Webinar Series

Data Use Skills
Featuring Data from Natural History Collections

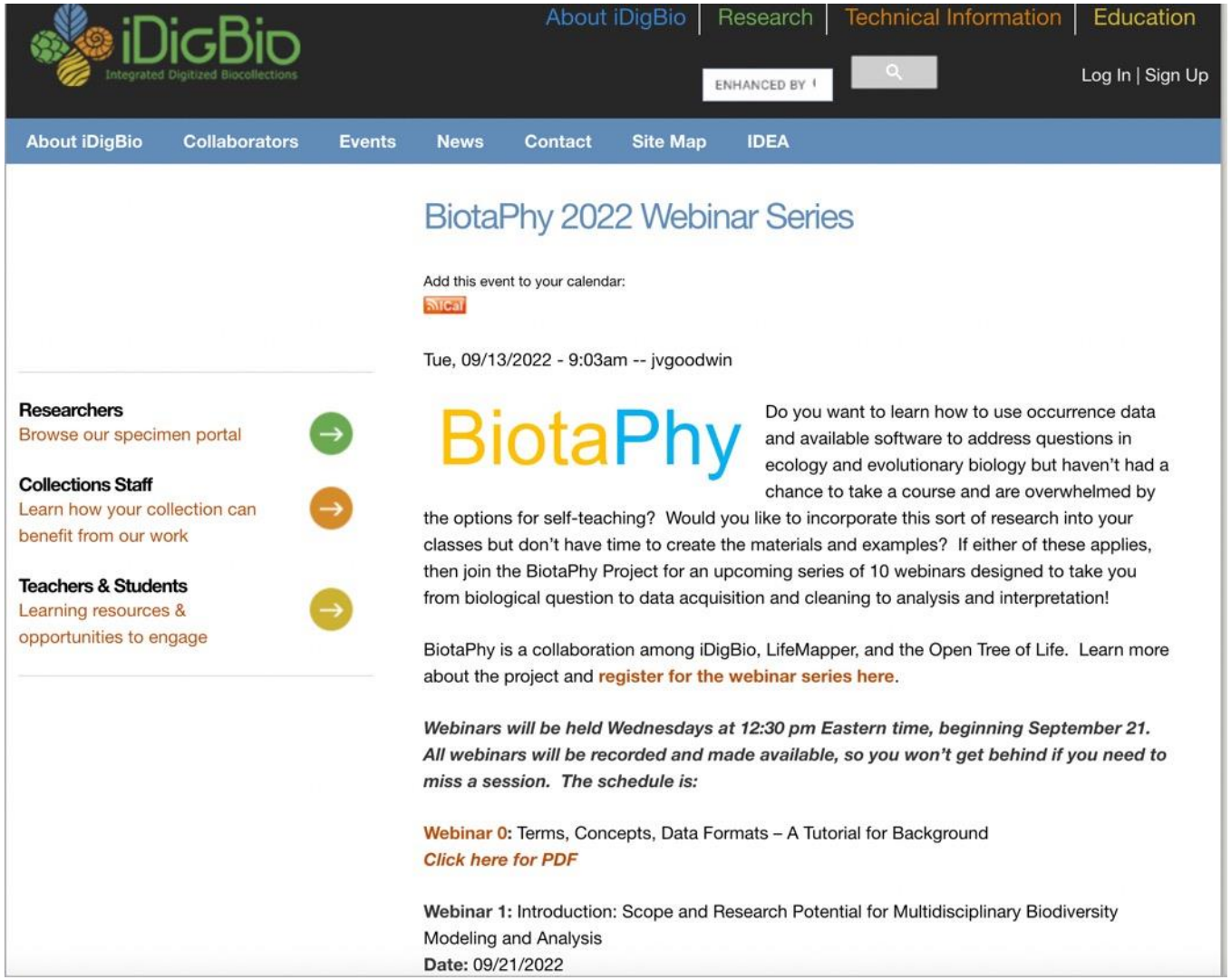
September 21-November 30, 2022

https://www.idigbio.org/content/biotaphy-2022-webinar-series

iDigBio:

https://www.idigbio.org/content/biotaphy-2022-webinar-series

iDigBio.org



Thank You

Maria Cortez
Aimee Stewart

Jill Goodwin
Gil Nelson

Webinar 4

Georeferencing with GEOLocate

Goals

How to use locality information in specimen records to obtain georeferences (i.e., latitude and longitude)

Learning Objectives

Biological Objectives:

- ✓ **Introduction to basic geospatial concepts and their relevance in biodiversity analyses**

- ✓ **Introduction to georeferencing and its importance for biodiversity analyses**

Learning Objectives

Technical Objectives:

- ✓ **Introduce GEOLocate as a tool to recover geographic coordinates based on locality information.**
- ✓ **Demonstrate and practice how to use GEOLocate**
- ✓ **Practice how to manually add modified georeferences to the spreadsheet containing geographic coordinates.**

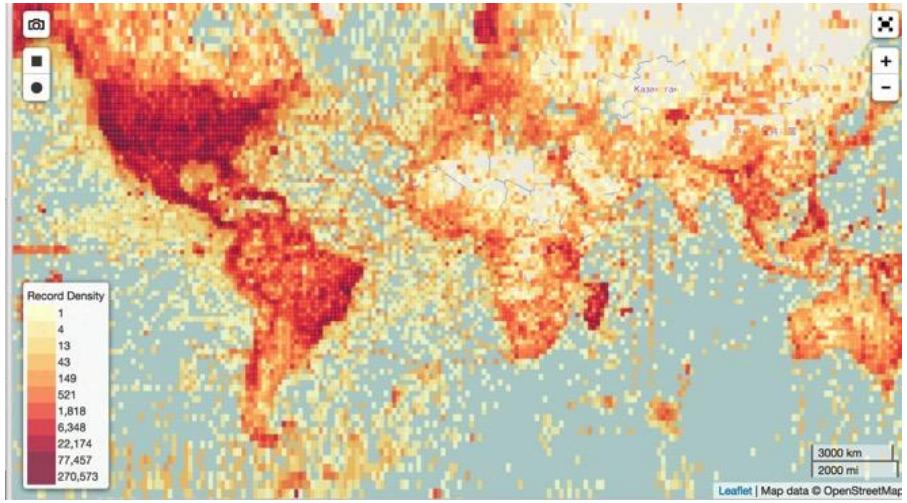
Webinar organization

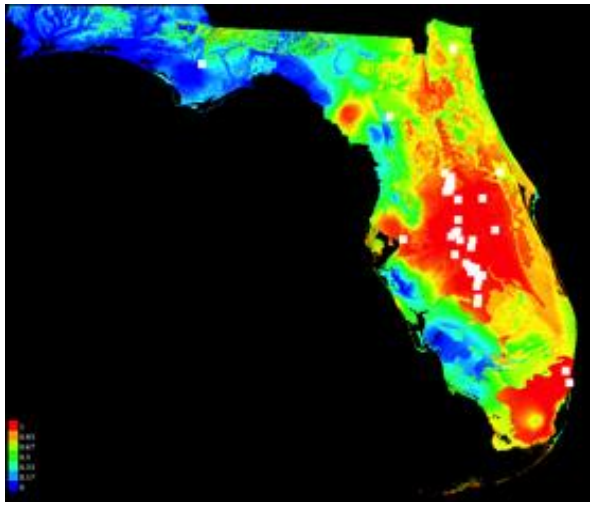
1. Exploring Concepts: understanding spatial concepts and GEOLocate
2. Demonstration: how to use GEOLocate
3. Exercises: using GEOLocate to recover geographic coordinates and merging the information recovered to your dataset!
4. Session Summary, Q&A and Discussion

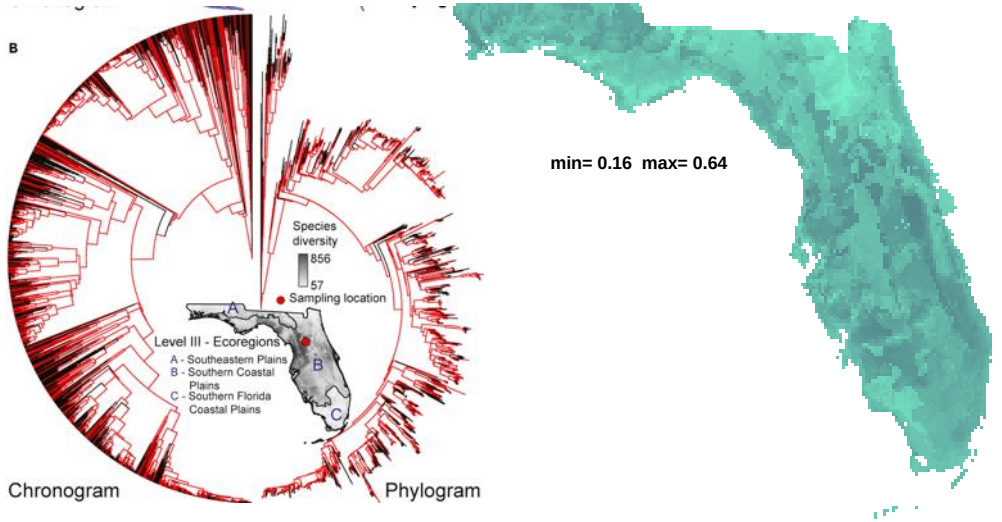
Learning Objectives

Biological Objectives:

- ✓ **Introduction to basic geospatial concepts and their relevance in biodiversity analyses**







Learning Objectives

Biological Objectives:

- ✓ **Introduction to basic geospatial concepts and their relevance in biodiversity analyses**

- ✓ **Introduction to georeferencing and its importance for biodiversity analyses**

Exploring Concepts

What is a georeference?

