Fundamentals of Georeferencing Using GEOLocate

Nelson E. Rios

Tulane University

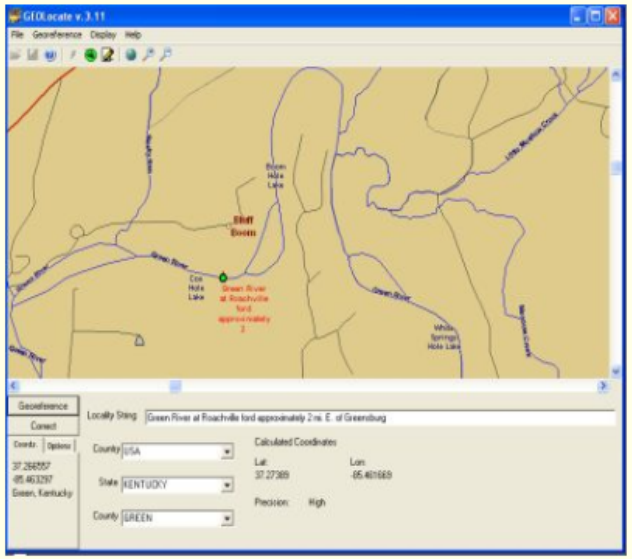
Georeferencing

The Problem

3 billion specimens with ½ billion collecting events

GEOLocate

A Platform for Georeferencing Natural History Collections Data



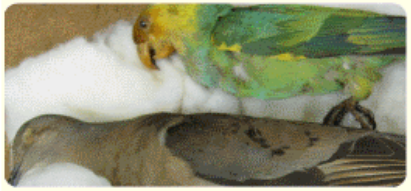
For Users:

- Overview
- GEOLocate Web Application
- Collaborative Georeferencing
- GEOLocate 3.xx (standalone)
  - Global Expansion
- Education & Outreach

Brief overview (video) of the GEOLocate Project.

For Developers:

- SOAP Services
- JSON/GeoJSON
- Embeddable Web Client



Web Application

Georeference collections data using your web browser. Quick and easy georeferencing.

Web Services

Integrate georeferencing into your own databases and applications using GEOLocate webservices.

Desktop Application

The original standalone desktop application.

Collaborative Georeferencing

Build communities, share data, relate records across collections and improve verification efficiency.

Tulane University

NSF

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

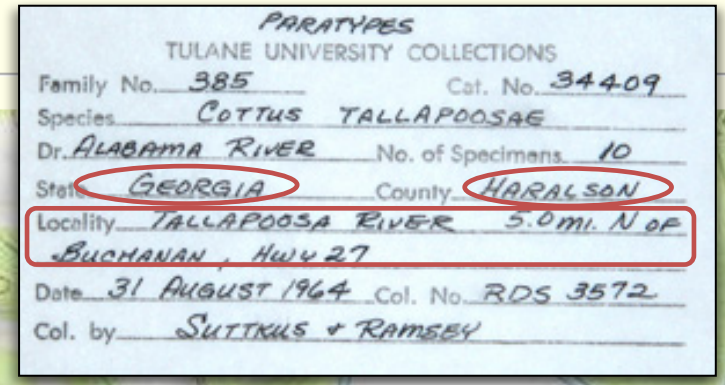
latitude: **30.88797**
longitude: **-89.83601**

uncertainty radius: **48m**

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

GEOLocate Web Application

1 possible location found.


PARATYPES
TULANE UNIVERSITY COLLECTIONS
Family No. 385 Cat. No. 34409
Species COTTUS TALLAPOOSAE
Dr. ALABAMA RIVER No. of Specimens 10
State GEORGIA County HARALSON
Locality TALLAPOOSA RIVER 5.0 mi. N of BUCHANAN, HWY 27
Date 31 AUGUST 1964 Col. No. RDS 3572
Col. by SUTTKUS & RAMSEY

Workbench | 1 possible location found

Georeference | Options | Draw polygon | Place marker | Measure

Locality String: Tallapoosa River 5.0 mi. N of Buchanan, Hwy 27

Country: UNITED STATES OF AMERICA

State: Georgia

County: Haralson

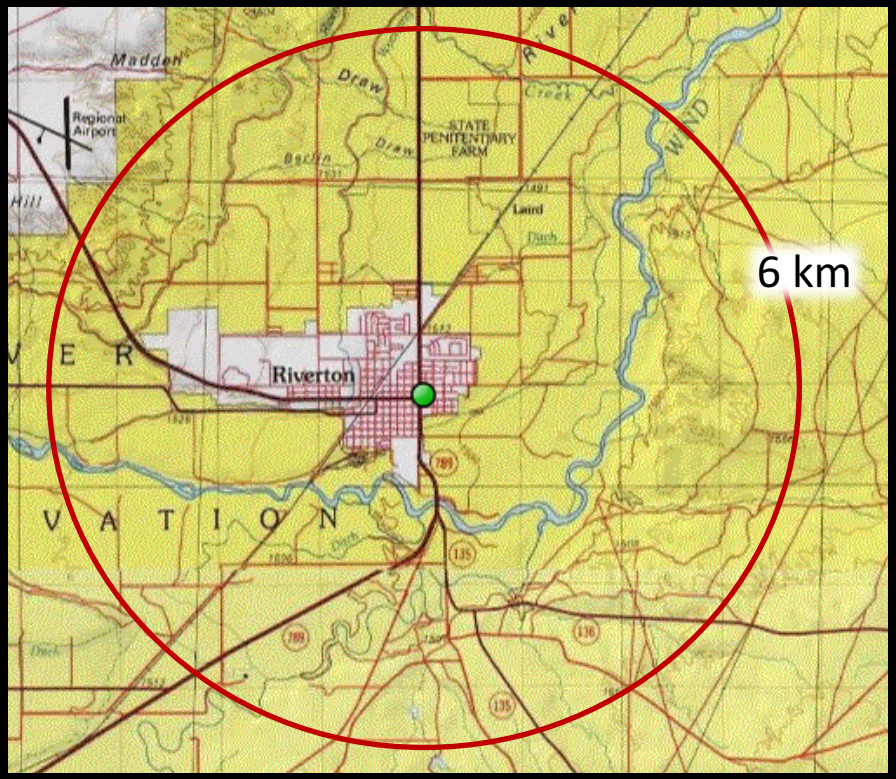
latitude: 33.863438 | longitude: -85.213334 | uncertainty: 90 m | error polygon

33.863438	-85.213334	90	Unavailable

Algorithm Performance

Versus Known U. S. Localities

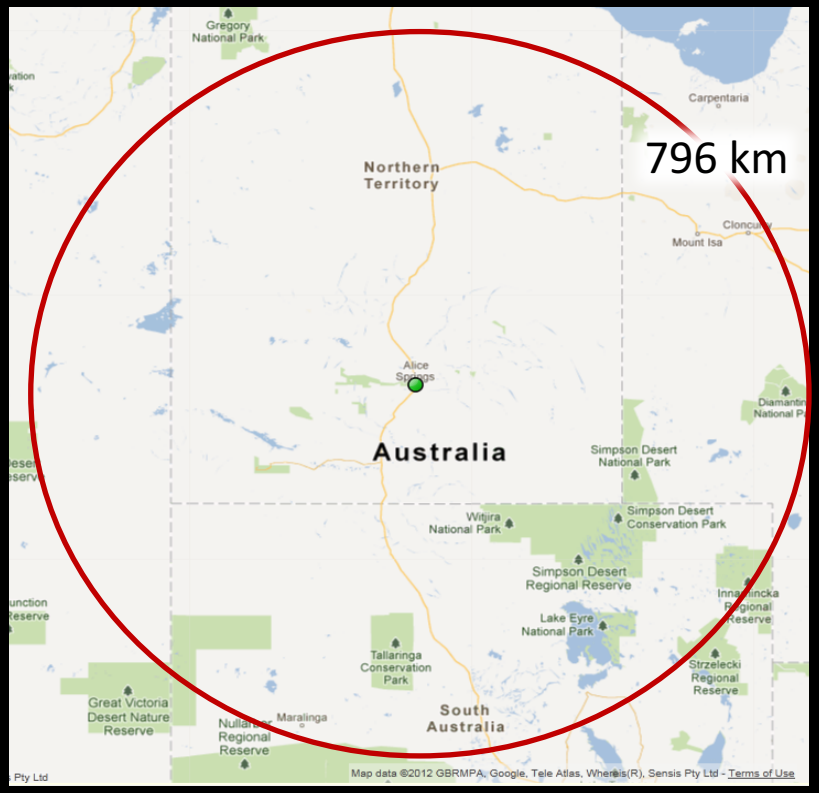
	% Found	Mean Dist. Off	Standard Error
GEOLocate	95%	6.1 km	2.1 km

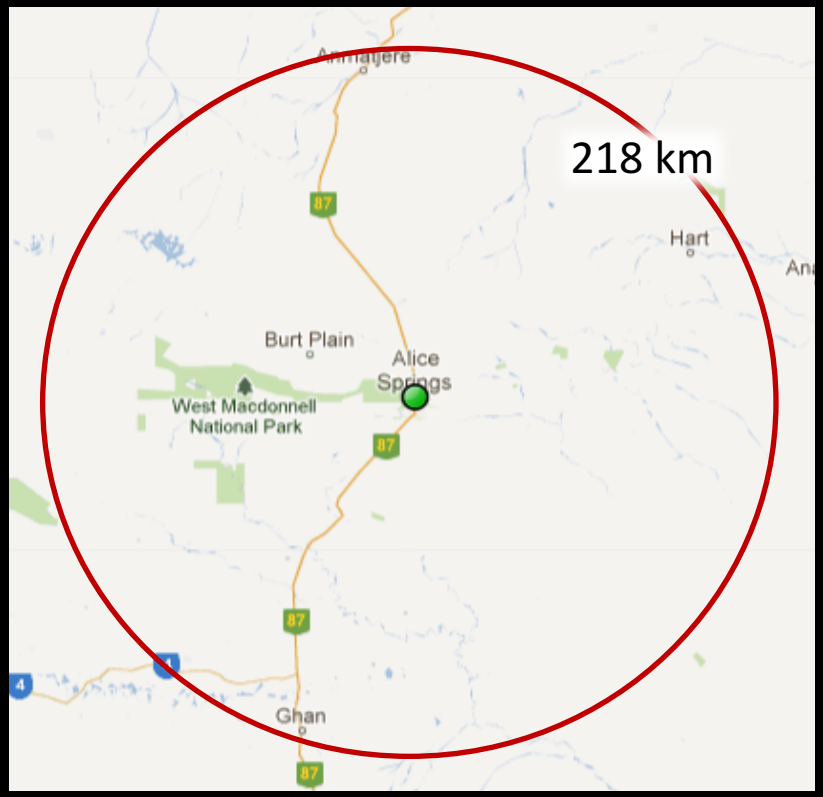


Algorithm Performance

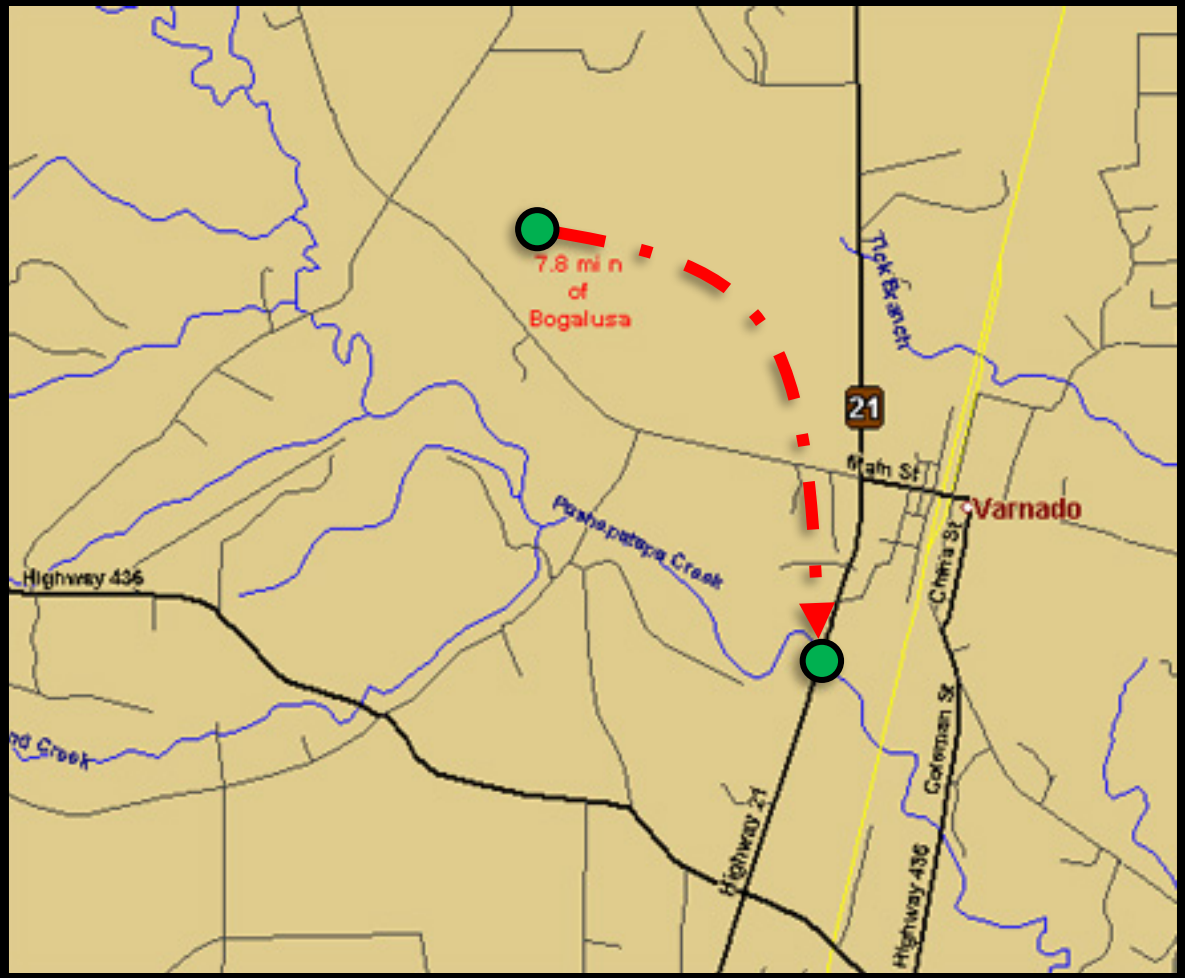
Versus Known Australian Localities

	% Found	Mean Dist. Off	Standard Error
GEOLocate	86 → 97%	**796 → 218 km**	154 → 48 km





