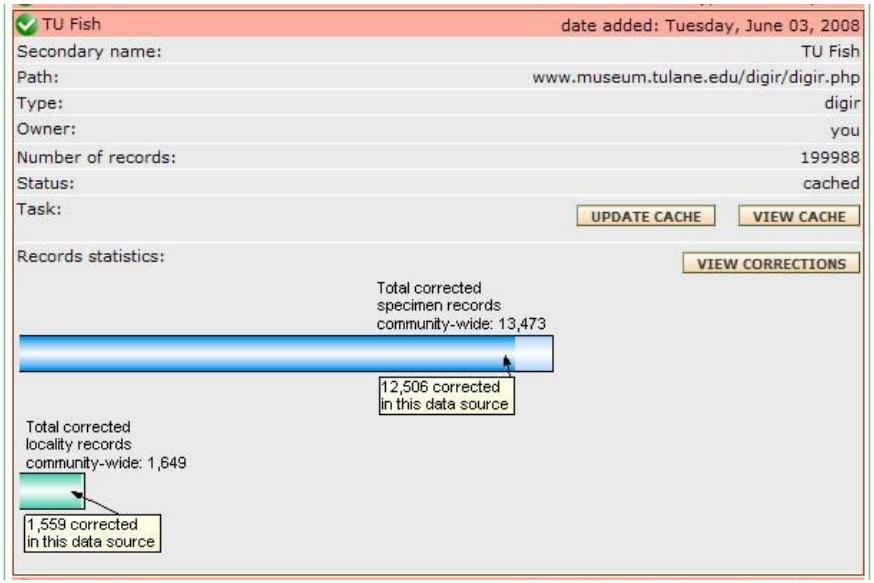


Assign all records from Kenya to experts on East African regions

Monitoring & Managing Progress







Data Repatriation

TU Fish	date added: Tuesday, June 03, 2008
Secondary name:	TU Fish
Path:	www.museum.tulane.edu/digir/digir.php
Type:	digir
Owner:	you
Number of records:	199988
Status:	cached
Task:	UPDATE CACHE VIEW CACHE
Records statistics:	VIEW CORRECTIONS

CoGe: Corrections details for the "TU Fish" data source - Mozilla Firefox

http://www.museum.tulane.edu/coge/protected/ShowDetails.aspx?.wid=3909f73d&total=1559&show=135&obj=ds

Showing 1 to 50 of 1559. DOWNLOAD CORRECTIONS

	Data source	ContinentOcean	Country	StateProvince	County	Locality
-	TU Fish	North America	USA	Utah	Washington	Santa Clara River, 2.75 mi. above Gunlock.

Correction order (most recent first)	Corrected by	Corrected longitude	Corrected latitude	Correction remarks	Date corrected
1	abentley	-113.748037	37.299871		10/27/2008 9:48:20 PM

+	TU Fish	North America	USA	Mississippi	Copiah	Trib. to Homochitto River, 4.6 mi. NE Caseyville.
+	TU Fish	North America	USA	Mississippi	Lawrence	Pearl River at mouth of Fair River.
-	TU Fish	North America	USA	Florida	Lake	Lake Harris near Leesburg.

Correction order (most recent first)	Corrected by	Corrected longitude	Corrected latitude	Correction remarks	Date corrected
1	abentley	-81.85438	28.783233		10/27/2008 9:49:33 PM
2	abentley	-81.87806	28.81056		10/27/2008 9:49:00 PM