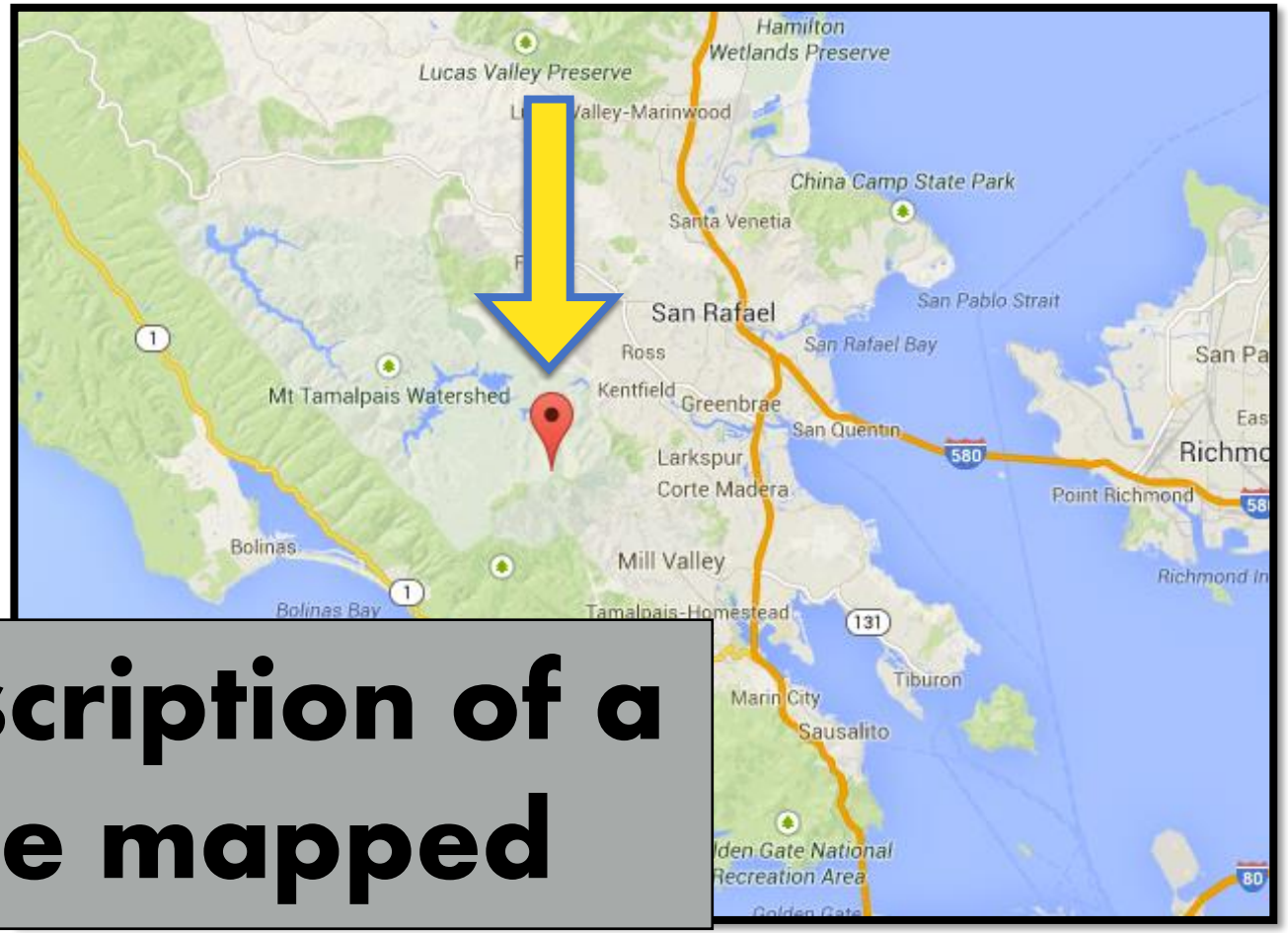
Exploring Concepts

What is a georeference?

Species: *Polypodium californicum*
State: CA
County: Marin
Locality: Marin Municipal Water District Lands: Mount Tamalpais. Middle Peak Rd. culvert RT-3

Exploring Concepts

What is a georeference?

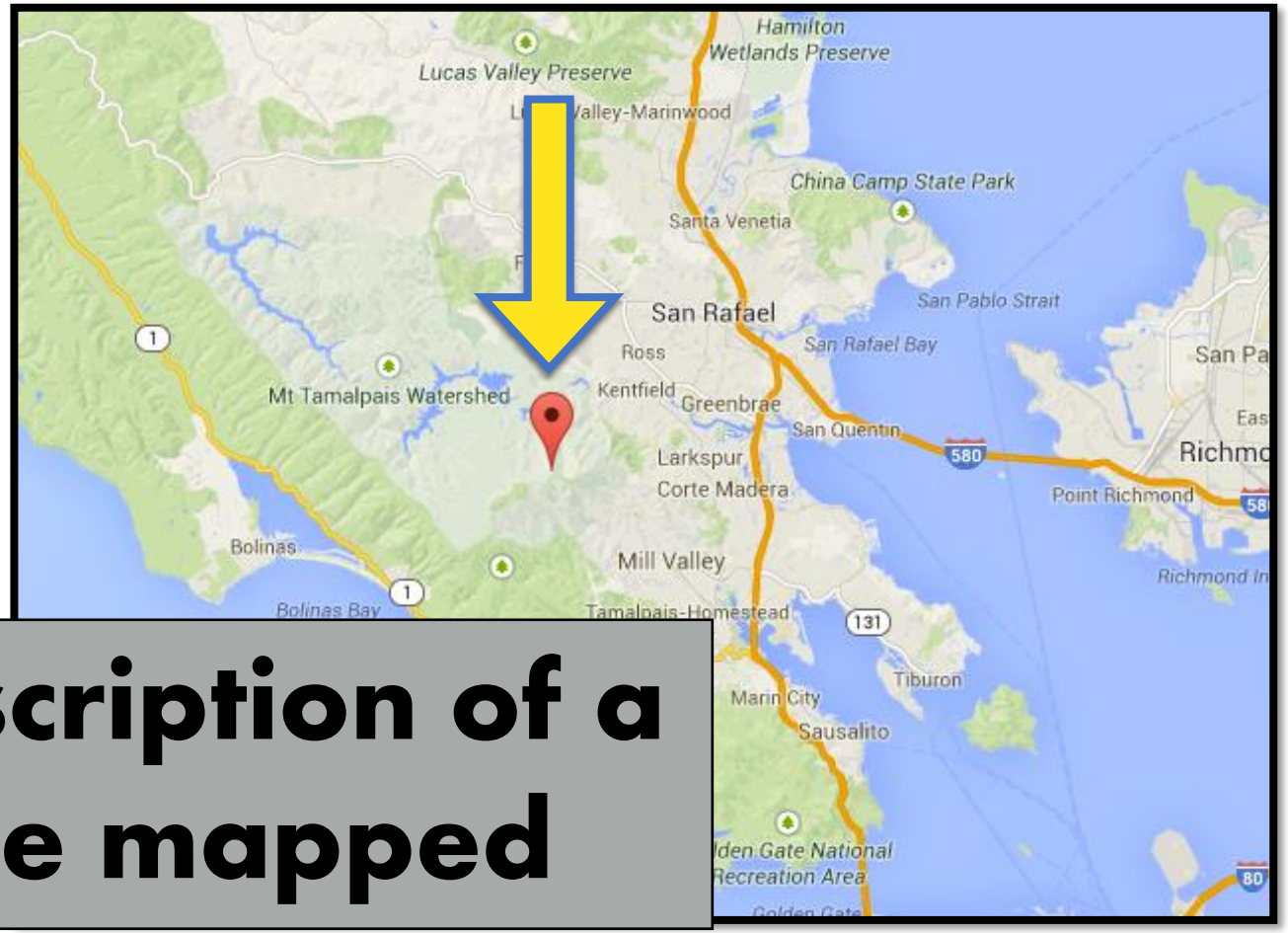


A numerical description of a place that can be mapped

Exploring Concepts

What is a georeference?

37.93, -122.59



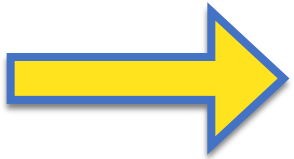
A numerical description of a place that can be mapped

Exploring Concepts

What is a georeference?

process of turning a textual description into a numerical description

Species: *Polypodium californicum*
State: CA
County: Marin
Locality: Marin Municipal Water District Lands: Mount Tamalpais. Middle Peak Rd. culvert RT-3



37.93, -122.59

Why georeference?

- Add more occurrences to a data set
- Correct geographic and specimen identification data = dependable occurrence record
- Provide uncertainty data, which allow points to be evaluated with regard to fitness for research applications and the resulting quality of output

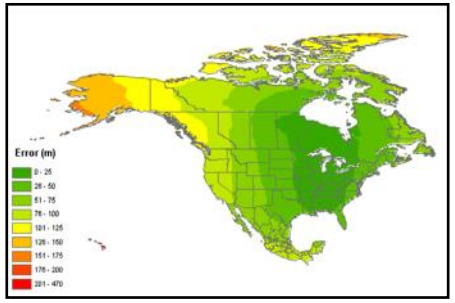
Exploring Concepts

Sources of uncertainty

- Coordinate uncertainty
- Map scale
- GPS accuracy
- Unknown datum
- Imprecision in direction measurements
- Imprecision in distance measurements
- Extent of locality

20° 30’ N 112° 36’ W

Scale	Uncertainty (ft)	Uncertainty (m)
1:1,200	3.3 ft	1.0 m
1:2,400	6.7 ft	2.0 m
1:4,800	13.3 ft	4.1 m
1:10,000	27.8 ft	8.5 m
1:12,000	33.3 ft	10.2 m
1:24,000	40.0 ft	12.2 m
1:25,000	41.8 ft	12.8 m
1:63,360	106 ft	32.2 m
1:100,000	167 ft	50.9 m
1:250,000	417 ft	127 m



Exploring Concepts

Sources of uncertainty

- **Coordinate uncertainty**
- Map scale
- GPS accuracy
- Unknown datum
- Imprecision in direction measur
- Imprecision in distance measure
- Extent of locality

Precision	0 degrees Latitude	30 degrees Latitude	60 degrees Latitude	85 degrees Latitude
1.0 degrees	156904 m	146962 m	124605 m	112109 m
0.1 degrees	15691 m	14697 m	12461 m	11211 m
0.01 degrees	1570 m	1470 m	1247 m	1122m
0.001 degrees	157 m	147 m	125 m	113 m
0.0001 degrees	16 m	15 m	13 m	12 m
0.00001 degrees	2 m	2 m	2 m	2 m
1.0 minutes	2615 m	2450 m	2077 m	1869 m
0.1 minutes	262 m	245 m	208 m	187 m
0.01 minutes	27 m	25 m	21 m	19 m
0.001 minutes	3 m	3 m	3 m	2 m
1.0 seconds	44 m	41 m	35 m	32 m
0.1 seconds	5 m	5 m	4 m	4 m
0.01 seconds	1 m	1 m	1 m	1 m

Exploring Concepts

Sources of uncertainty

- Coordinate uncertainty
- **Map scale**
- GPS accuracy
- Unknown/incorrect datum
- Imprecision in direction measurements
- Imprecision in distance measurements
- Extent of locality