Verification & Adjustment of Automated Outputs



Computed coordinates are displayed on digital maps

Manual verification of each record

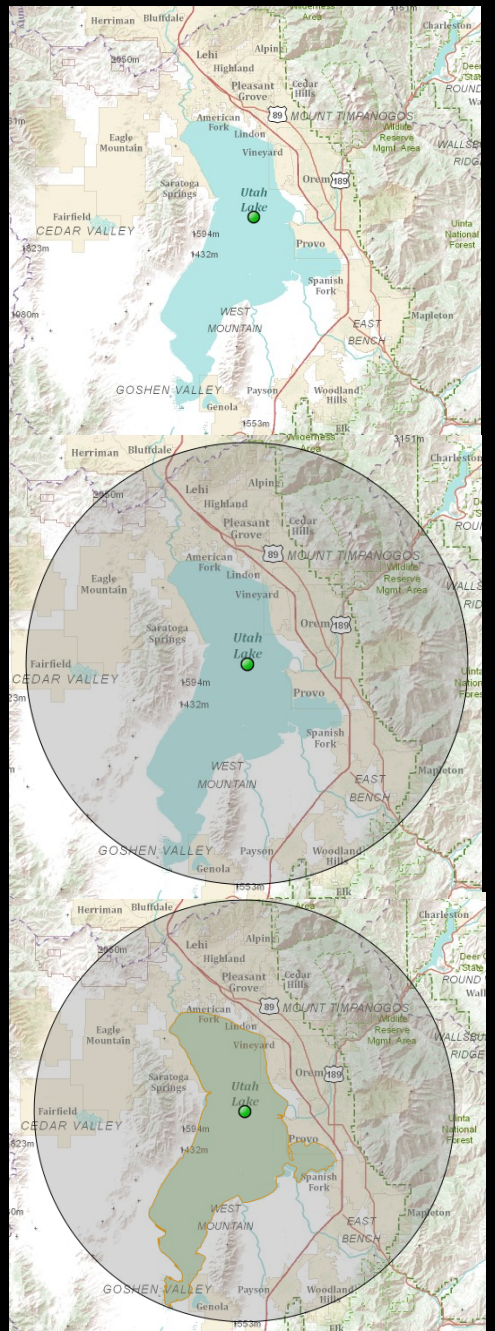
Drag and drop correction of records

Moving Beyond Points

- Point

Point & Radius

Point, Radius & Polygon

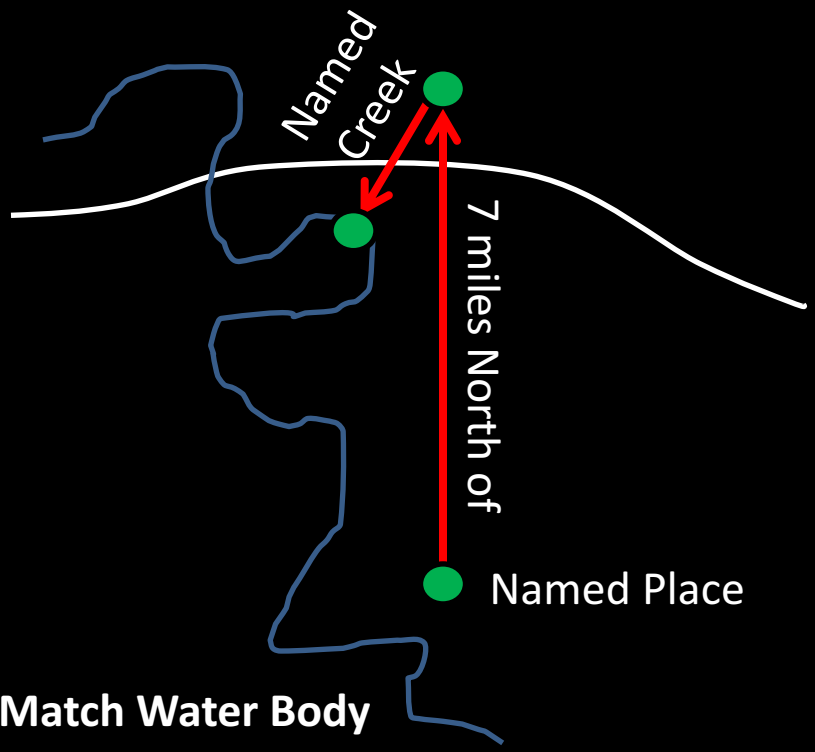


Georeferencing options

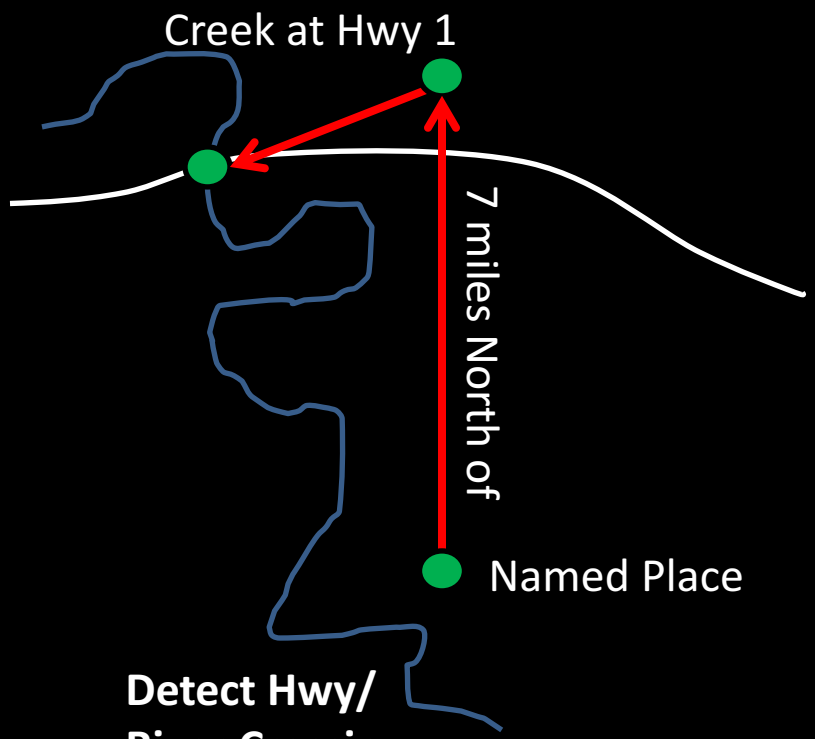
☑ Match Water Body | ☑ Detect Hwy/River Crossing | ☑ Do Uncertainty

☑ Do Error Polygon | ☐ Displace Polygon | ☐ Restrict to Lowest Adm. Unit

Language: English

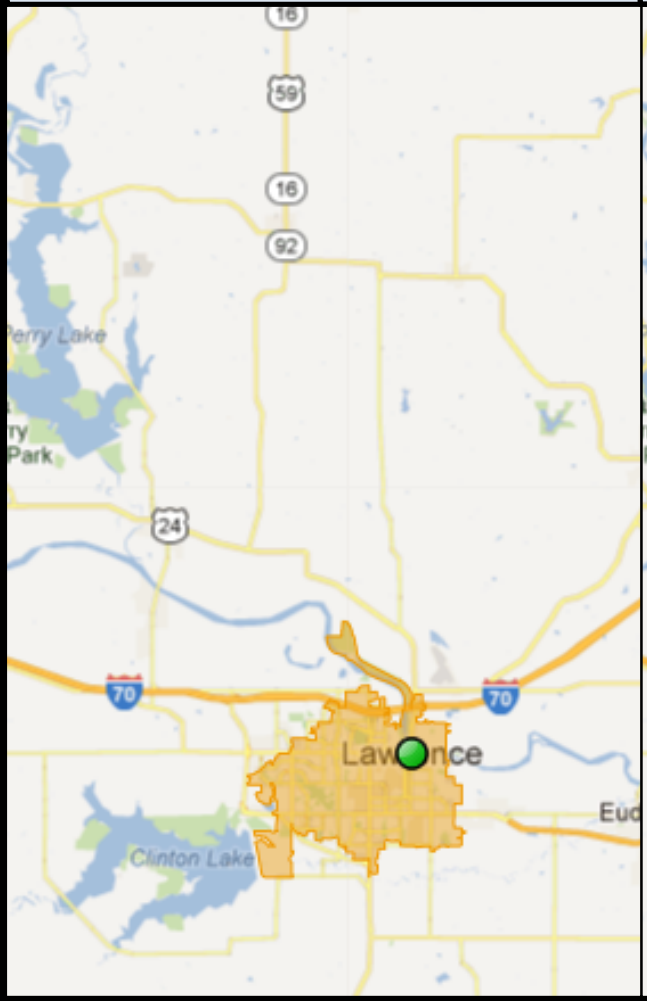
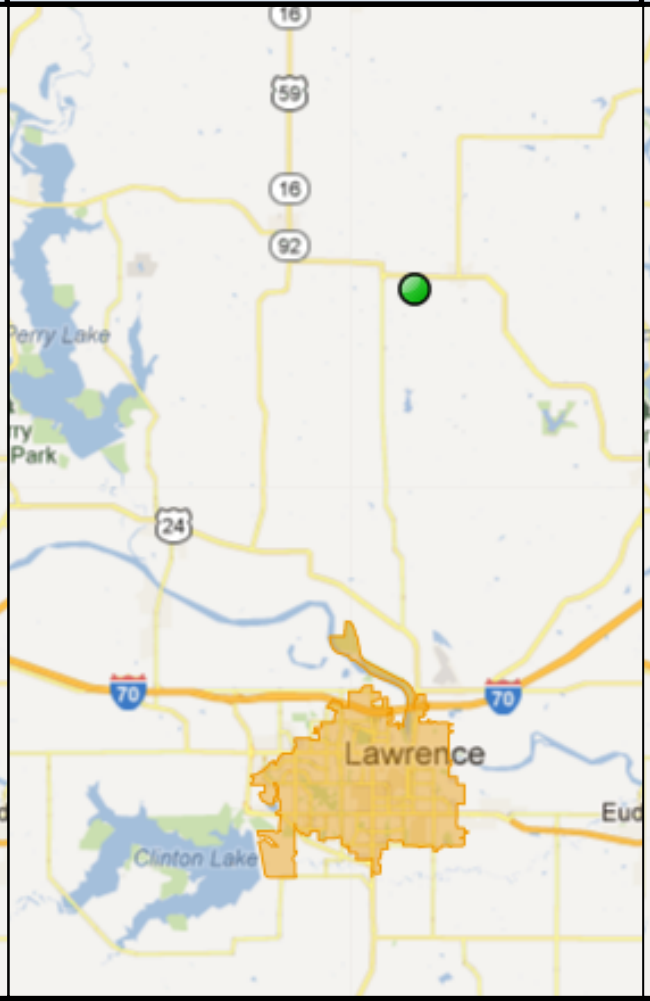
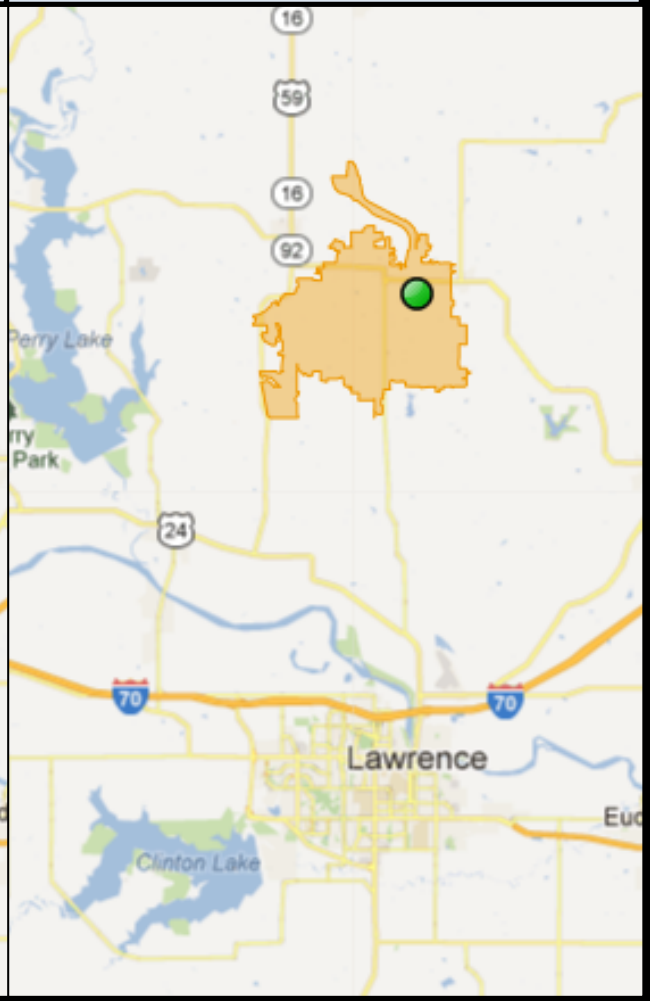


Match Water Body

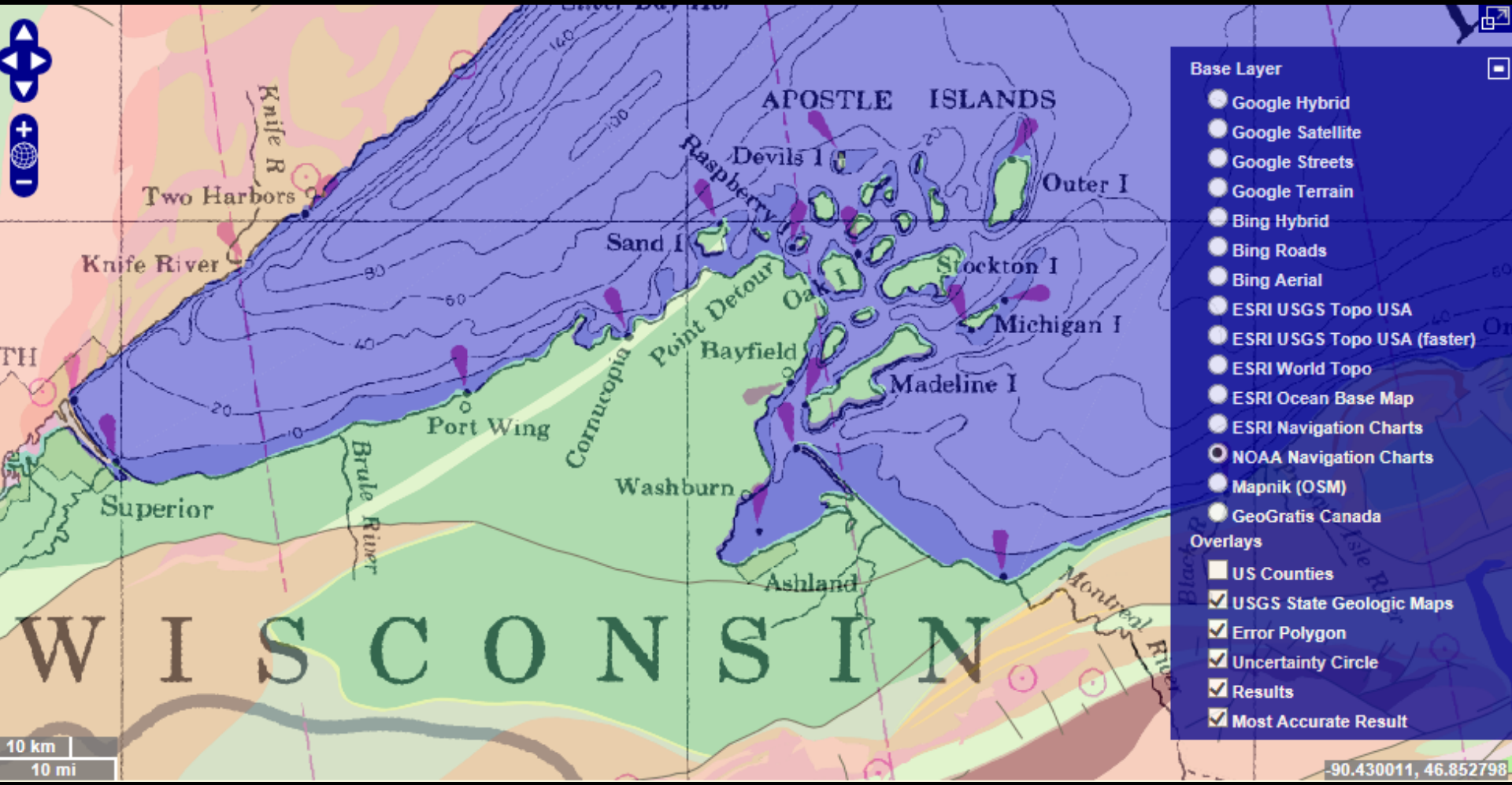


Detect Hwy/ River Crossing

Generating Polygons:

Lawrence, KS	15 mi N of Lawrence, KS no displacement	15 mi N of Lawrence, KS displaced polygon
		

Visualization: Base Layers



Primary Clients

The following based web clients are available to allow you to georeference data directly from your web browser:

Standard Client

Simply type in your locality description and get back georeferenced results. Start here if you are new to GEOLocate.