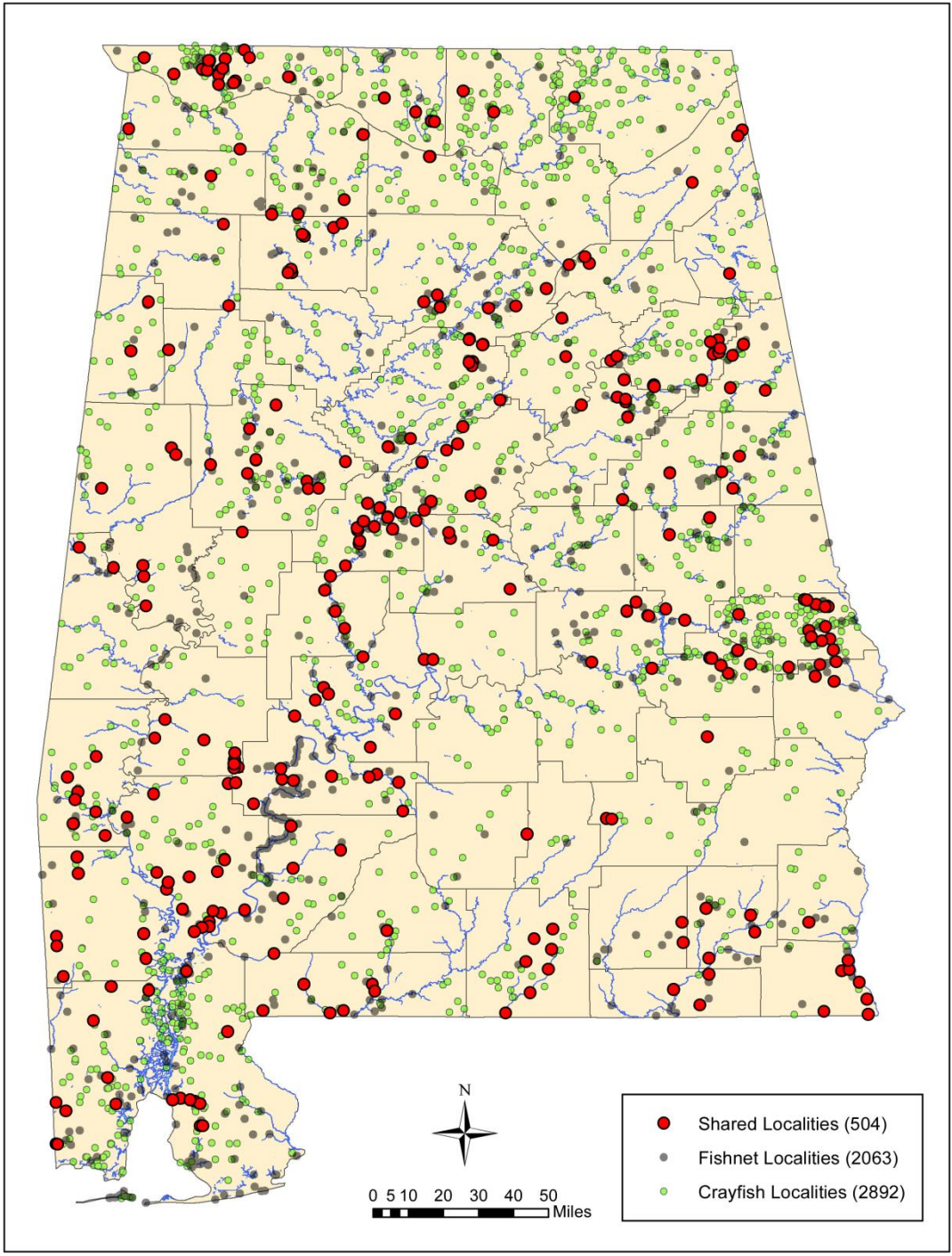
+	TU Fish	North America	USA	Mississippi	Copiah	Bayou Pierre, 1.7 mi. W of Carpenter or 5.8 mi. E of Carlisle.
+	TU Fish	North America	USA	Louisiana	Orleans	South shore of Lake Pontchartrain near I-10 bridge.
+	TU Fish	North America	USA	Mississippi	Simpson	Strong River at rapids upstream from Hwy. 28 bridge, 2 mi. W of Pinola.
+	TU Fish	North America	USA	North Carolina	Macon	Little Tennessee River, 5 mi. NW Franklin, Hwy. 28.
+	TU Fish	North America	USA	Louisiana	Orleans	Lake Pontchartrain at south end of causeway on Rt. 11
+	TU Fish	North America	USA	Louisiana	La Fourche	Ditch near canal, 5.7 mi. south of Golden Meadows on Rt. 78
+	TU Fish	North America	USA	Mississippi	Marion	Upper Little River tributary of Pearl River, 4.8 mi. SE of Columbia, MS on I
+	TU Fish	North America	USA	Mississippi	Marion	Trib. of Pearl R. 4.8 N of Sandy Hook on Rt. 35
+	TU Fish	North America	USA	Alabama	Mobile	Trib. of Escatawpa River 6.8 mi. E of Hurley, MS
+	TU Fish	North America	USA	Mississippi	Marion	Trib. of Pearl River 6/10 mi. N of LA state line on Rt. 35
+	TU Fish	North America	USA	Florida	Bay	Trib. of Econfina River at Youngstown on Rt. 231
+	TU Fish	North America	USA	Mississippi	Marion	Trib. of Pearl River 1.2 mi. N of MS state line on Rt. 35
+	TU Fish	North America	USA	Louisiana	St. Tammany	Trib. of Pearl River 12.7 mi. NW of town of Pearl River on Rt. 58.
+	TU Fish	North America	USA	Mississippi	Marion	Trib. of Pearl River 2.7 mi. N of Sandy Hook on Rt. 35.
+	TU Fish	North America	USA	Mississippi	Hancock	Long Bay, trib. to Turtleskin Cr. 11.2 mi. E of Picayune or 2 mi. W of turn
+	TU Fish	North America	USA	Florida	Jackson	Chipola River 1.5 mi. N of Marianna

Leveraging CoGe Process and System for Georeferencing Efficiency

- Beyond efficiencies of the locality matching system, Fishnet2 localities are a useful resource for georeferencing invertebrate collection localities.
- 149,642 localities currently in Fishnet2 in 18 states considered here as part of SE region.
- Many of these are replicates, and many are same places invertebrates have been collected.

Fish-Crayfish Locality Comparison

- Compared dataset of 2892 georeferenced crayfish localities from Crayfishes of Alabama project to 2063 georeferenced Alabama fish localities in Fishnet2
- Buffered lat/long by 90 meters to account for error in coordinate determinations (~ 3 seconds).
- 504 localities (18%) overlapped (were the same).



Proposal Development Plan

- Identify PIs and develop workflows for digitizing various preparations of specimens.
- Decide on digitization approach and perform task analyses/time trials.
- Develop project implementation plan.
- Develop budgets.
- Hopefully submit in October 2013.

Thanks for your attention!