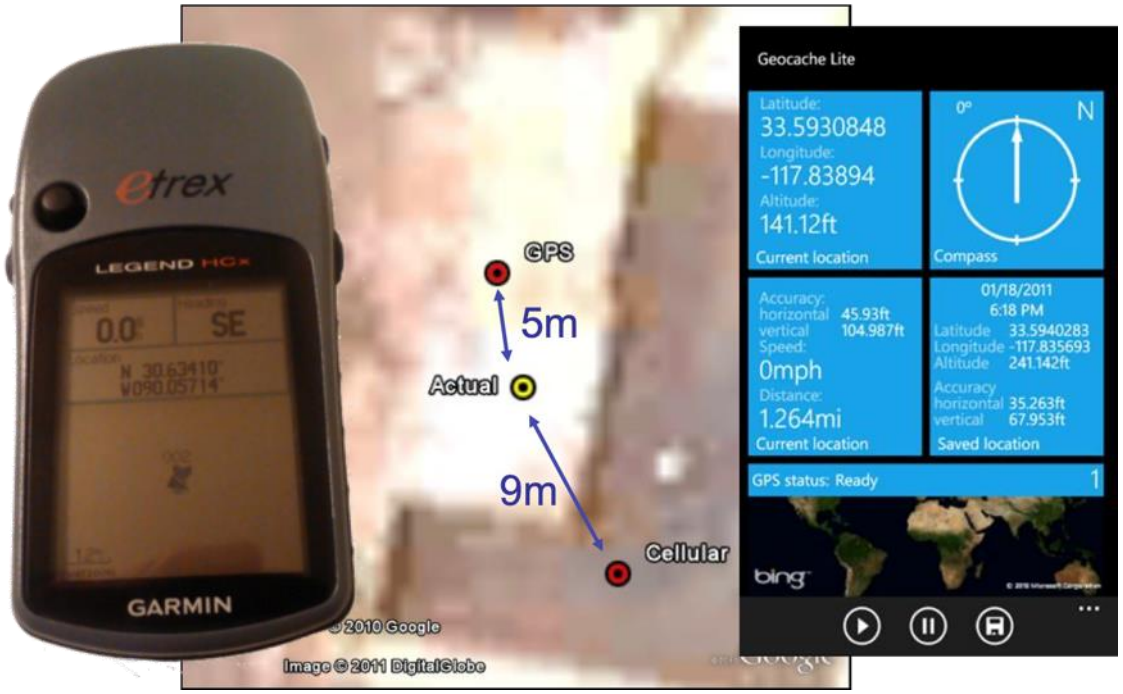
Scale	Uncertainty (ft)	Uncertainty (m)
1:1200	3.3 ft	1.0 m
1:2400	6.7 ft	2.0 m
1:4800	13.3 ft	4.1 m
1:10,000	27.8 ft	8.5 m
1:12,000	33.3 ft	10.2 m
1:24,000	40.0 ft	12.2 m
1:25,000	41.8 ft	12.8 m
1:63,360	106 ft	32.2 m
1:100,000	167 ft	50.9 m
1:250,000	417 ft	127 m

USGS Standard Map Accuracy

Exploring Concepts

Sources of uncertainty

- Coordinate uncertainty
- Map scale
- **GPS accuracy**
- Unknown/incorrect datum
- Imprecision in direction measurements
- Imprecision in distance measurements
- Extent of locality



Sources of uncertainty

- Coordinate uncertainty
- Map scale
- GPS accuracy
- **Unknown/incorrect datum**
- Imprecision in direction measurements
- Imprecision in distance measurements
- Extent of locality

Datum

What is a datum?

A geodetic datum is an abstract coordinate system with a reference surface (such as sea level) that serves to provide known locations to begin surveys and create maps. In this way, datums act similar to starting points when you give someone directions.

Geodetic datum

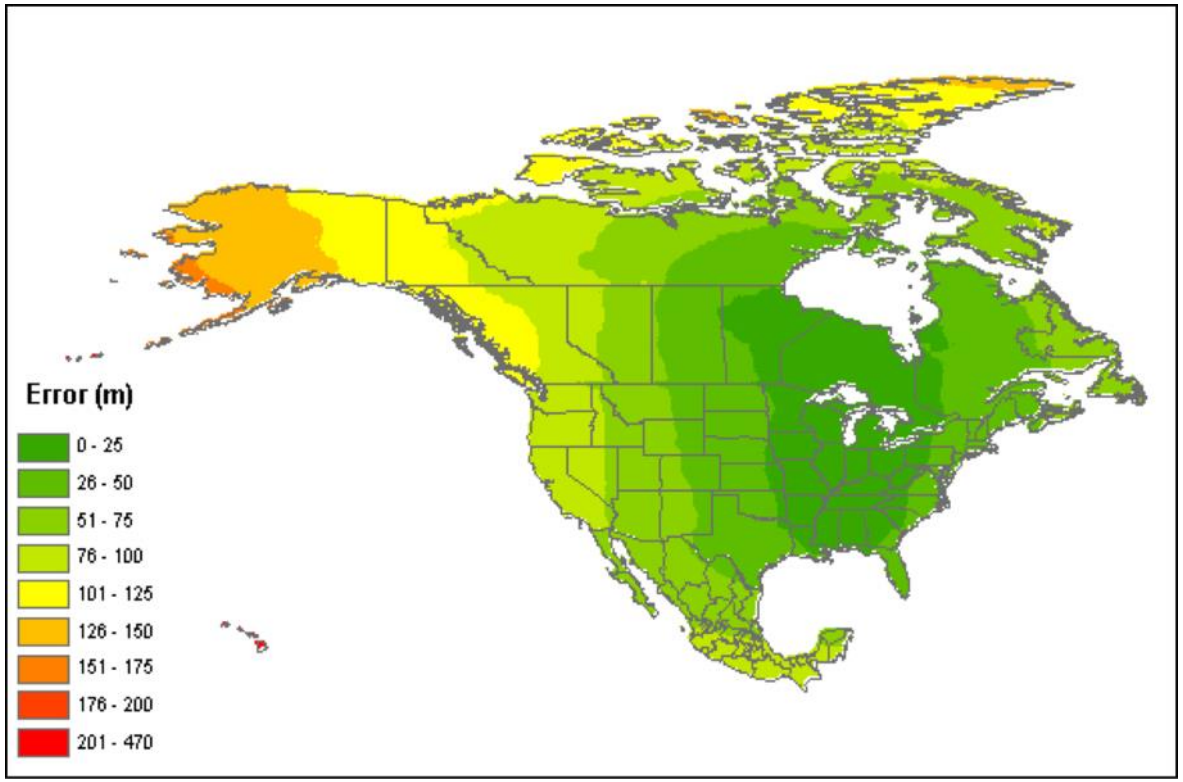
A model of the Earth used for geodetic calculations (related to Earth's figure, orientation, and gravity). A geodetic datum describes the size, shape, origin, and orientation of a coordinate system for mapping the surface of the Earth.

https://oceanservice.noaa.gov/facts/datum.html#:~:text=A%20geodetic%20datum%20is%20an,when%20you%20give%20someone%20directions.

Exploring Concepts

Sources of uncertainty

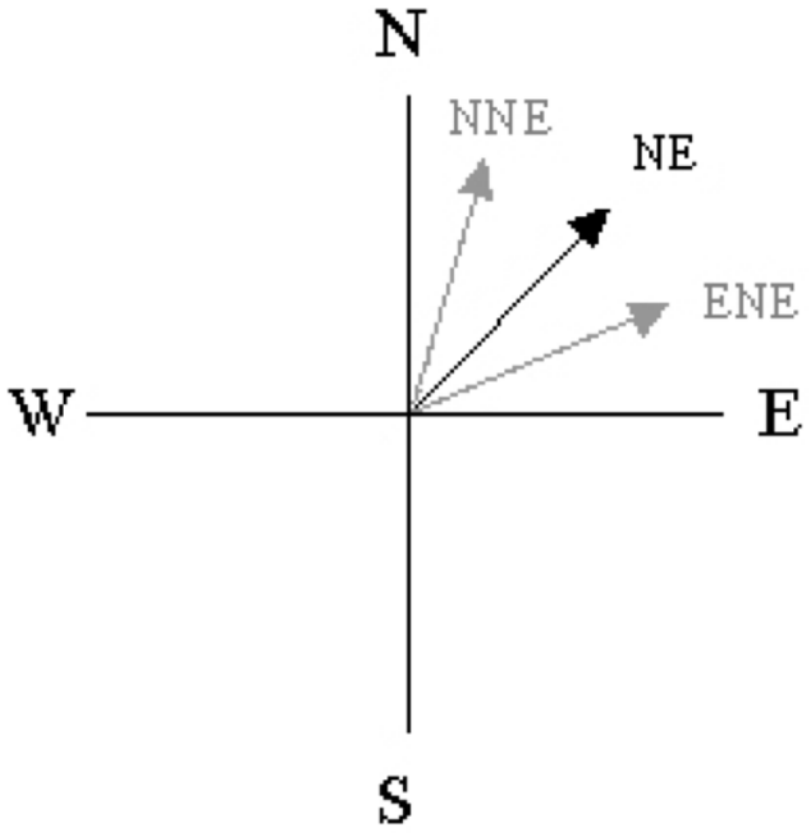
- Coordinate uncertainty
- Map scale
- GPS accuracy
- **Unknown/incorrect datum**
- Imprecision in direction measurements
- Imprecision in distance measurements
- Extent of locality



Error assuming NAD_{27} vs NAD_{83} or WGS_{84}

Sources of uncertainty

- Coordinate uncertainty
- Map scale
- GPS accuracy
- Unknown/incorrect datum
- **Imprecision in direction measurements**
- Imprecision in distance measurements
- Extent of locality