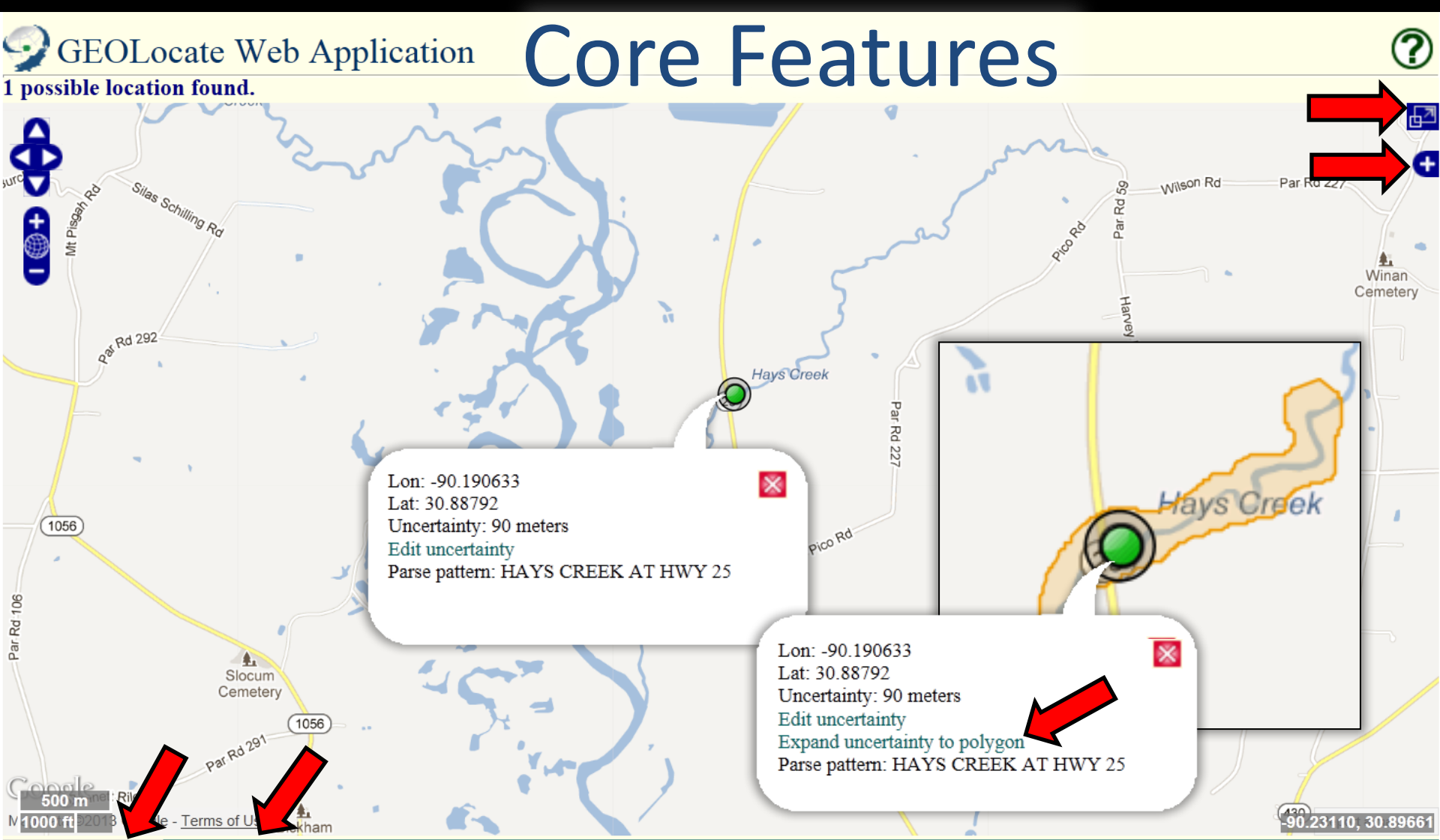
Batch (File Based) Client

Allows you to upload a .csv file and batch process it. (file formatting instructions)

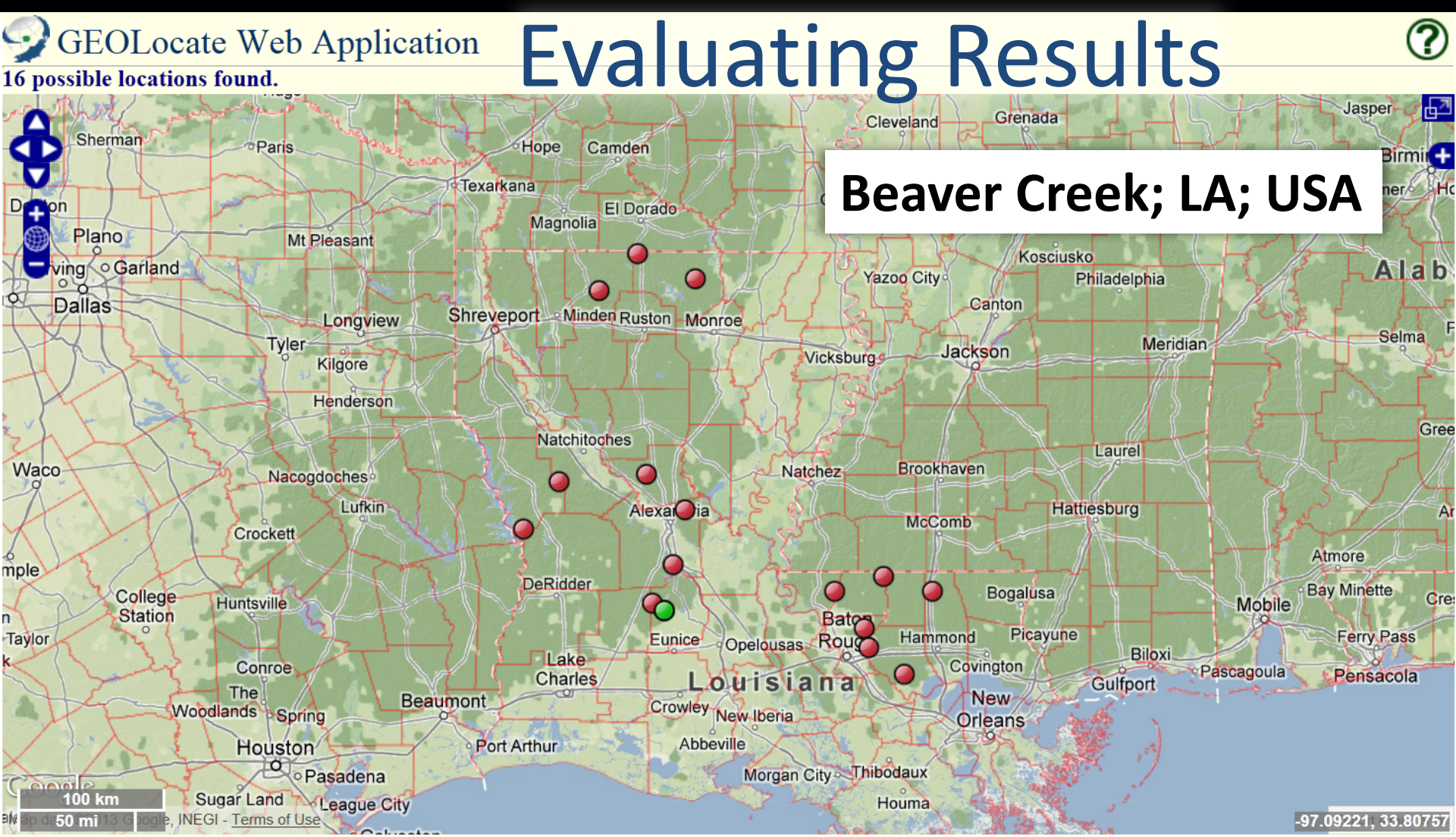
Collaborative Georeferencing Client

Utilizes the collaborative georeferencing framework. Ideal for largescale multi-institution projects. (https link)

Core Features



Evaluating Results



Batch Georeferencing

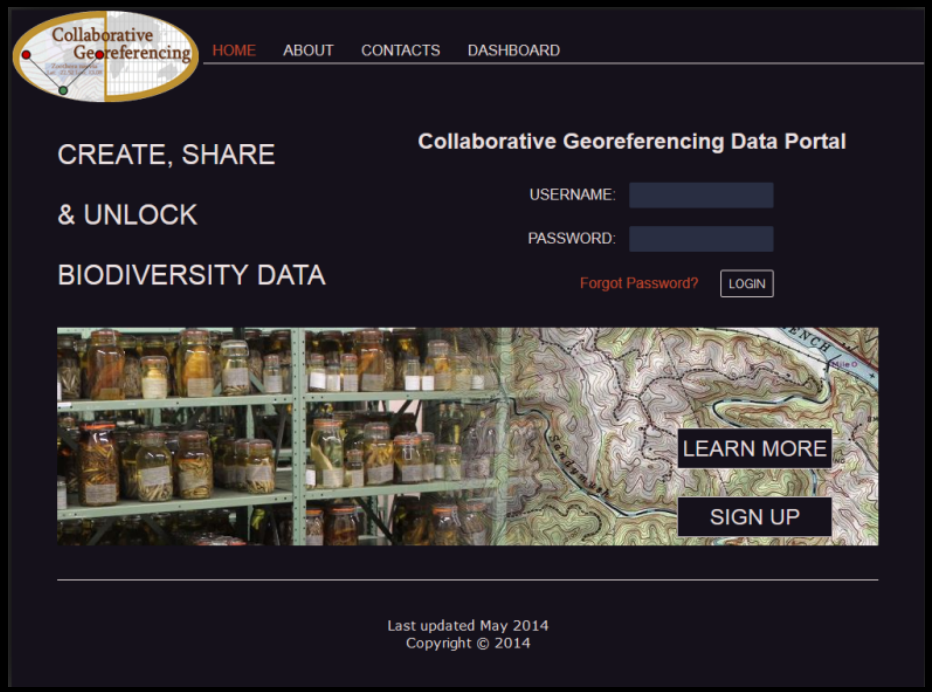
Workbench | 2 possible locations found

Show 8 records | Page Georeference | Georeference | Options | Correct | Draw polygon | Place marker | Measure

Locality	Country	StateProvince	County	Latitude	Longitude	Corrected	precision
Chambers Spring Road 2.5 km S of Hwy 412, 8.0 km E of Siloam Springs, T17N, R33W,	USA	Arkansas	Benton	36.188027	-94.451005	no	High(89)
Osage Creek, 1.0 mile N on gravel road to bridge crossing; gravel road jcts with	USA	Arkansas	Benton	36.189077	-94.395375	no	High(97)
Yocum Creek, near Oak Grove (Pass 11a), Sec. 30	USA	Arkansas	Carroll	36.454986	-93.322008	no	Low(35)
Village Creek State Park, S of driving range, Sec. 6	USA	Arkansas	Cross	35.16111	-90.70833	no	Low(39)
Sugar Creek, Hwy 163 at Bay Village, Sec. 4	USA	Arkansas	Cross	35.44909	-90.67533	no	High(100)
Buck Creek, 8.0 miles SE Corydon	USA	Indiana	Harrison	38.155118	-86.014724	no	High(88)
E Branch Mill Creek, Hessdale Road, 4.0 km S of Allendorph, Sec. 36	USA	Kansas	Wabaunsee	39.003564	-96.277745	no	High(88)
Blissdale Creek, Hillside National Wildlife Refuge, 500 m SW of Blissdale on Blis	USA	Mississippi	Holmes	33.083754	-90.224633	no	High(84)

Search: | File management | Phoxinus_erythrogaster_Locations_MMN... | Showing 1 to 8 of 44 records