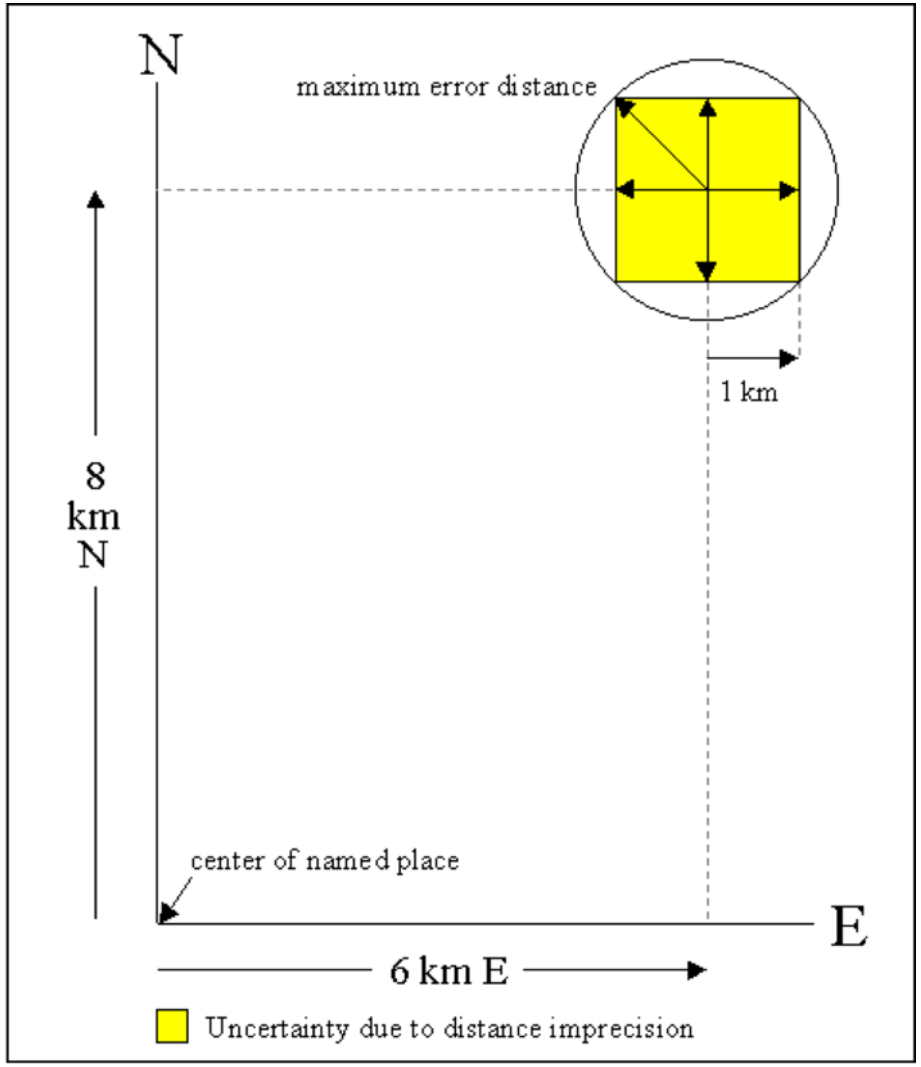


N > NE > ENE

Exploring Concepts

Sources of uncertainty

- Coordinate uncertainty
- Map scale
- GPS accuracy
- Unknown/incorrect datum
- Imprecision in direction measurements
- **Imprecision in distance measurements**
- Extent of locality



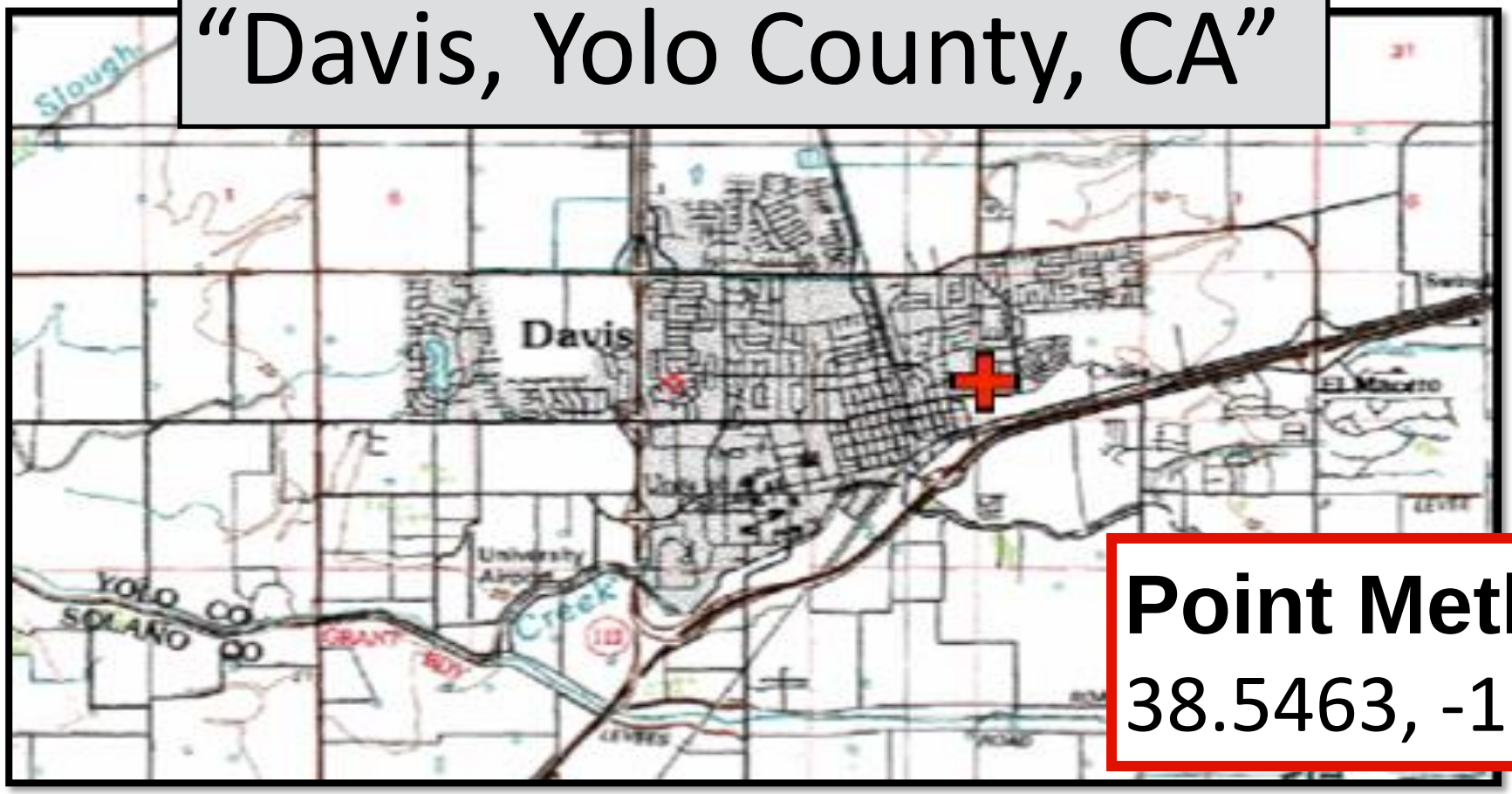
Sources of uncertainty

- Coordinate uncertainty
- Map scale
- GPS accuracy
- Unknown/incorrect datum
- Imprecision in direction measurements
- Imprecision in distance measurements
- **Extent of locality**

Exploring Concepts

Sources of uncertainty: Extent of locality

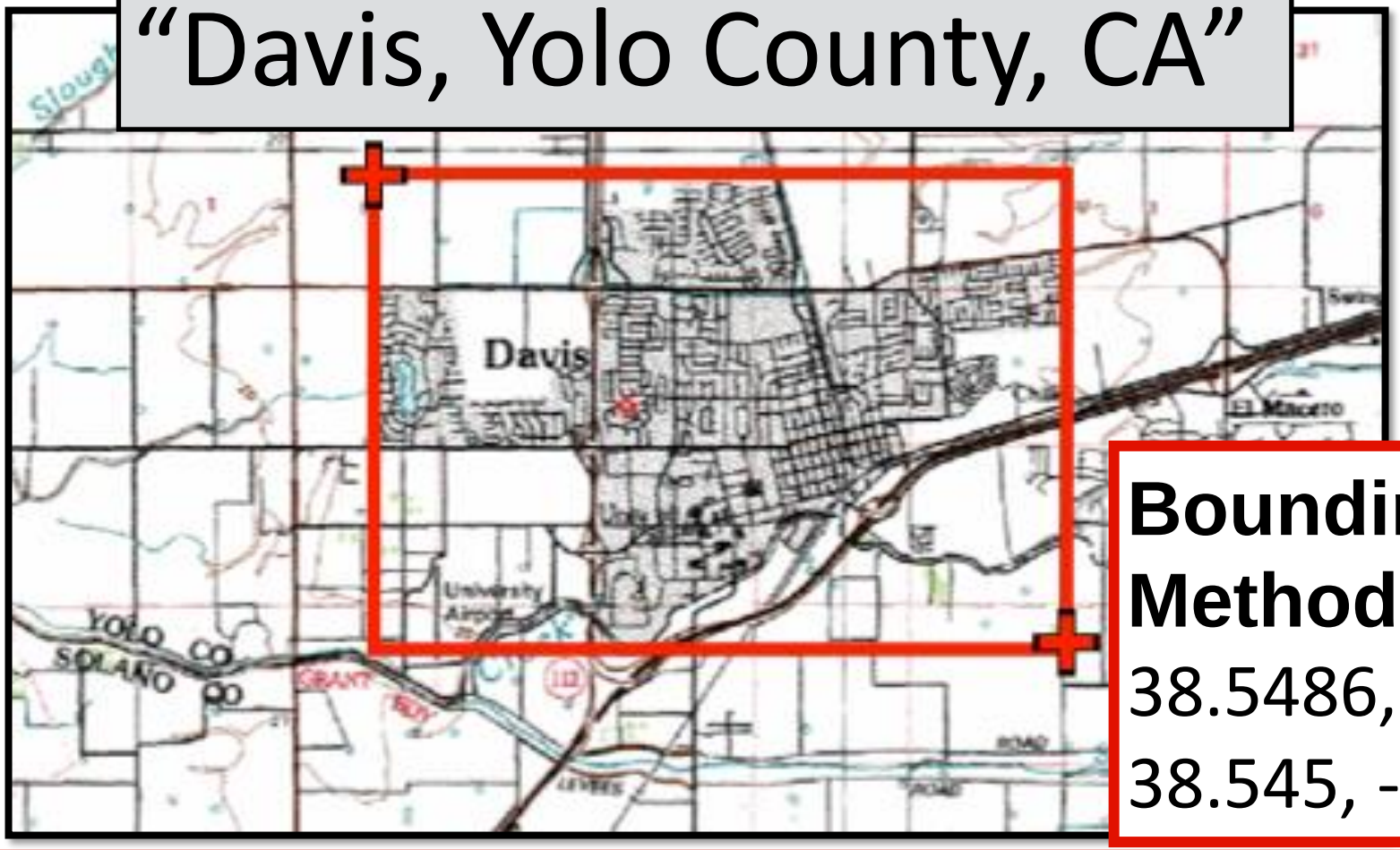
"Davis, Yolo County, CA"



Point Method
38.5463, -121.7425

Sources of uncertainty: Extent of locality

"Davis, Yolo County, CA"



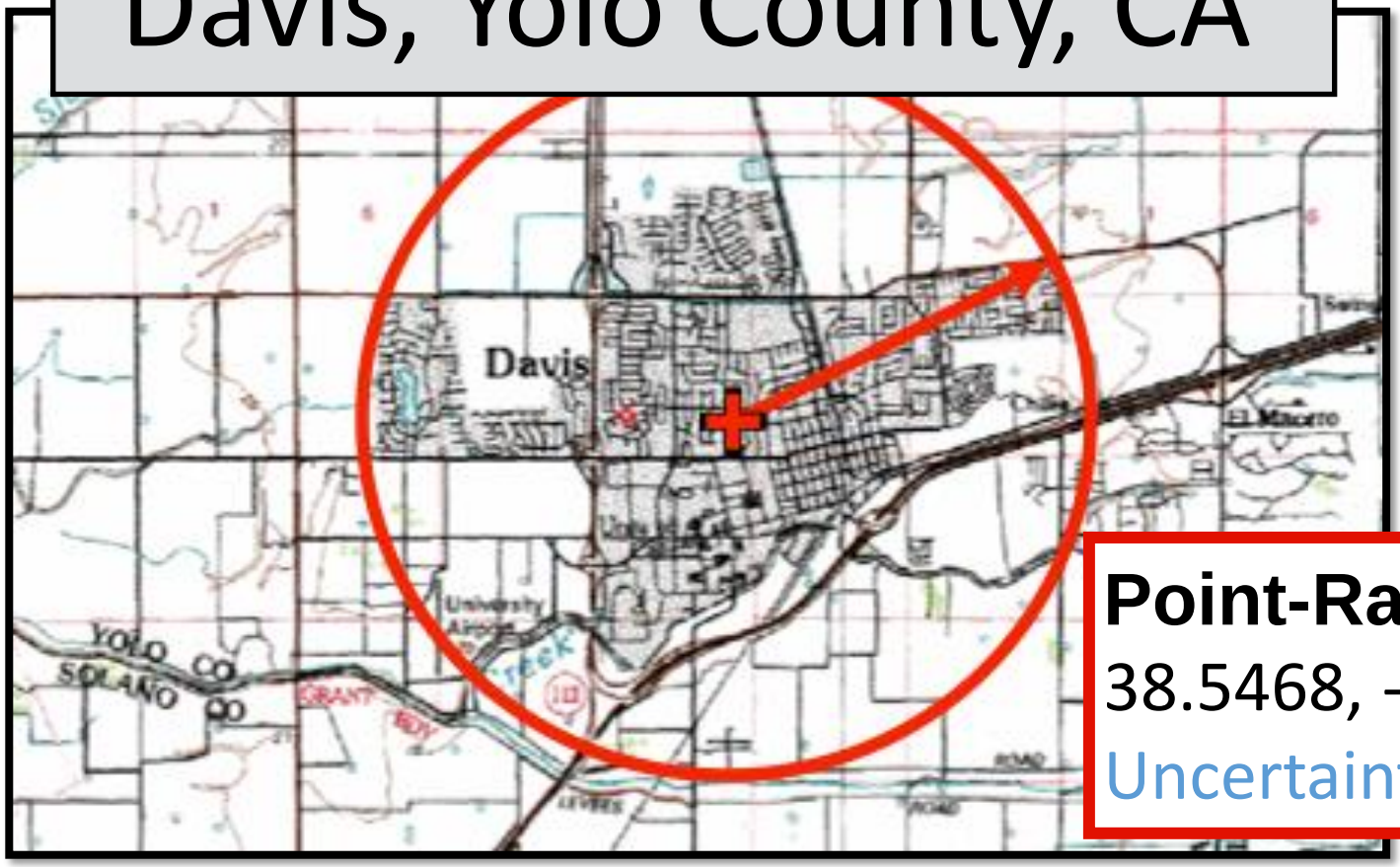
Bounding-Box Method

38.5486, -121.7542

38.545, -121.7394

Sources of uncertainty: Extent of locality

"Davis, Yolo County, CA"



Point-Radius Method

38.5468, -121.7469

Uncertainty radius: 8325 m

Exploring Concepts

What is the ideal georeference?

What is the ideal georeference?

A numerical description of a place that can be mapped

What is the ideal georeference?

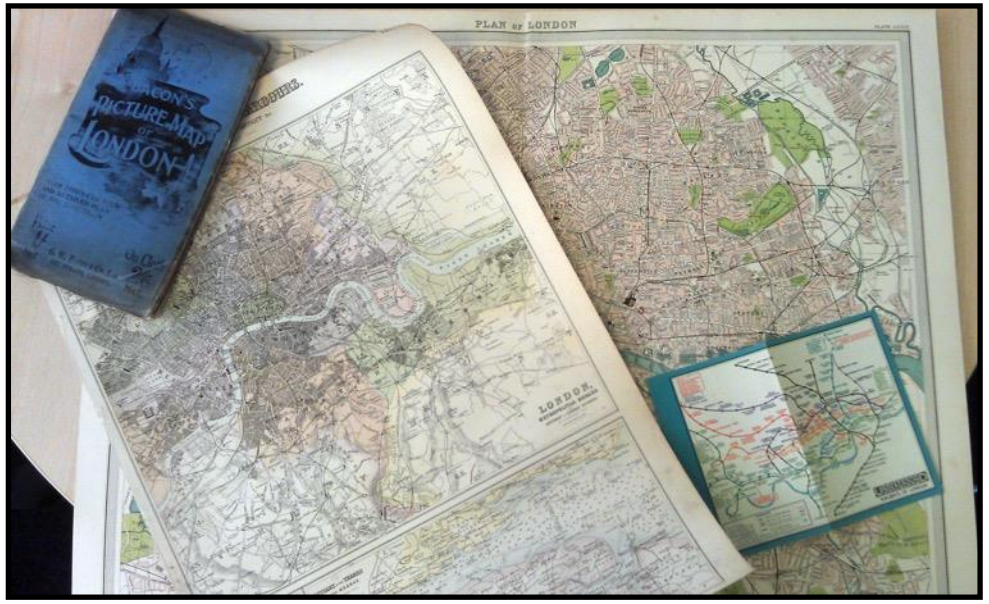
A numerical description of a place that can be mapped

and that describes the spatial extent of a locality and its associated uncertainties

How to do it?

Paper Maps

- Time-consuming
- Good quality paper maps may be hard to find



Internet Resources

How to do it?

Google Maps

- Search and directions
  - Free text search
  - Directions for traveling by car, bike, public transportation, or foot
  - Data compiled from different sources
- Maps
  - Views: map, satellite, terrain, Google Street View
  - Sources of maps indicated at bottom

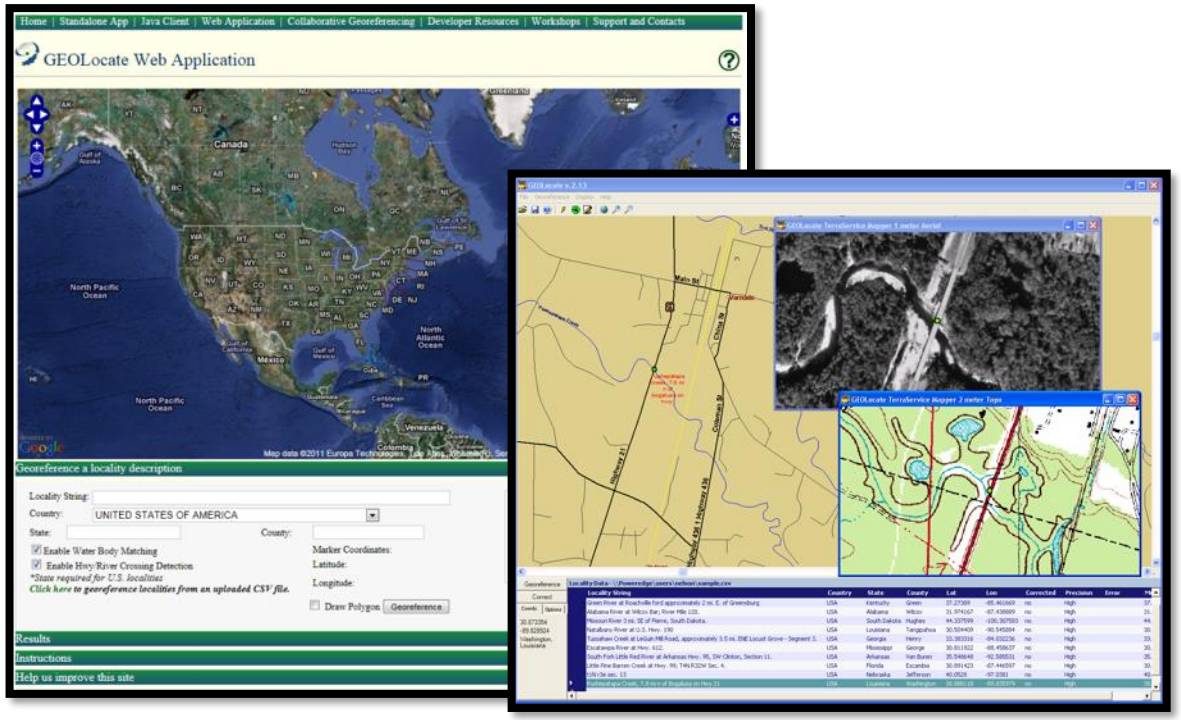
Exploring Concepts

How to do it? GEOLocate

- Software services for georeferencing of natural history collection data
- Automated georeferencing
- Verification and correction
- Multi-lingual
- Interoperability
- Training
- Uncertainty determination
- Batch processing
- Collaborative georeferencing
- Geographic visualization
- Kml export

GEOLocate

https://www.geo-locate.org/



Exploring Concepts

How to do it?

Falling Rain

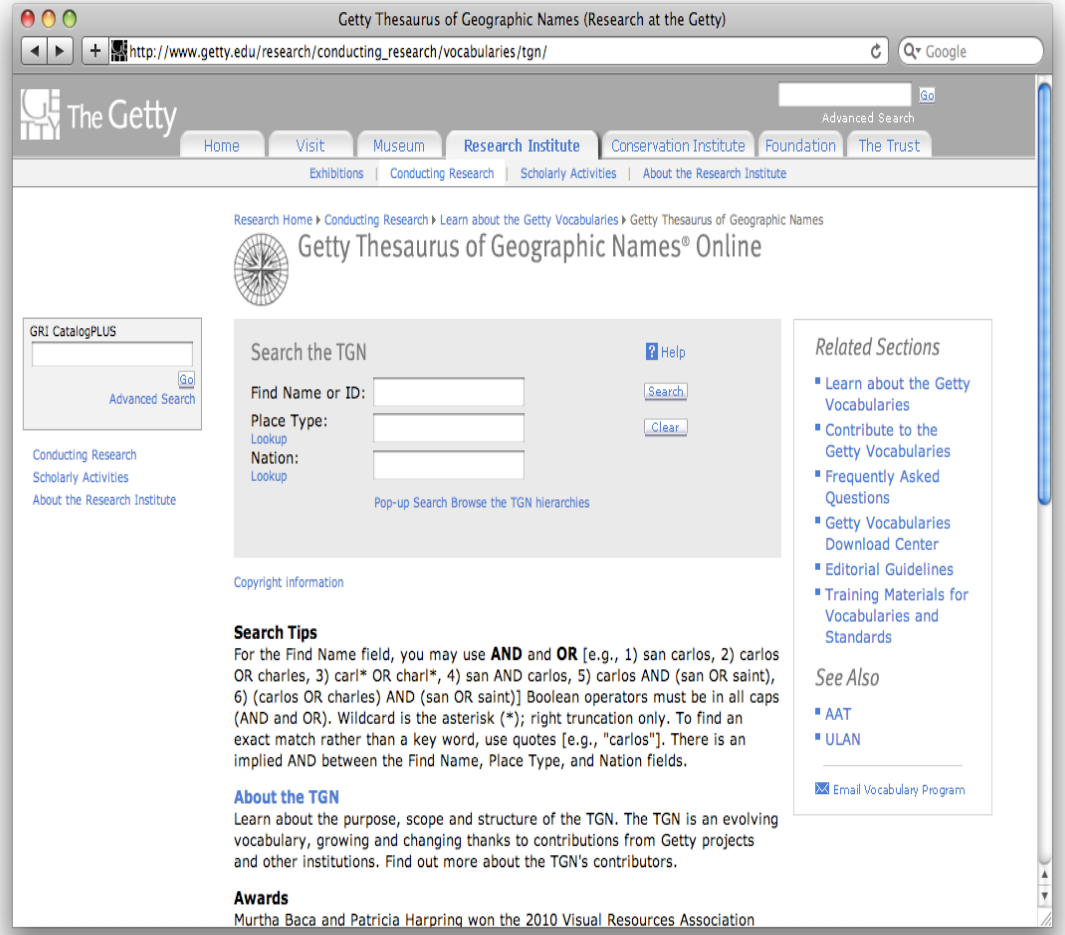
- Worldwide gazetteer for cities and towns
- Great for hard-to-find localities, especially outside US
  - No search
  - Provides hierarchy, alternative names, topo maps, altitude, weather information, and location of nearby towns
  - Example: Qaryeh-ye Gol'alam, Velayat-e Lowgar, AF

www.fallingrain.com

How to do it?