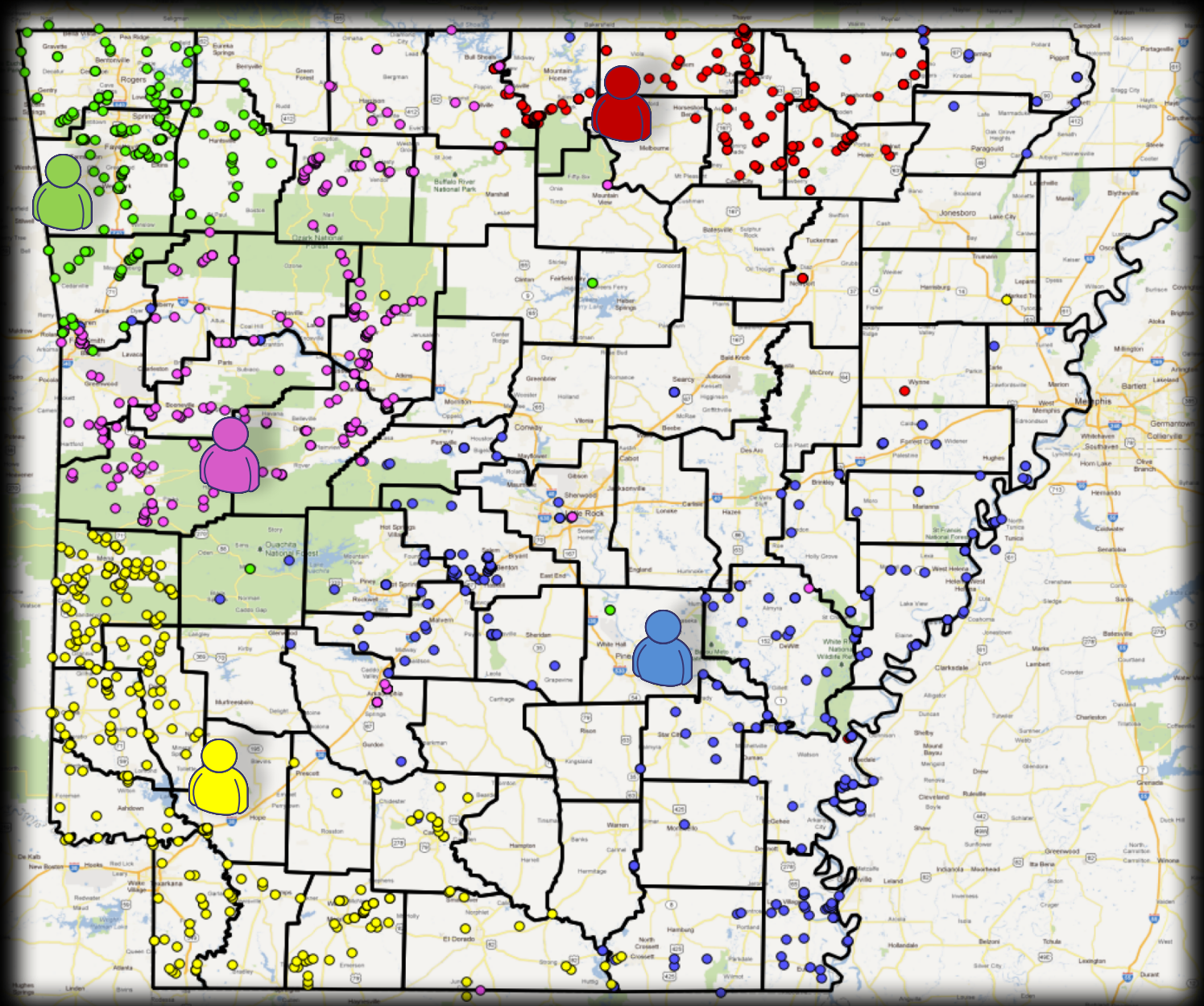
Collaborative georeferencing



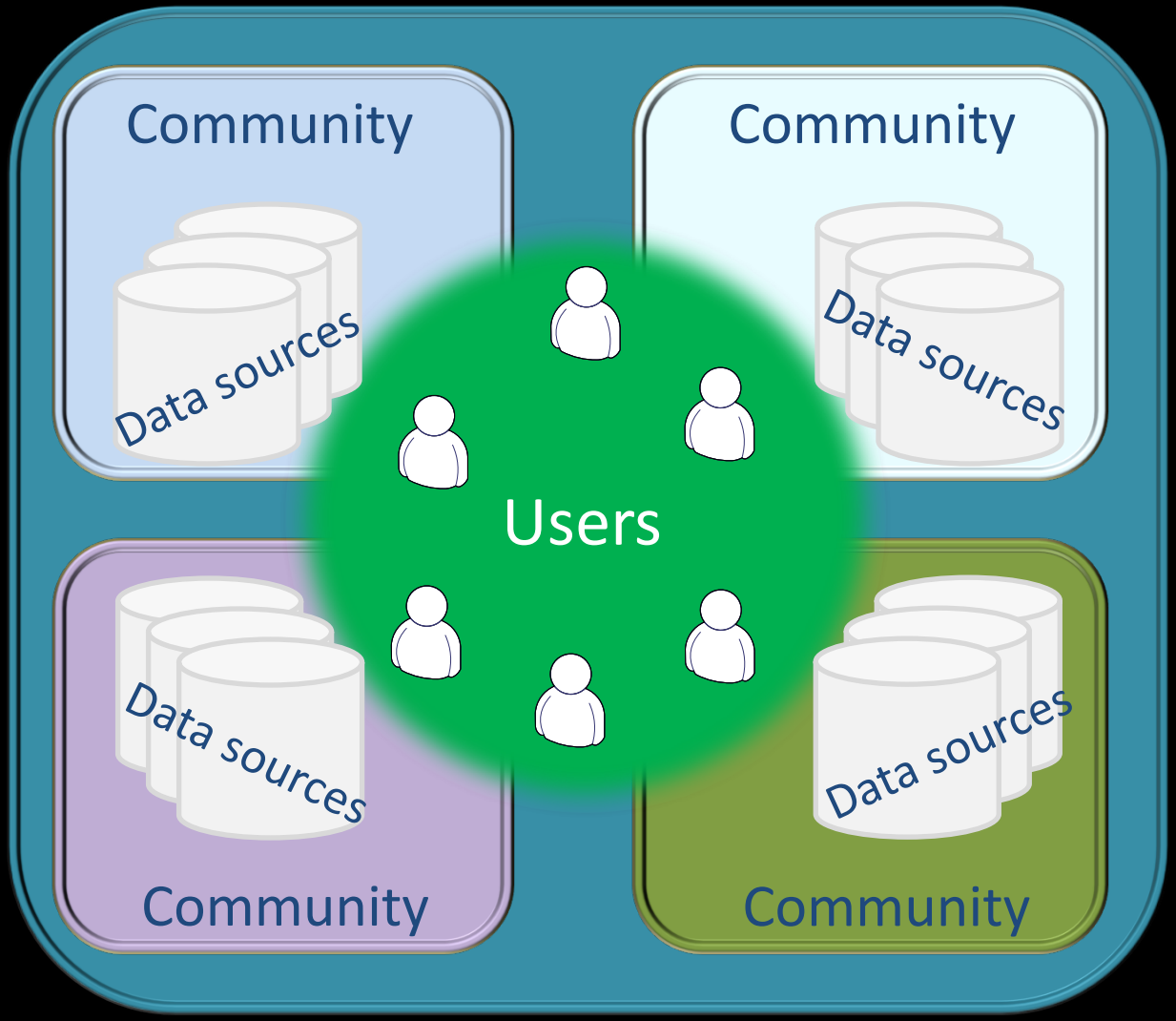
- Data management platform for large scale georeferencing
- Increased output by taking advantages similarities across collections
- Distribution of workloads to appropriate expertise

Similarity Matching

Distributing Workloads



Organizational Units



FishNet 2

DATA SOURCES MEMBERS SNAPSHOT SETTINGS

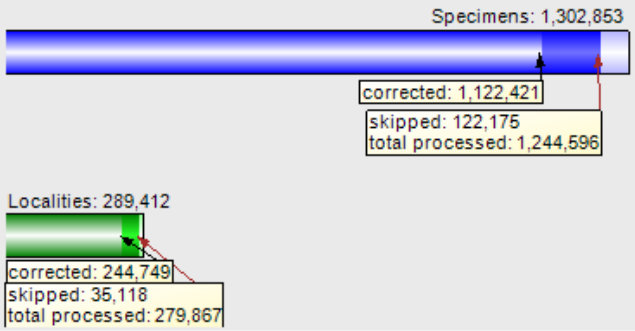
Data source management operations

- Add new community data source via CSV files

Click on an item's header to expand/collapse its content. sort by date

community-wide

Community Owner:	you
Community description:	Collaborative georeferencing of data from FishNet 2
Number of community data sources:	168
Records statistics:	

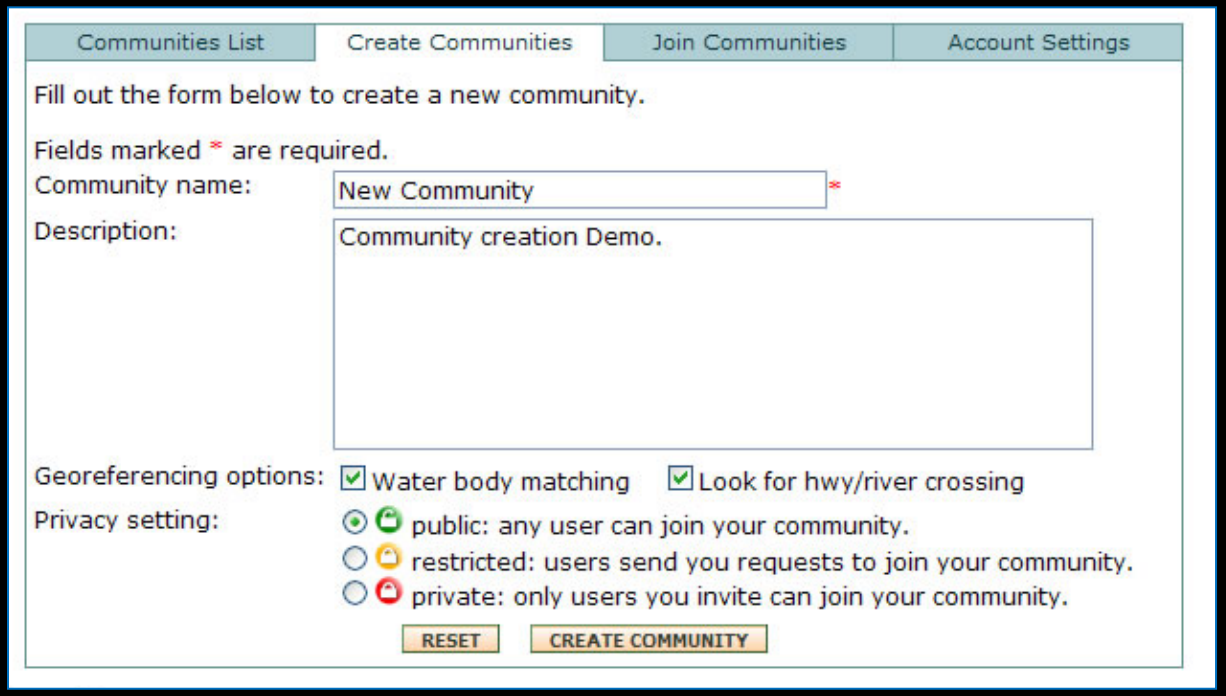


VIEW RECORDS

DOWNLOAD RECORDS

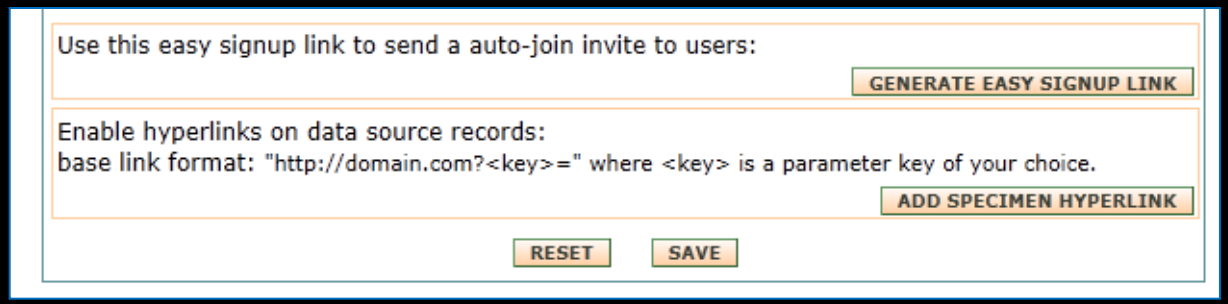
Iowa_with_county	date added: Monday, January 07, 2013	delete
Maine_with_county	date added: Monday, January 07, 2013	delete
Minnesota_with_county	date added: Monday, January 07, 2013	delete
Montana_with_county	date added: Monday, January 07, 2013	delete
Nevada_with_county	date added: Monday, January 07, 2013	delete