Getty Thesaurus of Geographic Names

- Worldwide gazetteer by The Getty
- Useful for alternative and old names
  - Feature types
  - Geographical hierarchy
  - Degrees-minutes, not coordinates
    - Use recent name and search in Google Maps
  - Example: New Amsterdam, US



http://bit.ly/Getty-TGN

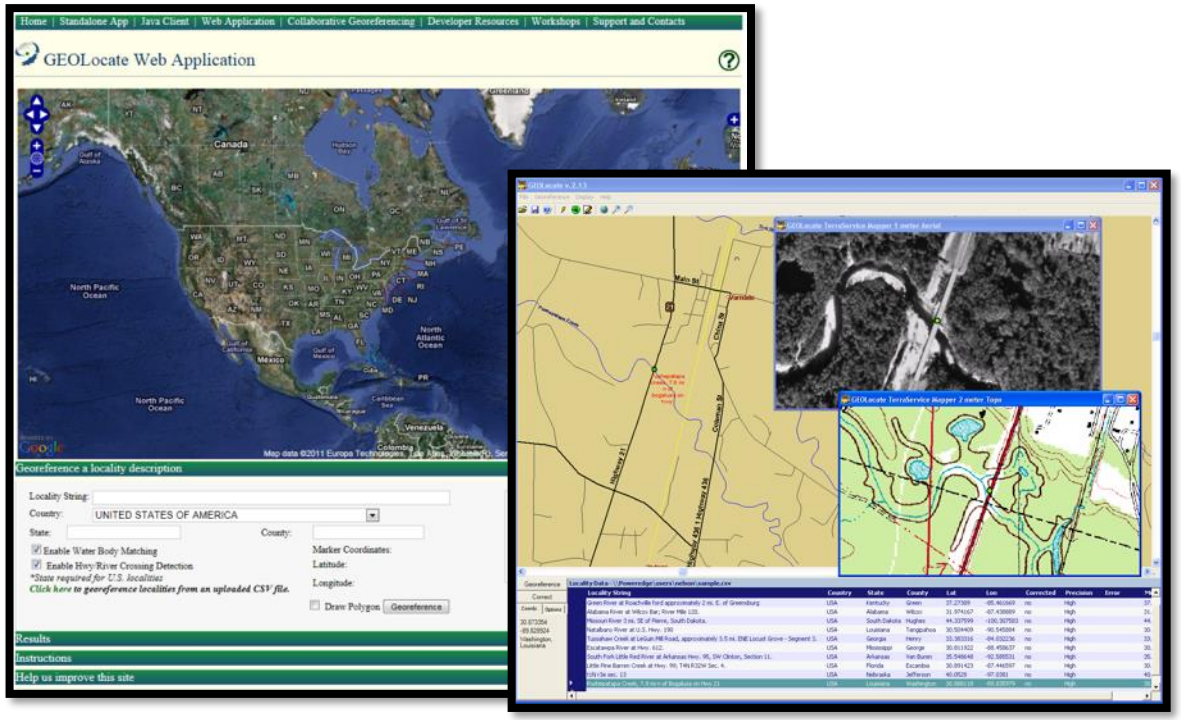
Exploring Concepts

How to do it? GEOLocate

- Software services for georeferencing of natural history collection data
- Automated georeferencing
- Verification and correction
- Multi-lingual
- Interoperability
- Training
- Uncertainty determination
- Batch processing
- Collaborative georeferencing
- Geographic visualization
- Kml export

GEOLocate

https://www.geo-locate.org/



Demonstration: How to use GEOLocate

- ✓ Access GEOLocate web interface and learn how to navigate it

Demonstration: How to use GEOLocate

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

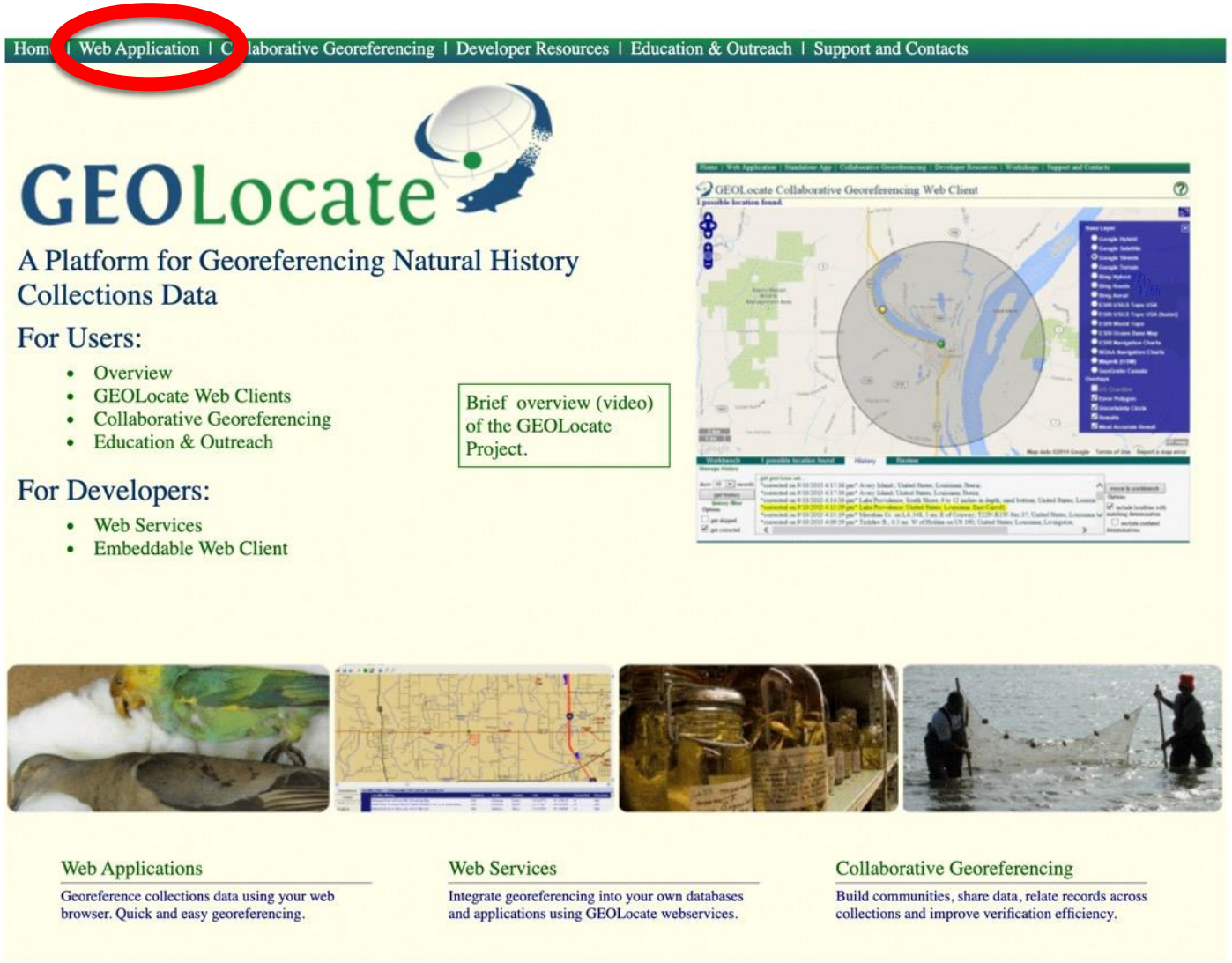
latitude: **30.88797**
longitude: **-89.83601**