Georeferencing Communities



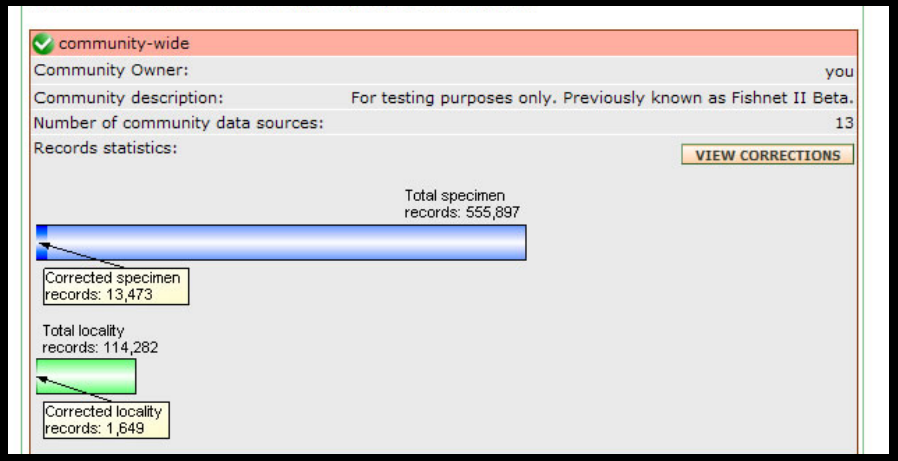
Create Communities

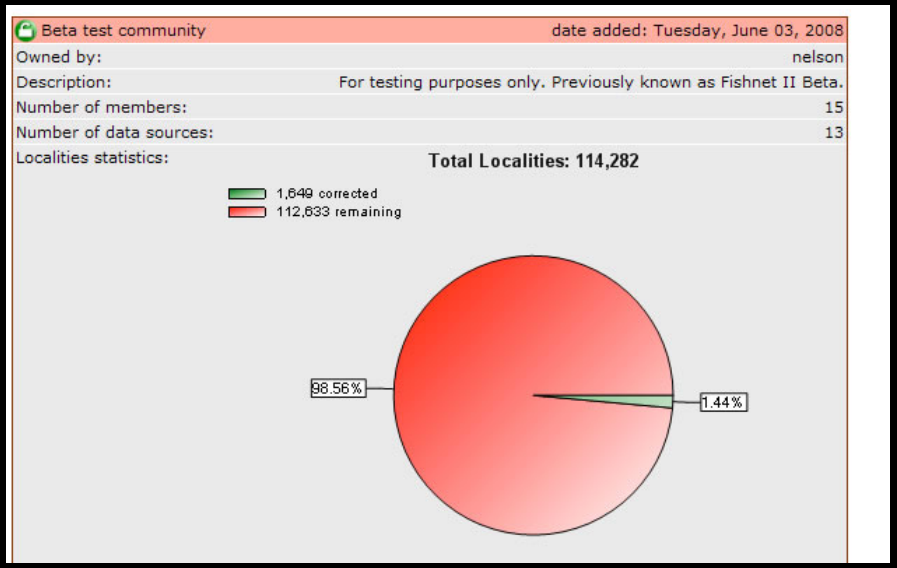
Data Sources

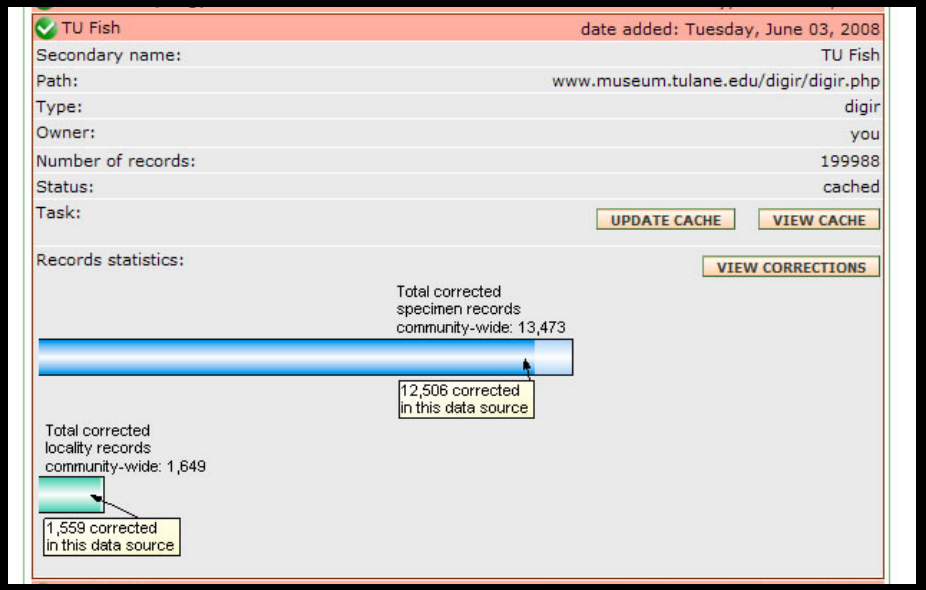


Add New Users & Link Out

Monitoring & Managing Progress

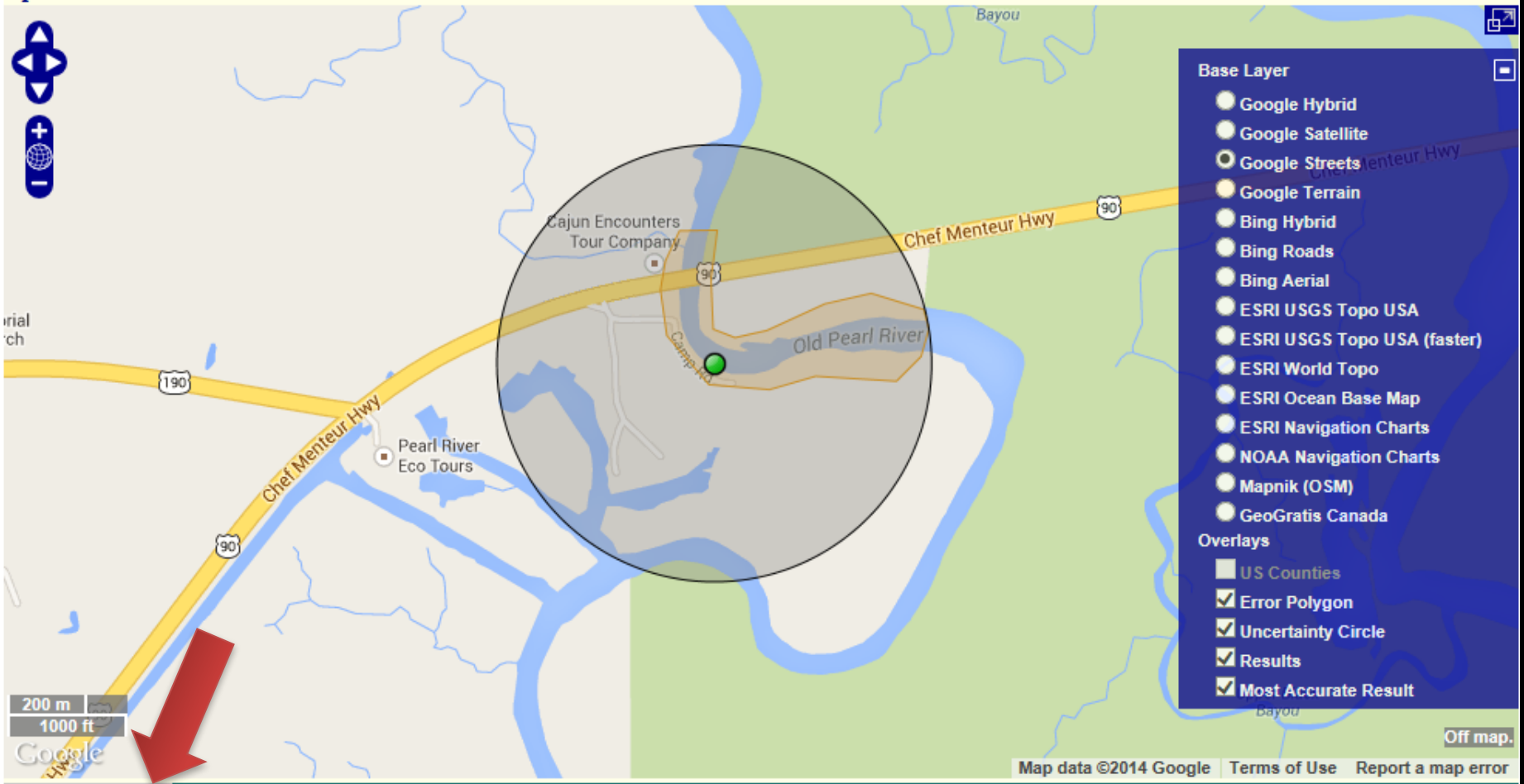






GEOLocate Collaborative Georeferencing Web Client

1 possible location found.



Workbench | 1 possible location found | History | Review

Community: FishNet 2

Clear Polygon | Draw polygon | Place marker | Measure

Next Record(s) | Correct | Skip Selected | Add Comments | Clear Workbench | Correction Report

Luggar Landing, West Pearl River, Starlding; United States; Louisiana; ;

FMNH14004 Fundulus chrysotus

Similar Records(5)

- Luggar Landing, West Pearl River near Starlding; United States; Louisiana; ;
- West Pearl River, Stardling, Lugger Landing, Deer Island; United States; Louisiana; ;
- Luggar Landing, West Pearl River, near Starlding; United States; Louisiana; ;

Logout | logged in as nelson

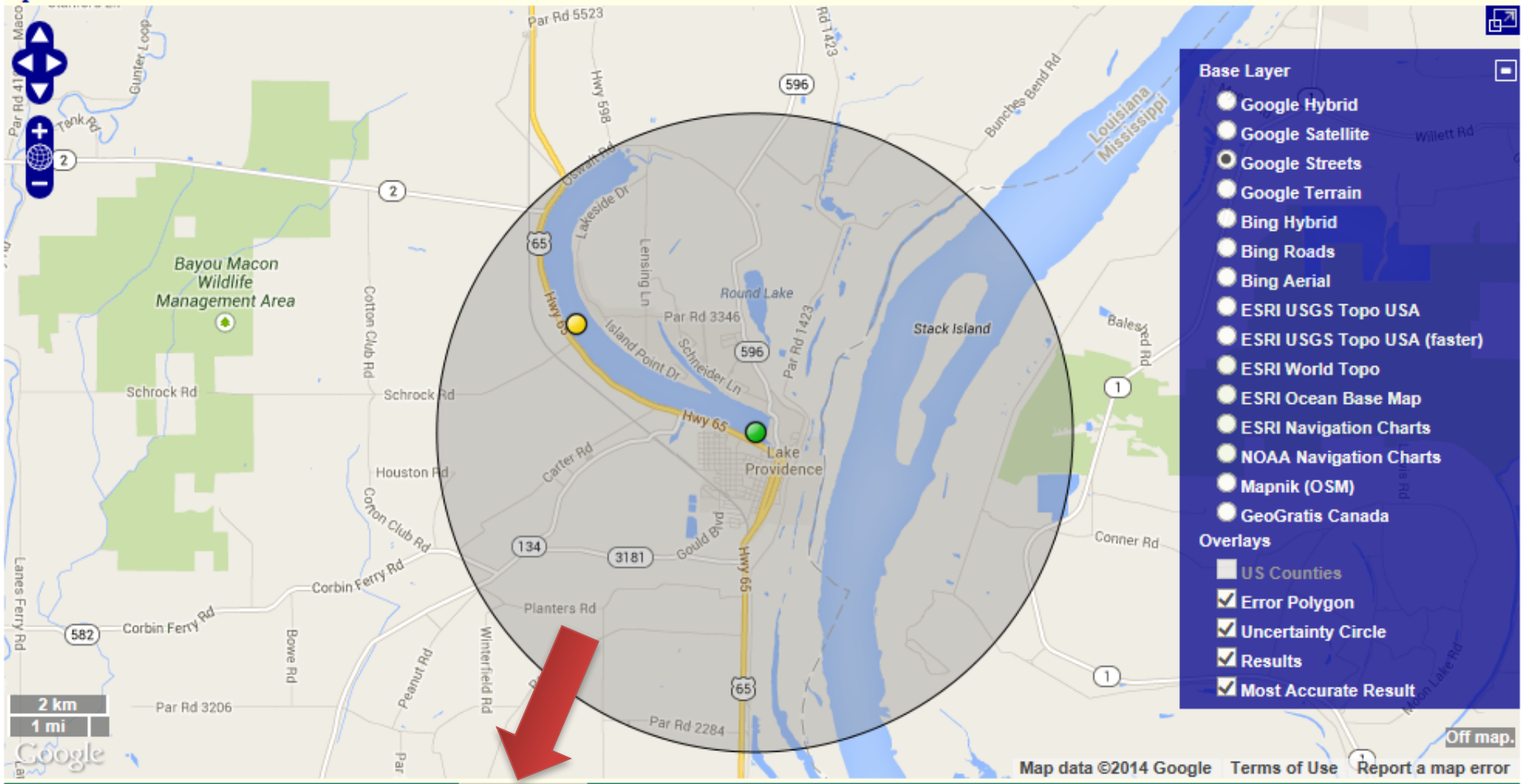
Calculated Coordinates

Lat: 30.229134
Lon: -89.668154
U. Radius: 581 m

Manual Edit

GEOLocate Collaborative Georeferencing Web Client

1 possible location found.



Workbench | 1 possible location found | History | Review

Manage History

show 10 records

get history

history filter

Options

- [ ] get skipped
- [x] get corrected

...get previous set...

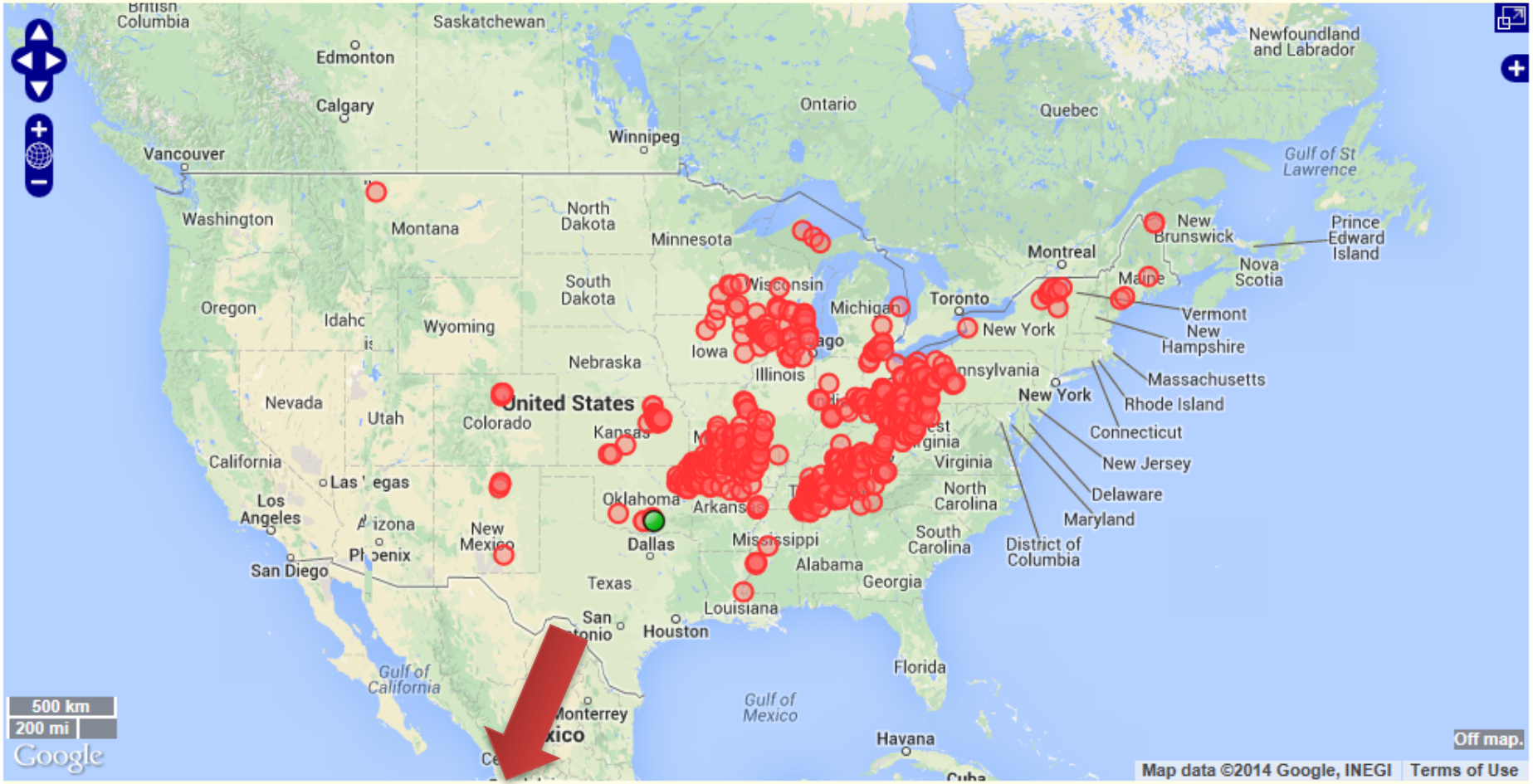
corrected on 9/10/2013 4:17:36 pm Avery Island.; United States; Louisiana; Iberia;
corrected on 9/10/2013 4:17:36 pm Avery Island; United States; Louisiana; Iberia;
corrected on 9/10/2013 4:14:36 pm Lake Providence, South Shore; 6 to 12 inches in depth; sand bottom; United States; Louisia
corrected on 9/10/2013 4:13:59 pm Lake Providence; United States; Louisiana; East Carroll;
corrected on 9/10/2013 4:11:19 pm Meridian Cr. on LA 348, 1 mi. E of Conway; T22N-R1W-Sec.17; United States; Louisiana;
corrected on 9/10/2013 4:09:59 pm Tickfaw R., 0.5 mi. W of Holden on US 190; United States; Louisiana; Livingston;

move to workbench

Options

- [x] include localities with matching determination
- [ ] exclude outdated determinations

GEOLocate Collaborative Georeferencing Web Client



Workbench | Results | History | Review

Active Related

1 collecting event at that location.