uncertainty radius: **48m**

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

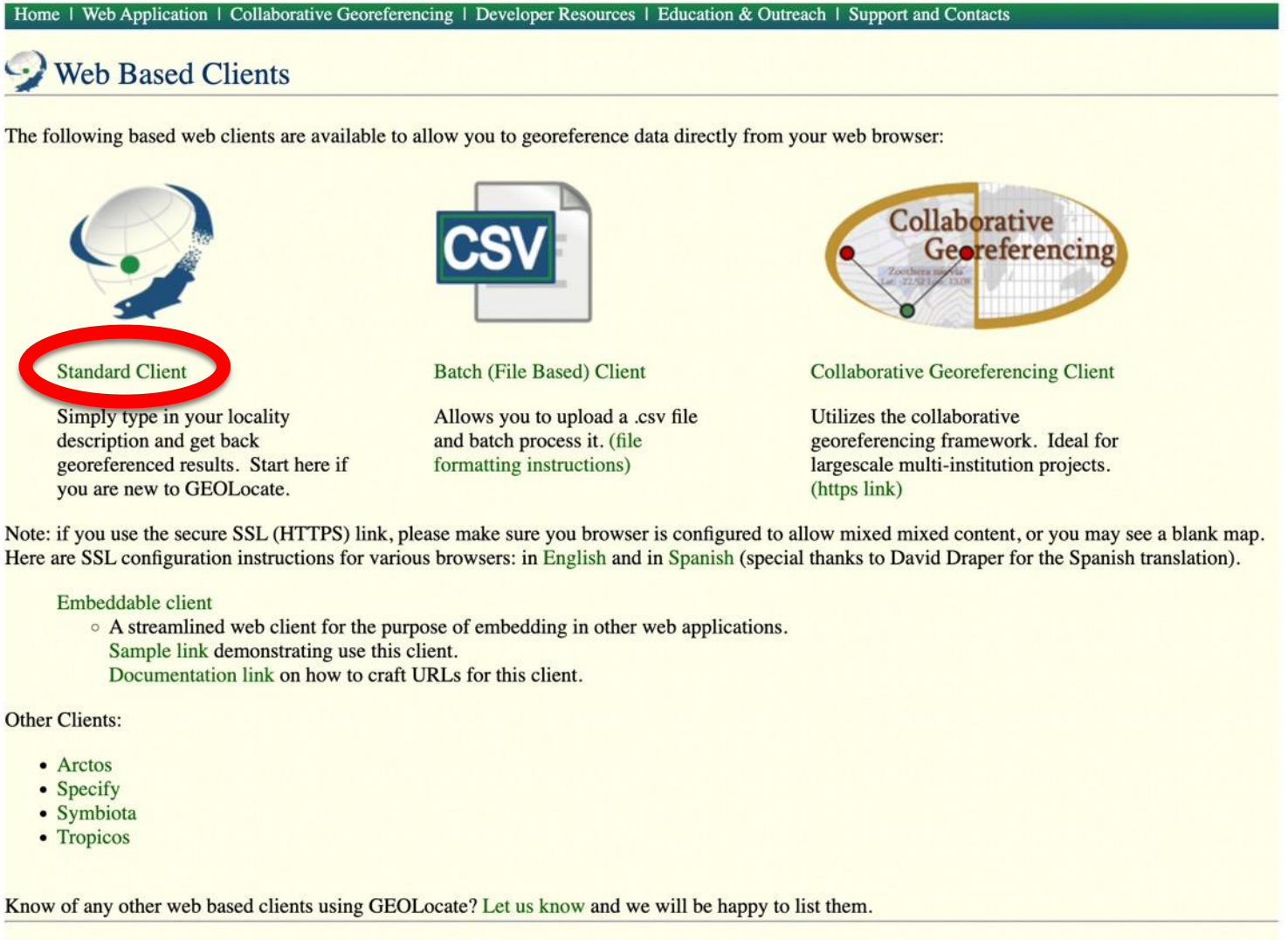
Demonstration: How to use GEOLocate

GEOLocate Workflow



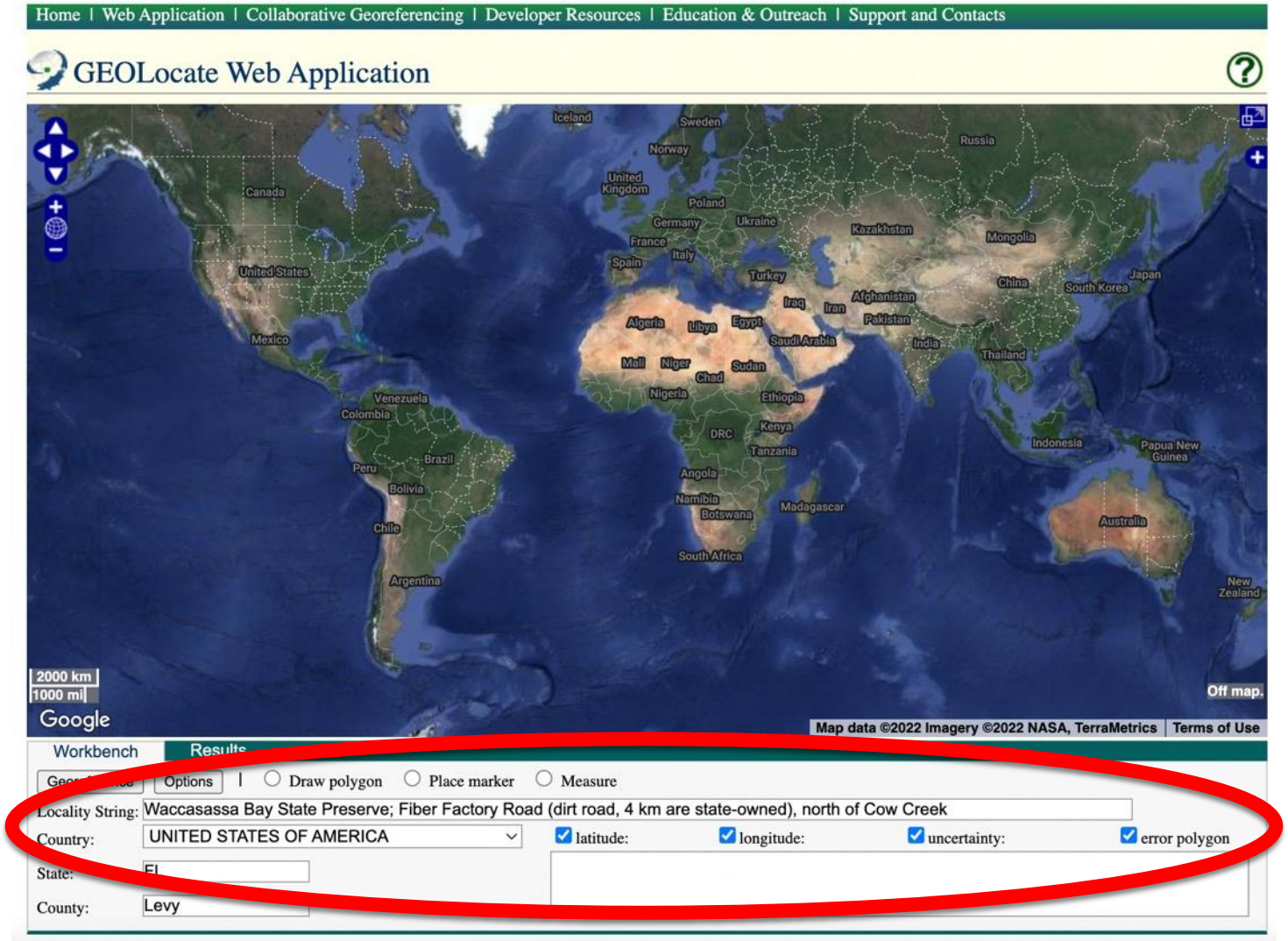
Demonstration: How to use GEOLocate

GEOLocate Workflow



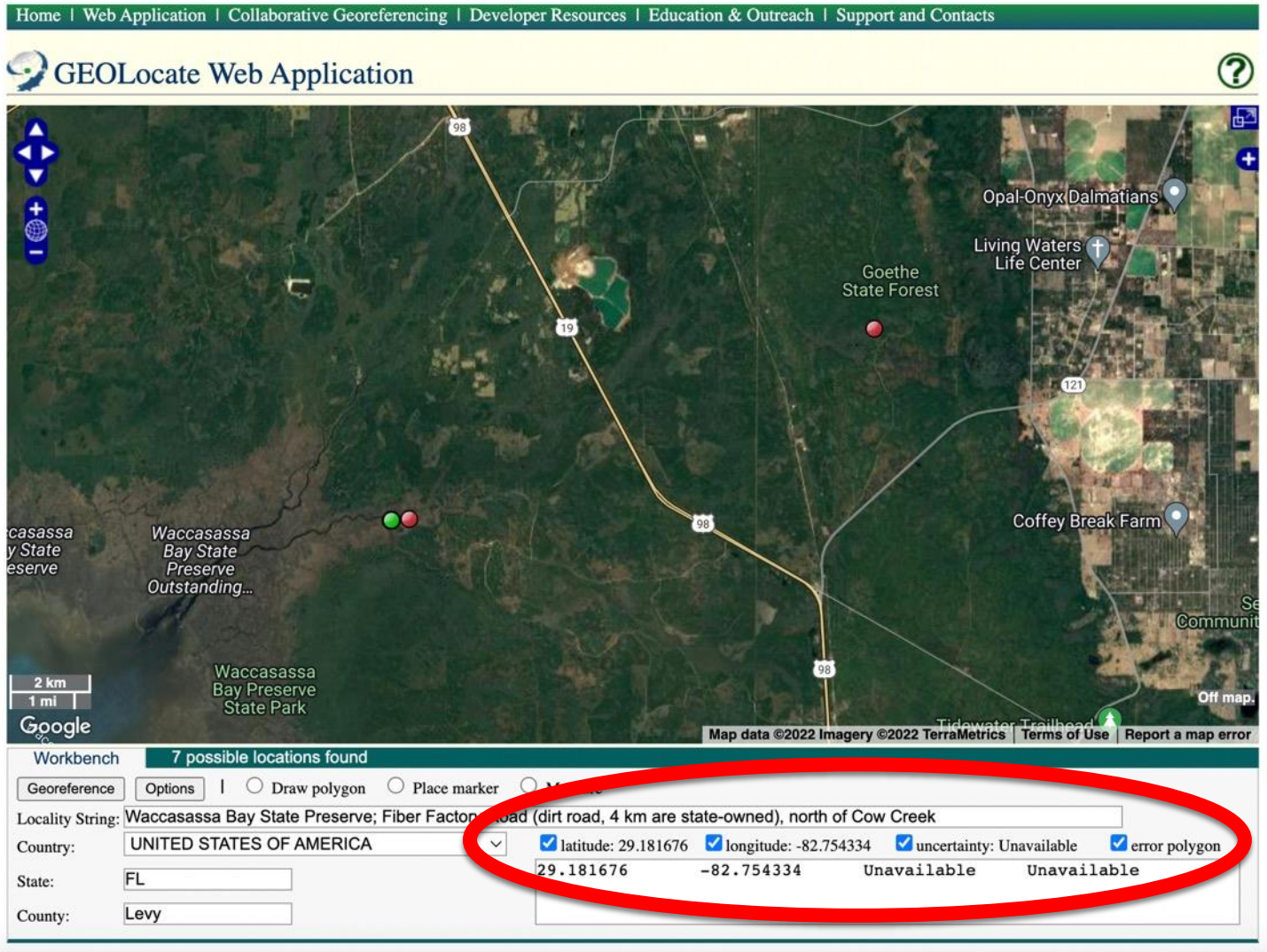