Blue River Prosper 1 mi E of St Rte 99 & Connerville.; United States; Oklahoma; Johnston

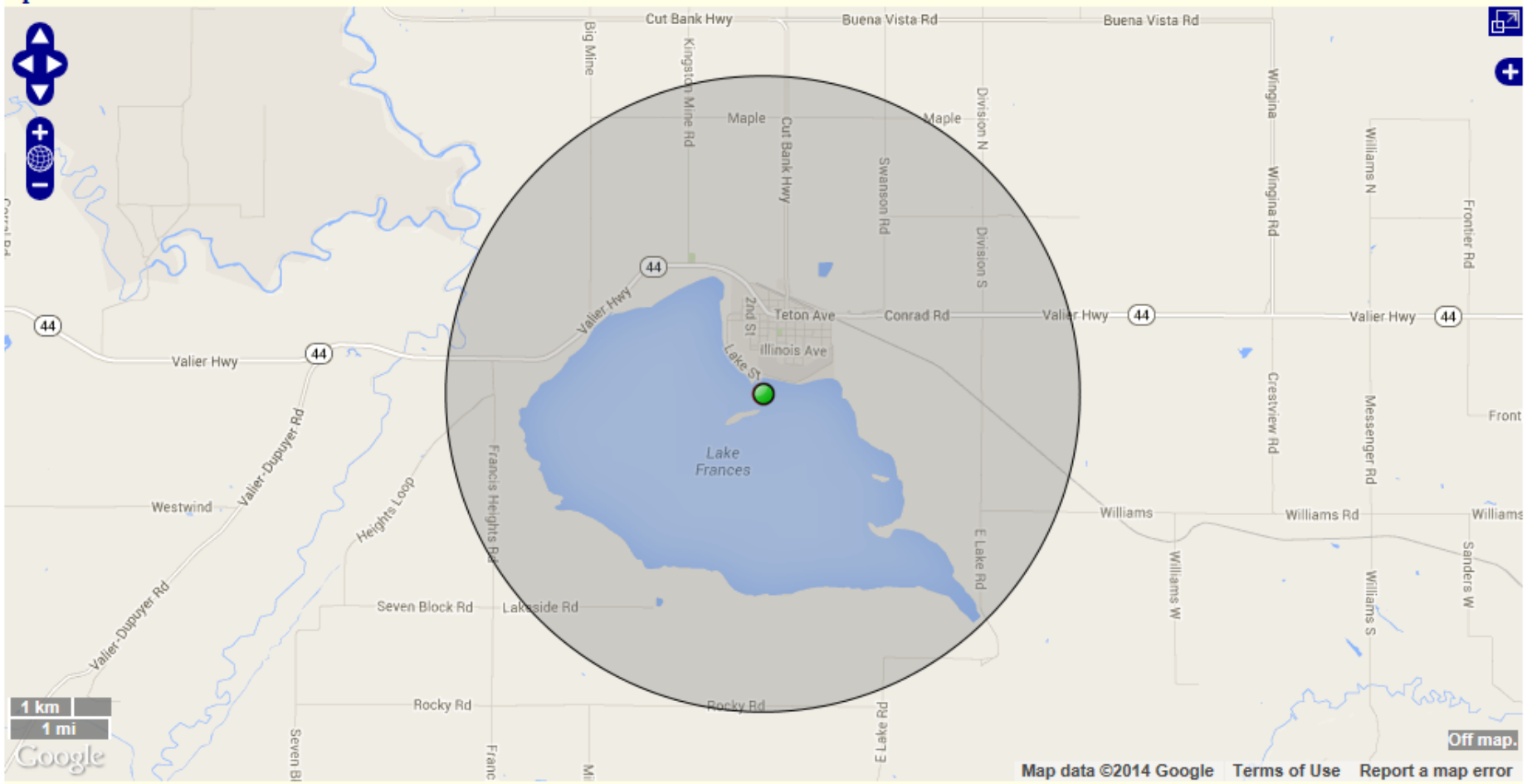
Reviewing records for scientific name 'PHOXINUS ERYTHROGASTER'. Edit...

Comments:

< Page 1 of 1 > move to workbench

GEOLocate Collaborative Georeferencing Web Client

1 possible location found.



Workbench | 1 possible location found | History | Review

Community: FishNet 2

Draw polygon | Place marker | Measure

Next Record(s)
Correct
Skip Selected
Add Comments
Clear Workbench
Correction Report

LAKE FRANCES, 5 MI. W OF VALIER, IRRIGATION DITCH FEEDS THE LAKE AND FLOWS OUT OF

- *UWFCUW 002716 CATOSTOMUS CATOSTOMUS*
- *UWFCUW 002717 CATOSTOMUS COMMERSONI*
- *UWFCUW 002719 COUESIUS PLUMBEUS*
- *UWFCUW 002715 ONCORHYNCHUS CLARKII*

Logout | logged in as nelson

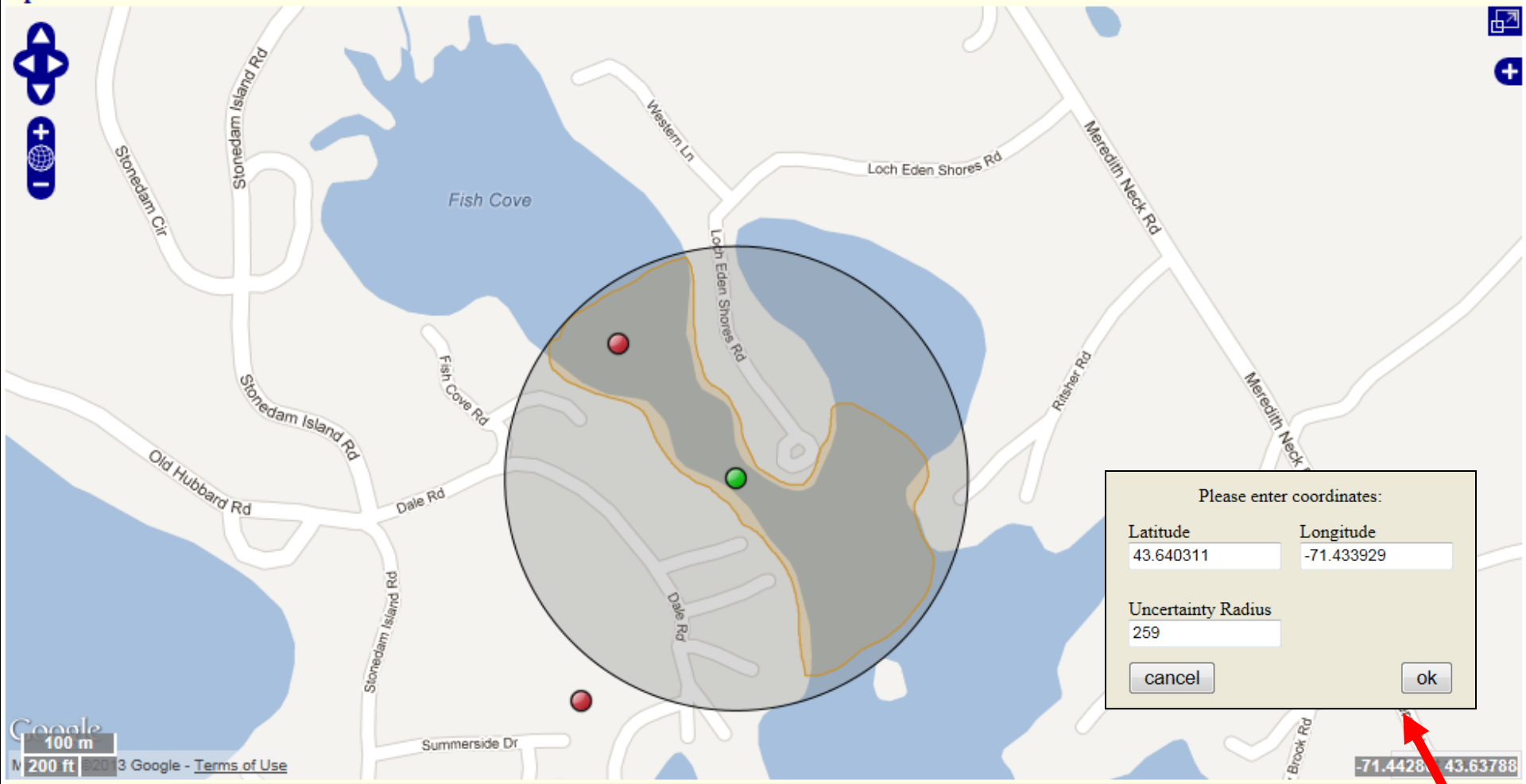
Calculated Coordinates

Lat: 48.296377
Lon: -112.255363
U. Radius: 5214 m

Manual Edit

GEOLocate Collaborative Georeferencing Web Client

6 possible locations found.



Workbench | 6 possible locations found | History

Community: FishNet 2

Clear Polygon | Draw polygon | Place marker | Measure

Next Record(s)
Correct
Skip Selected
Add Comments
Correction Report
Logout

Lake Winnipesaukee, Fish Cove near mouth of T 108, 3.5 mi ESE of Meredith, Merrimack Watershed, elev. 50

Similar Records(2)

Lake Winnipesaukee, Smalls Cove, Merrimack system, 5.5 mi E of Gilford, elev. 504 ft; United States; New Ha

UMMZ140837 Catostomus commersonii commersoni

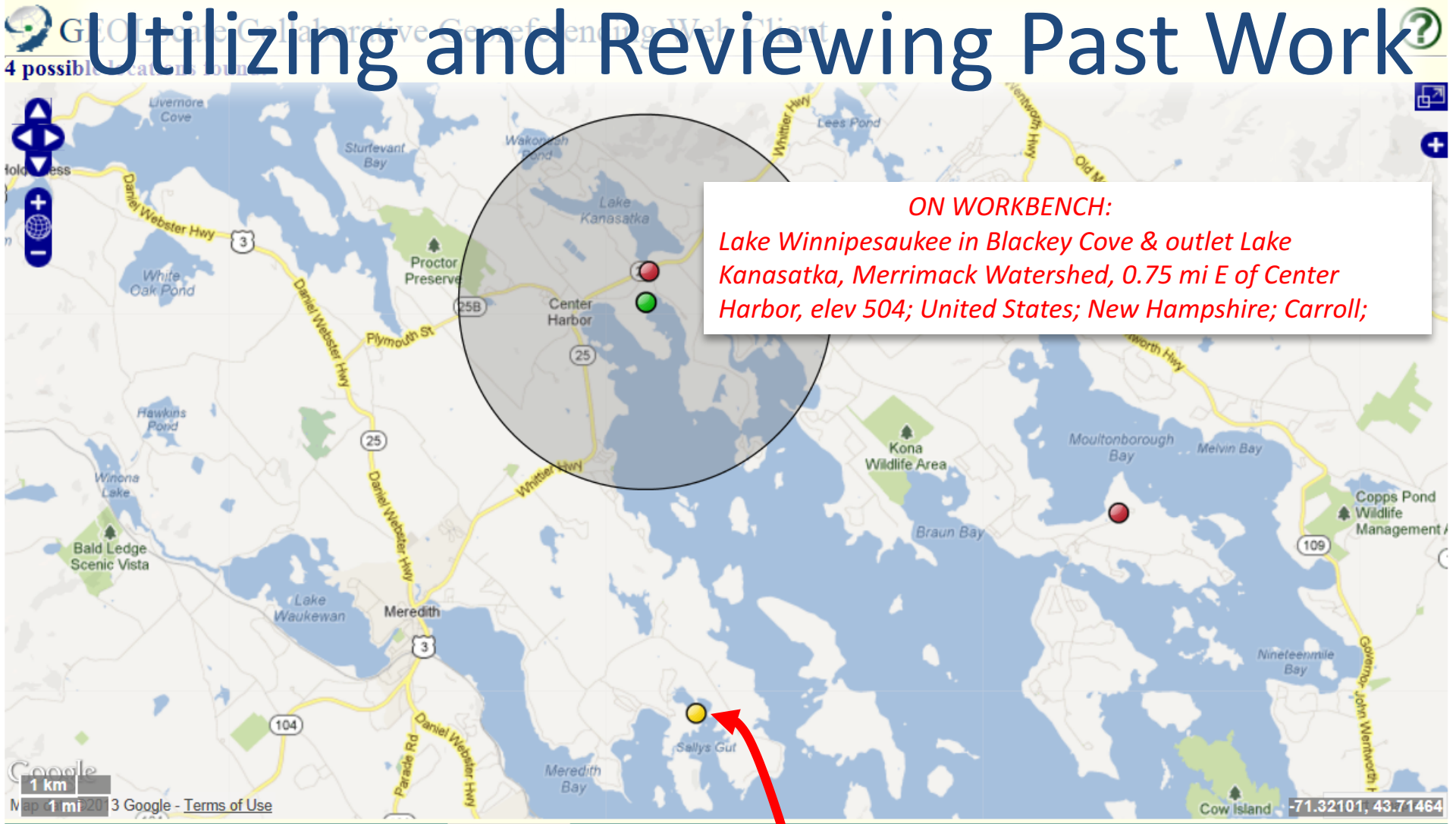
UMMZ140838 Esox niger

logged in as: nelson

Calculated Coordinates
Lon: -71.433929
Lat: 43.640311
U. Radius: 259 m

Manual Edit

Utilizing and Reviewing Past Work



Workbench | 4 possible locations found | History

show 100 records

get history

Options

- [ ] get skipped
- [x] get corrected

corrected on 3/2/2013 10:33:16 am Lake Winnipesaukee, Fish Cove near mouth of T 108, 3.5 mi ESE of Meredith, Merrimack
corrected on 3/1/2013 2:02:10 pm Calcasieu River, 6.3 miles W of Jct of LA Hwy 113 and US 165.; United States; Louisiana; R
corrected on 3/1/2013 2:02:10 pm Calcasieu River, 6.3 mi. W. of Jct. of LA Hwy 113 & US 165; United States; Louisiana; Rap
corrected on 3/1/2013 1:48:30 pm Calcasieu River on La 113; United States; Louisiana; Rapides;
corrected on 3/1/2013 1:48:30 pm Calcasieu River on LA 113; United States; Louisiana; Rapides;
corrected on 3/1/2013 1:48:30 pm Calcasieu River, on LA 113; United States; Louisiana; Rapides;
corrected on 3/1/2013 1:29:16 pm Soucook River, Merrimack system, 3 mi SSE of Concord, elev. 230 ft; United States; New H

move to workbench

Options

- [ ] include localities with matching determination
- [ ] exclude outdated determinations

Interoperability

- **Webservices**
  - **SOAP**
  - **JSON**
  - **GeoJSON**

- **Web Client**

Arctos

Symbiota

Specify 6

Google

```
HTTP/1.1 200 OK
Content-Type: text/xml; charset=utf-8
Content-Length: length

<?xml version="1.0" encoding="utf-8"?>
<Georef_Result_Set xmlns="http://www.museum.tulane.edu/webservices/">
  <EngineVersion>string</EngineVersion>
  <NumResults>int</NumResults>
  <ExecutionTimems>double</ExecutionTimems>
  <ResultSet>
    <WGS84Coordinate>
      <Latitude>double</Latitude>
      <Longitude>double</Longitude>
    </WGS84Coordinate>
    <ParsePattern>string</ParsePattern>
    <Precision>string</Precision>
    <Score>int</Score>
    <UncertaintyRadiusMeters>string</UncertaintyRadiusMeters>
    <UncertaintyPolygon>string</UncertaintyPolygon>
    <ReferenceLocation>string</ReferenceLocation>
    <DisplacedDistanceMiles>double</DisplacedDistanceMiles>
    <DisplacedHeadingDegrees>double</DisplacedHeadingDegrees>
    <Debug>string</Debug>
  </ResultSet>
  <ResultSet>
    .
    .
    .
  </ResultSet>
</Georef_Result_Set>
```

Specify 6.4.08 (EZDB)

File Edit Data System Tabs Help

Welcome Data Trees Reports Interactions Statistics Query Workbench SGR Lifemapper Search

Actions
- Import Data
- Import Images
- New Data Set
- Export Data Set
- Export MS Excel Template
- Create Data Set from Record Set

Reports
- Basic Label
- Data Set Summary
- Chart
- Top 10

Data Sets
- Sample2

Trash

	LocalityString	Country	State	County	Lat	Lon	uncertainty
1	Cable Station at Cape York	Australia	Queensland		-10.68850	142.53167	
2	Aberdeen Lake	Canada			64.46667	-99.00000	
3	Pioneer Park near Davis	USA	CA	Yolo			
4	Conroe	USA	TX				
5	Twin Creek at US Hwy 24	USA	CO	Teller			
6	Turkey Creek at Goodwin Road.	USA	AL	Jefferson			
7	Borden Creek at Forest Road 244, Sipsey Wilderness; 13.1 mi SSW M...	USA	AL	Lawrence			
8	Moon Lake	USA	MS	Coahoma			
9	Red River along right bank at River Mile 105; Station 1.	USA	LA	Rapides			
10	Bienville National Forest on US Hwy 20	USA	MS				
11	Zion National Park	USA	UT				
12	Whitetail state wildlife management area	USA	NE				
13	Lincoln Beach	USA	OR	Lincoln			
14	Green River near neatsville	USA	KY	Adair			
15	New Milan Reservoir, 1.75 miles north of milan	USA	MO				
16	Homochitto River, 1.5 mi. S of Bude, Hwy. 98.	USA	MS	Franklin			
17	Homochitto River 5 mi. E of Union Church, MS Hwy. 550.	USA	MS	Lincoln			
18	Mamala Bay	USA	HI				
19	Bay of bengal	India					
20	Henley On Thames	United Kingdom	Oxfordshire				

GEOLocate Progress

Process: 19 / 20

Cancel

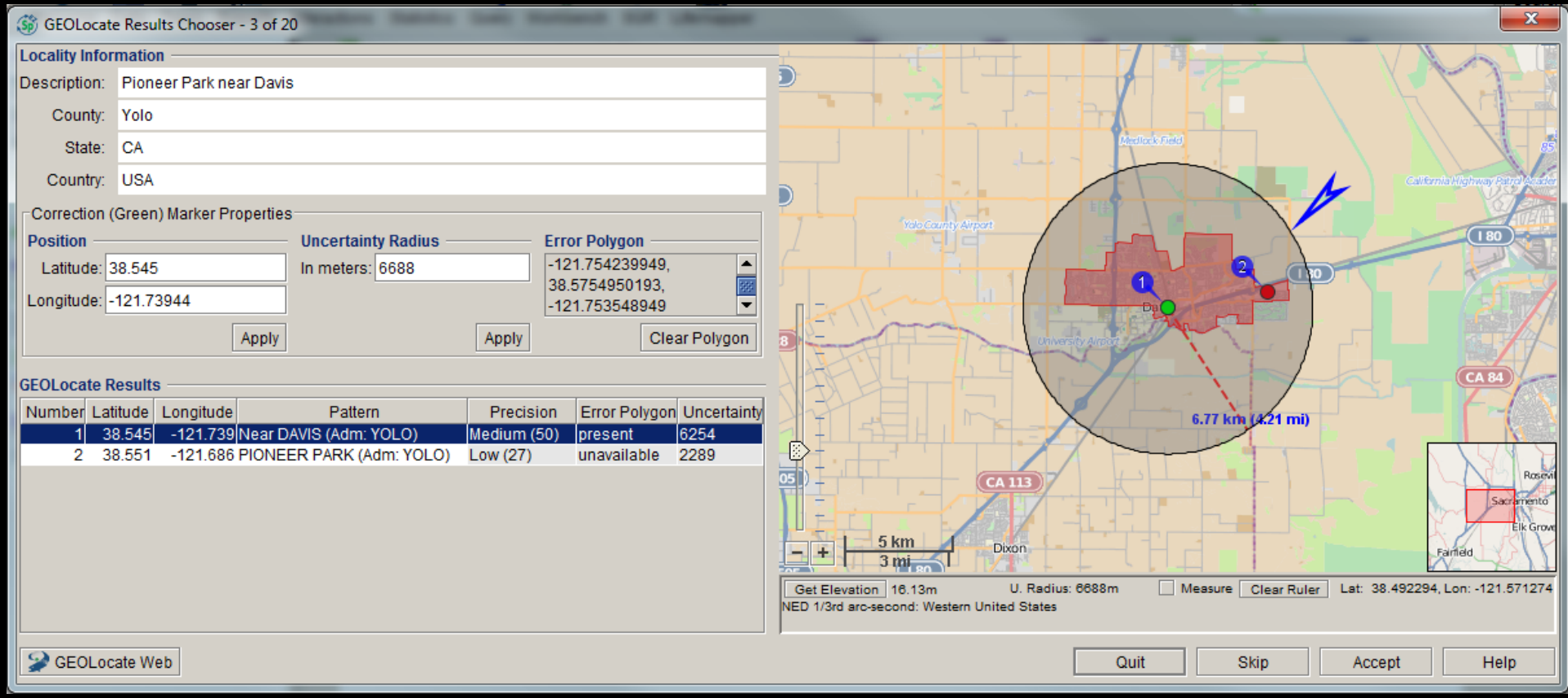
Save Form

Highlight Invalid Cells : 0 Highlight New Records

Welcome Sample2

Ichthyology DJ's Specify Test Collection djihbril

Specify 6 + GEOLocate



Kepler + GEOLocate

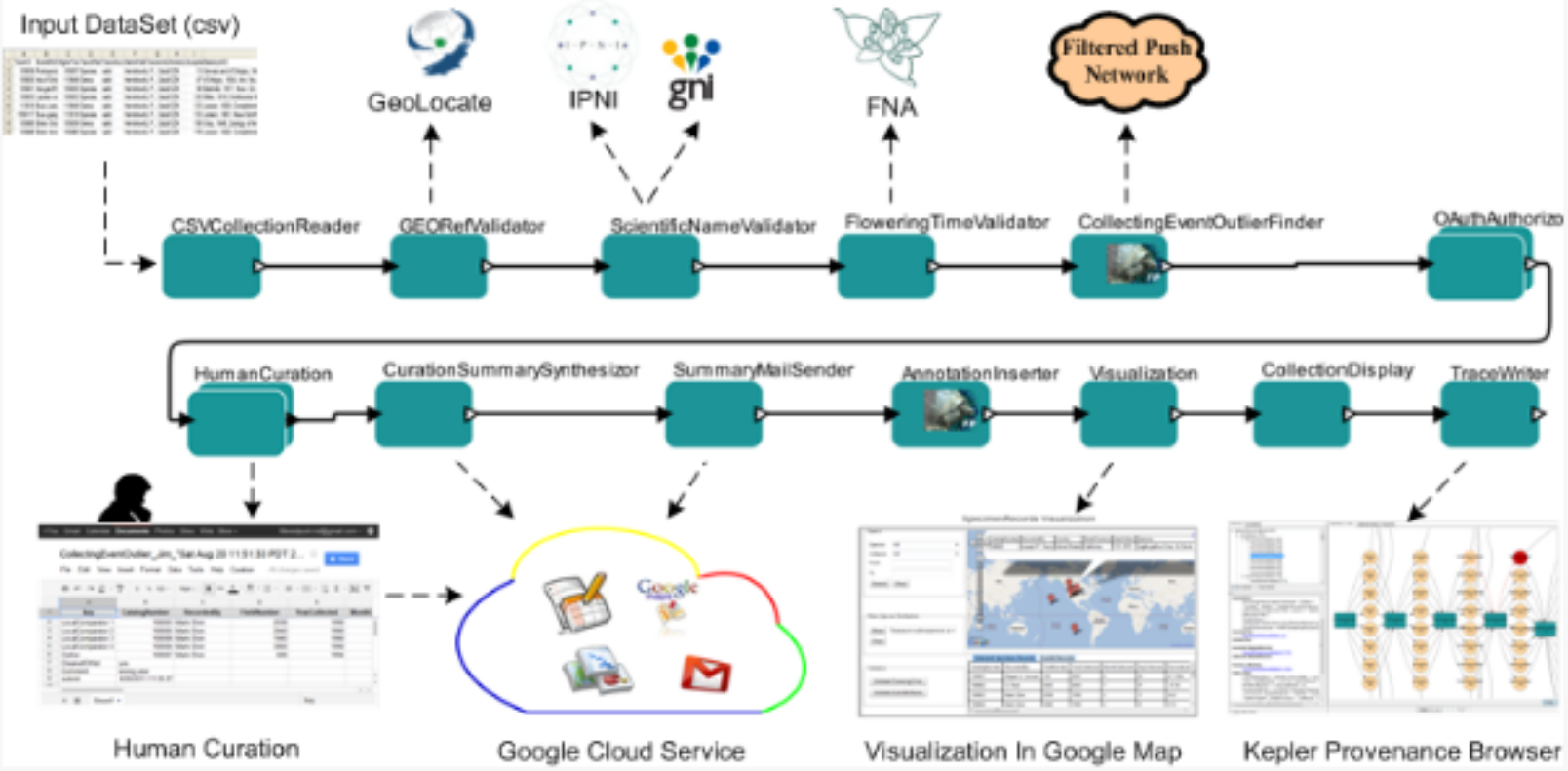
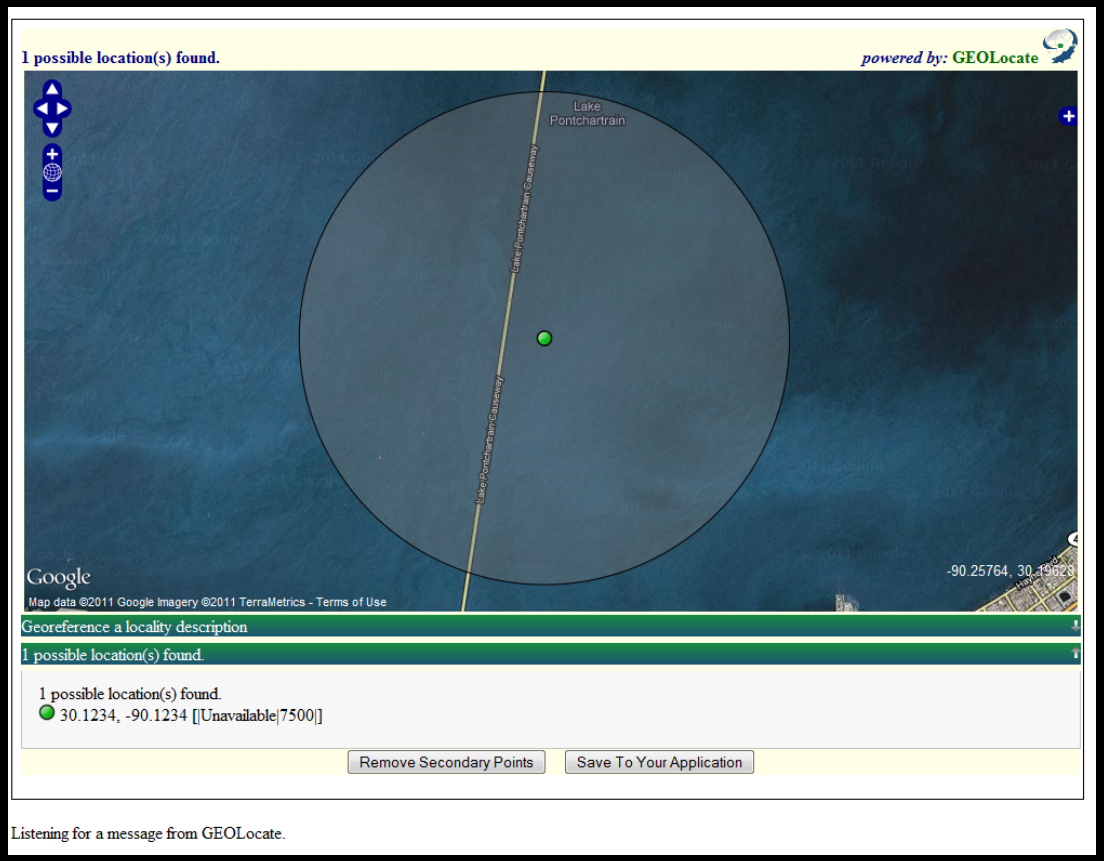


Fig1. Kepler Curation Workflow

```
library(RJSONIO)
library(RCurl)

setwd("E:/FishNet2GeoRef")

INPUTFILENAME="t_localities4GLC2.csv"
OUTPUTFILENAME="t_localities4GLC_out.csv"
OUTPUTFILENAMEFIRSTRESULT="t_localities4GLC_out_first_result.csv"


OPTIONS="&doduncert=true&dopoly=false&displacepoly=false"
glcIn= read.csv(INPUTFILENAME)
numGLCRuns = 0
recordCounter = 0
for (k in 1:nrow(glcIn)){
  print(k)
        Sys.sleep(3) #be nice and pause a few seconds between requests
        Country=glcIn[k,]$Country
        Locality=glcIn[k,]$Locality
        StateProvince=glcIn[k,]$StateProvince
        County=glcIn[k,]$County
        q=paste("http://www.museum.tulane.edu/webservices/geolocatesvcv2/glcwrap.aspx?country=",Cou
        q=gsub(' ','%20',q)

  tryCatch({
    JSONresponse = basicTextGatherer()
    curlPerform(url = q, writefunction = JSONresponse$update)
    glcRecNum = k
    glc = fromJSON(JSONresponse$value())
    numresults = glc$numResults
    if (numresults > 0){
      for (i in 1:numresults) {
        glcRank  = i
        glcLongitude = glc$resultSet$features[[i]]$geometry$coordinates[1]
        glcLatitude = glc$resultSet$features[[i]]$geometry$coordinates[2]
        glcPrecision = glc$resultSet$features[[i]]$properties$precision
        glcScore = glc$resultSet$features[[i]]$properties$score
        glcParsepattern = glc$resultSet$features[[i]]$properties$parsePattern
        glcUncert = glc$resultSet$features[[i]]$properties$uncertaintyRadiusMeters
        glcPoly = glc$resultSet$features[[i]]$properties$uncertaintyPolygon
        #if a polygon is present reformat coordinates to geolocate format-a comma delimited array
        if ("coordinates"%in%names(glcPoly)){
          sPoly = ''
```

Application Services: Web Client APIs



- URL API for user input & lightweight web clients

- Lightweight client, specifically designed for embedding into other web applications.
  - Two way communication between web sites uses JavaScript postMessage()
  - Compatible with all modern browsers:
    - IE 8.0+
    - Firefox 3.0+
    - Safari 4.0+
    - Chrome 1.0+
    - Opera 9.5+

UAM:Mamm Cat# Accn [customize form]

Collector 1 Copy2All | Collector 2 Copy2All
Collector 3 Copy2All | Collector 4 Copy2All
Collector 5 Copy2All

OtherID 1 original identifier
OtherID 2
OtherID 3
OtherID 4

Scientific Name
ID By Copy2All
Nature ID of kin
Date Copy2All
ID Remk

Higher Geog North America, United States, Louisiana, Washington Parish
Spec Locality Bogalusa
Existing LocalityID Pick Locality Existing EventID Pick Event
Verbatim Locality Bogalusa
Use Specloc
VerbatimDate --> Begin Copy2All End Copy2All
Coll. Meth.: Coll. Src.: wild caught
Habitat
Associated Species
Microhabitat
Elevation (min-max) between >>
CollEvntRemk
LocalityRemk - checkPicked

Original lat/long Units [geolocate] geolocate messages go here

Sex Date Sync Att. Detr Sync Att. Meth
len tail Hind Foot Ear From Notch Units Weight Date Determiner
mm g

Attribute	Value	Units	Date	Determiner	Method	Remarks

Entered By gravier
Disposition being processed
Condition unchecked
Relations

Spec Remark

Missing.... Status UAM MAMM TEMPLATE

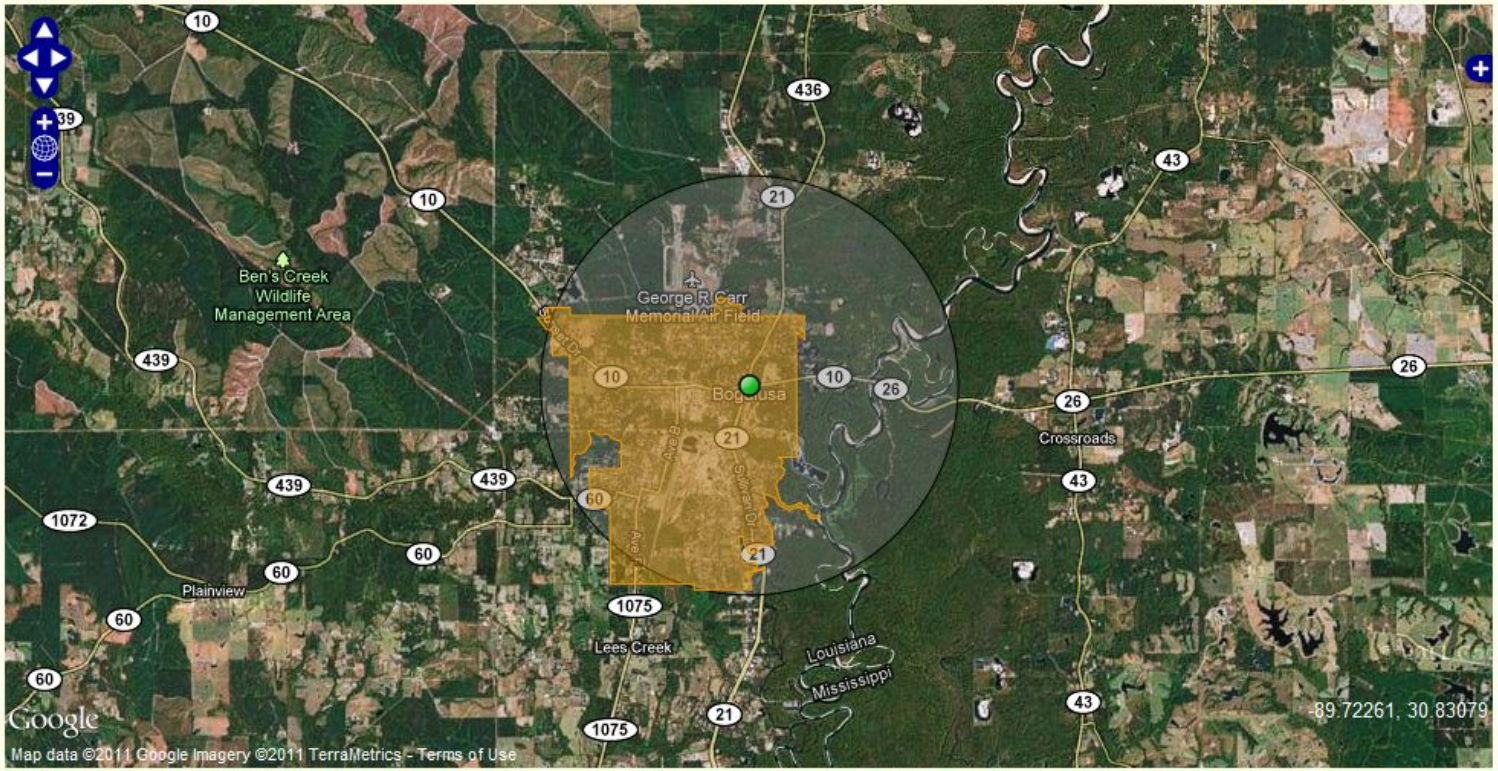
Part Name	Condition	Disposition	#	Barcode	Vial Label	Remark
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			
	unchecked	being processed	1			

Save This As A New Record | Edit Last Record

[help]

1 possible location(s) found.

powered by: **GEOLocate**



Georeference a locality description

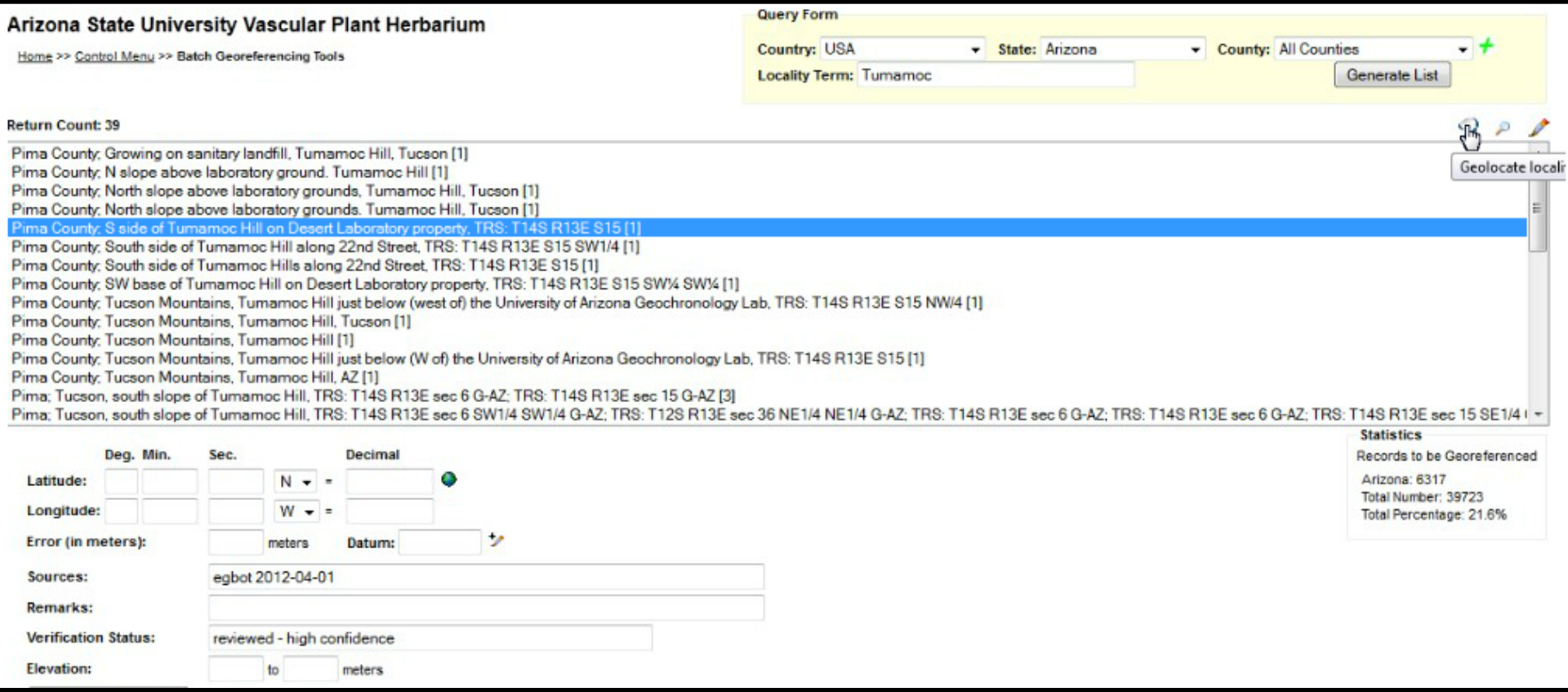
1 possible location(s) found.

1 possible location(s) found.

30.79083, -89.84861 [BOGALUSA|30.801736012,-89.8943429191...|4490|Low(23)]

Remove Secondary Points | Save To Your Application

Symbiota + GEOLocate

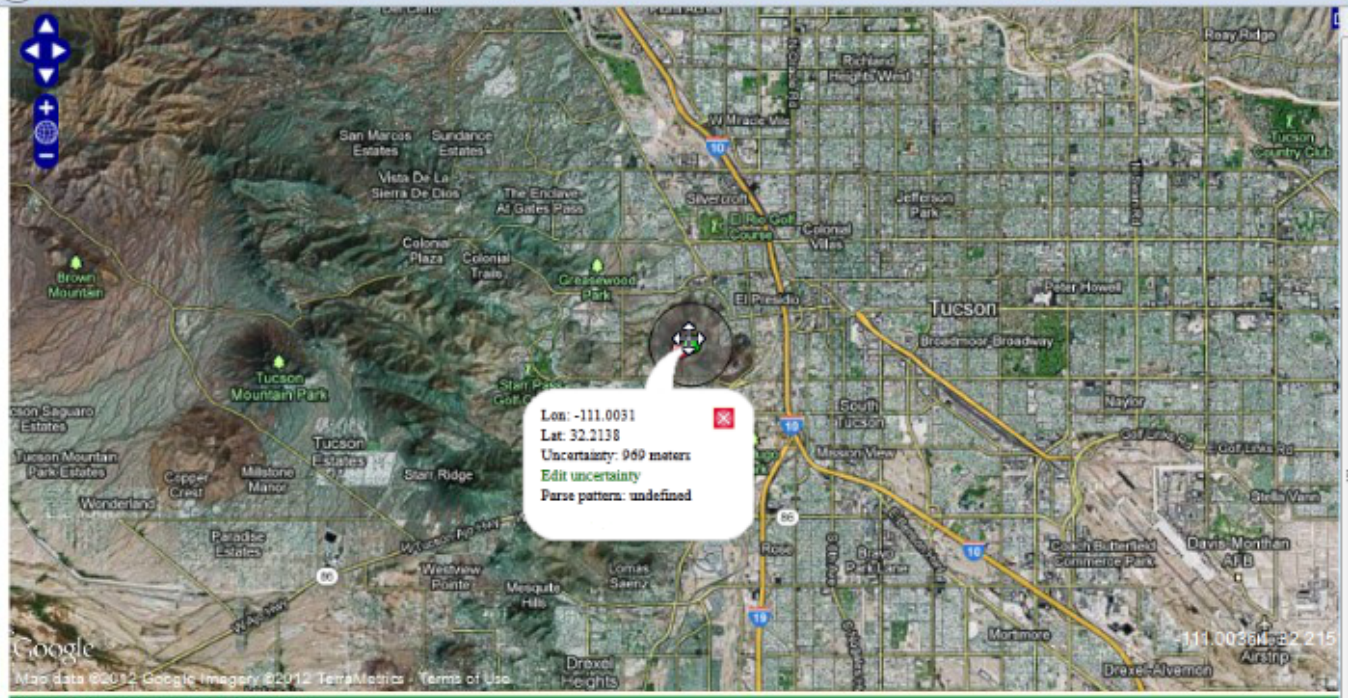


GEOLocate Tool - Mozilla Firefox

GEOLocate Tool

localhost:89/seinet/collections/georef/geolocate.php?country=USA&state=Arizona&county=&locality=

AVG Secure Search



Georeference a locality description

2 possible location(s) found.

2 possible location(s) found.

32.2138, -111.0031 [TRS|Unavailable|969|High(85)]

32.21278, -111.00556 [TUMAMOC HILL|Unavailable|1807|Low(27)]

Remove Secondary Points | Save To Your Application

County: All Counties

Generate List

13E sec 6 G-AZ; TRS: T14S R13E sec 15 SE1/4

Statistics

Records to be Georeferenced

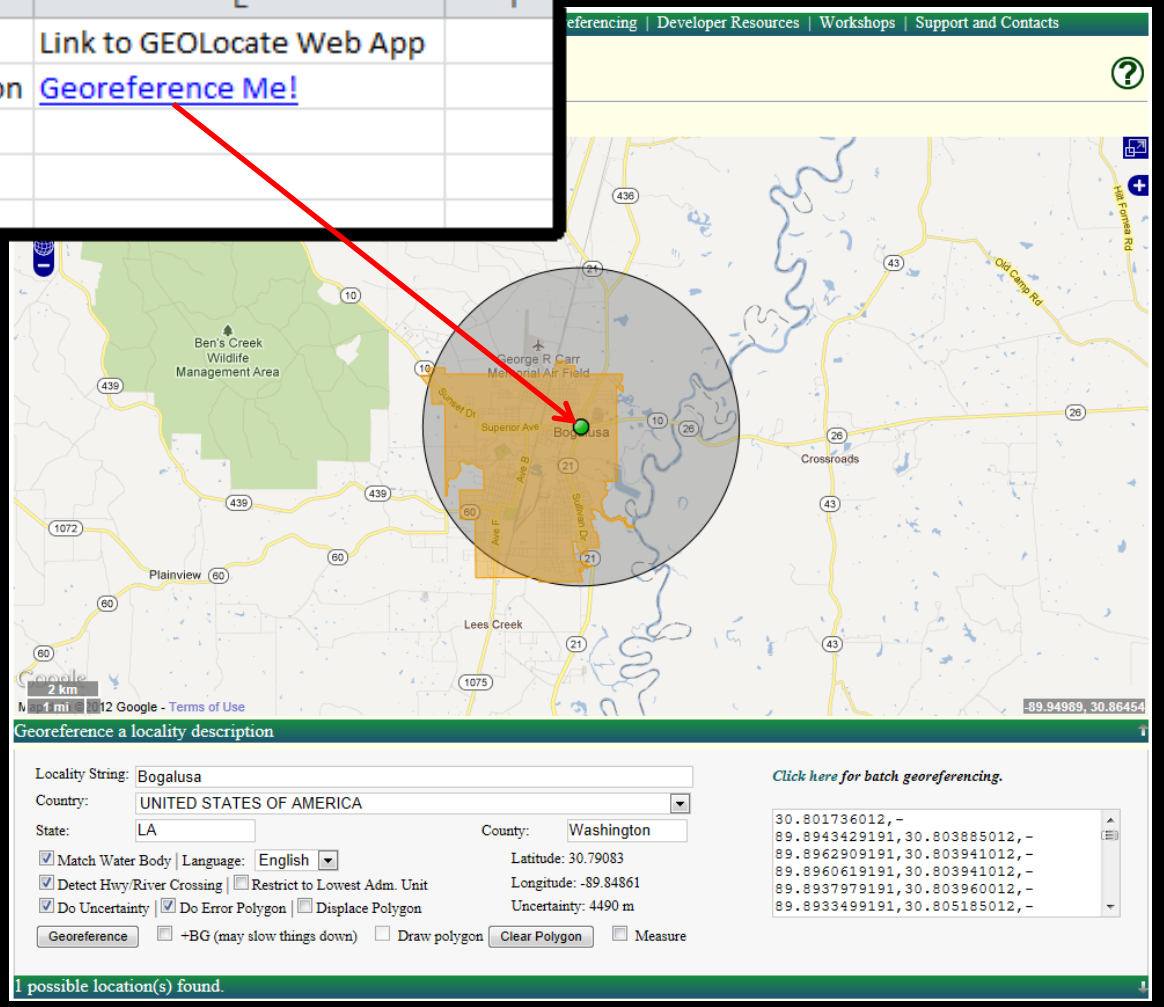
Arizona: 6317

Total Number: 39723

Total Percentage: 21.6%

Excel + GEOLocate

	A	B	C	D	E	F
1	Locality	Country	State	County	Link to GEOLocate Web App	
2	Bogalusa	USA	LA	Washington	Georeference Me!	
3						
4						
5						



http://www.museum.tulane.edu/geolocate/web/WebGeoref.aspx?v=1&Country=**USA**&State=**LA**&County=**Washington**&Locality=**Bogalusa**&georef=run