Demonstration: How to use GEOLocate

GEOLocate Workflow



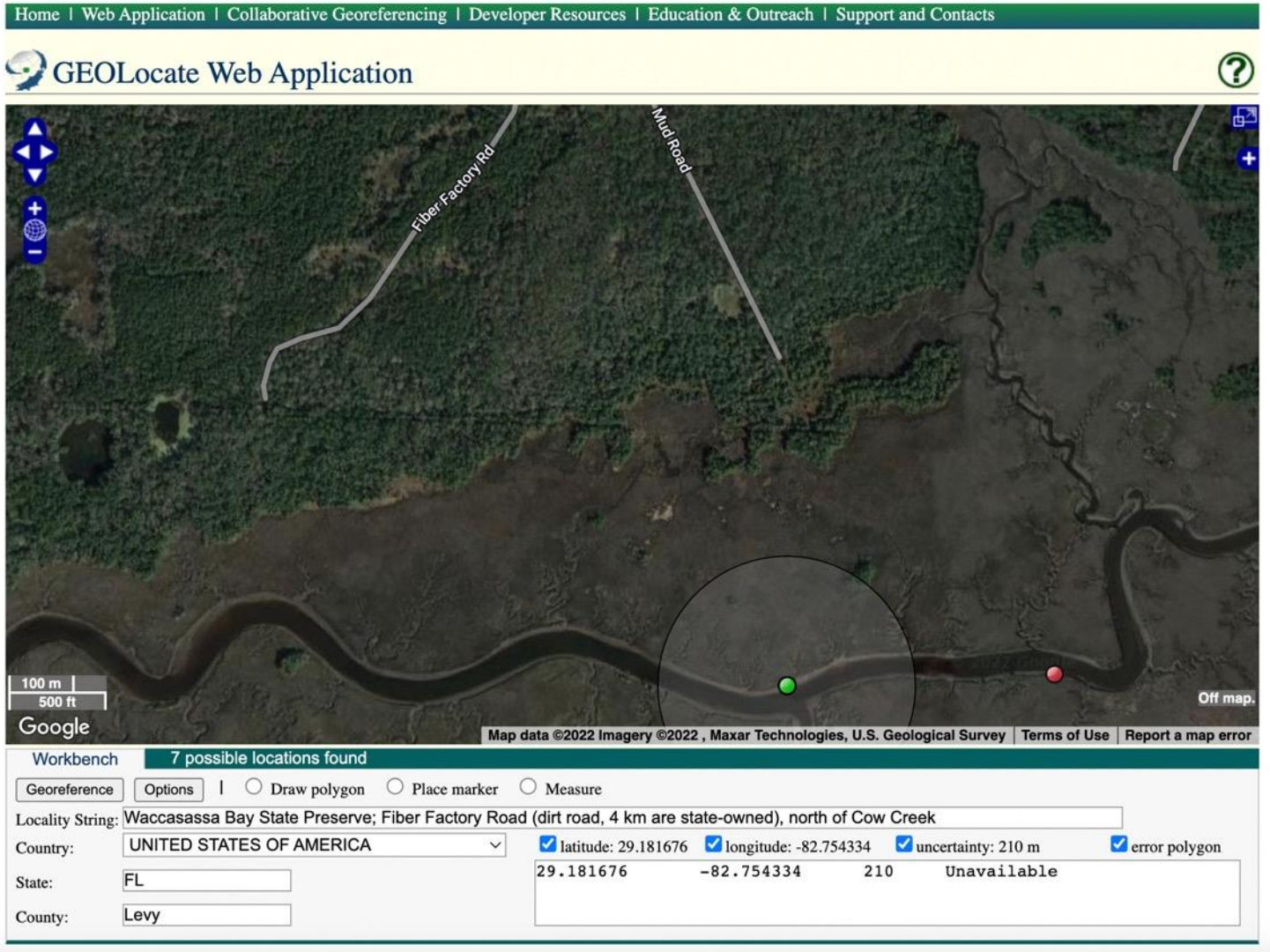
Demonstration: How to use GEOLocate

GEOLocate Workflow



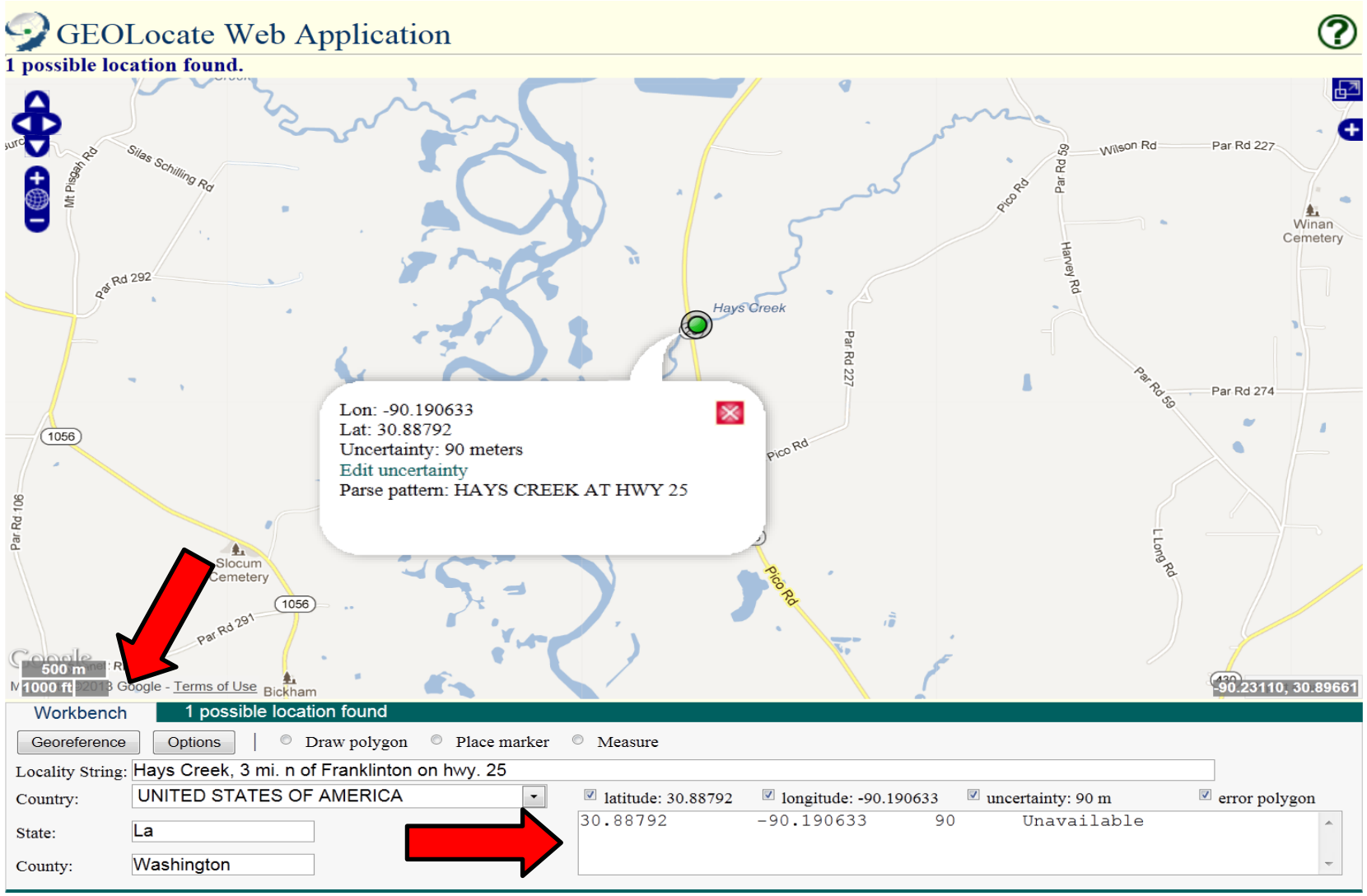
Demonstration: How to use GEOLocate

GEOLocate Workflow



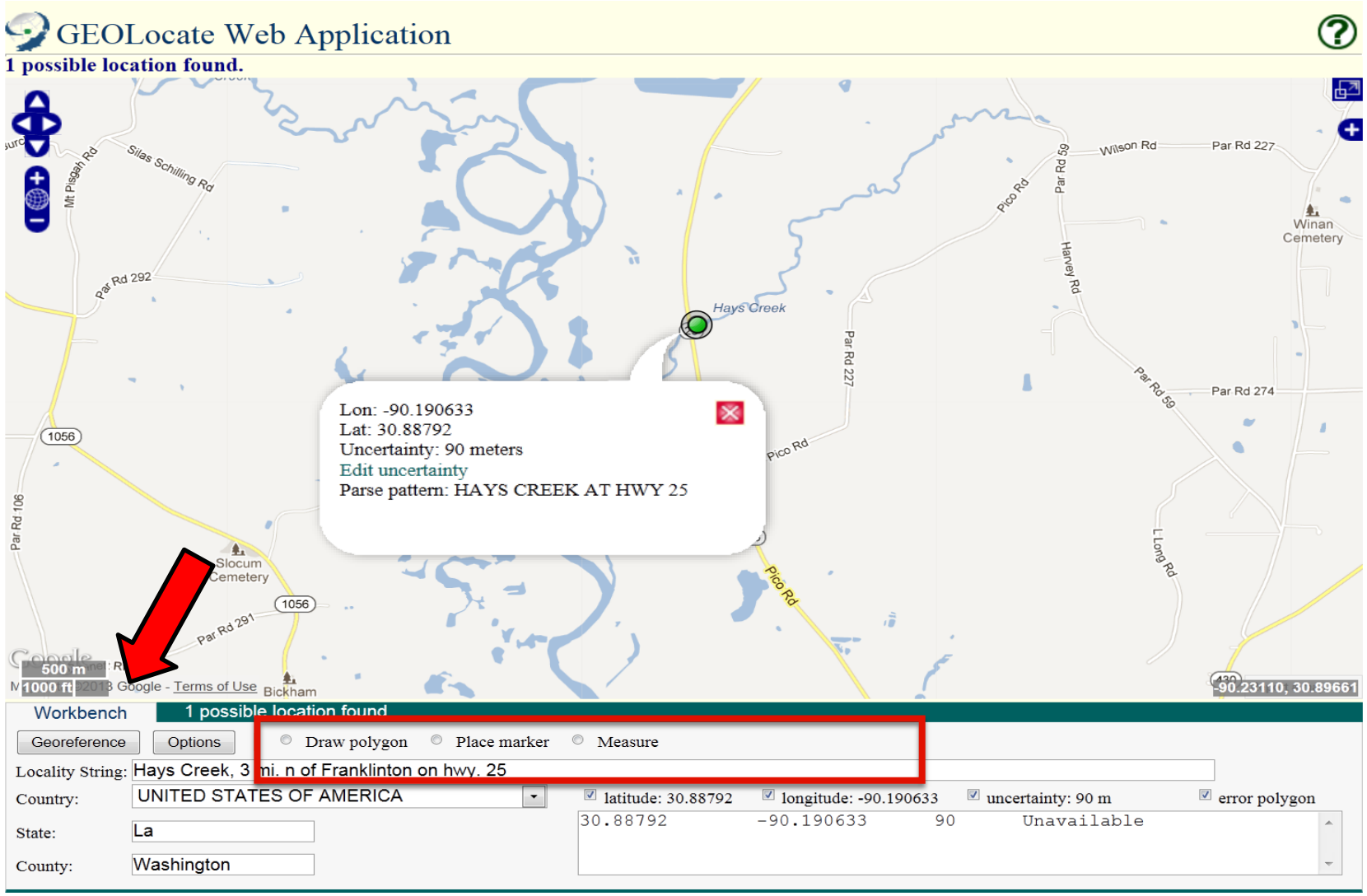
Demonstration: How to use GEOLocate

GEOLocate Workflow



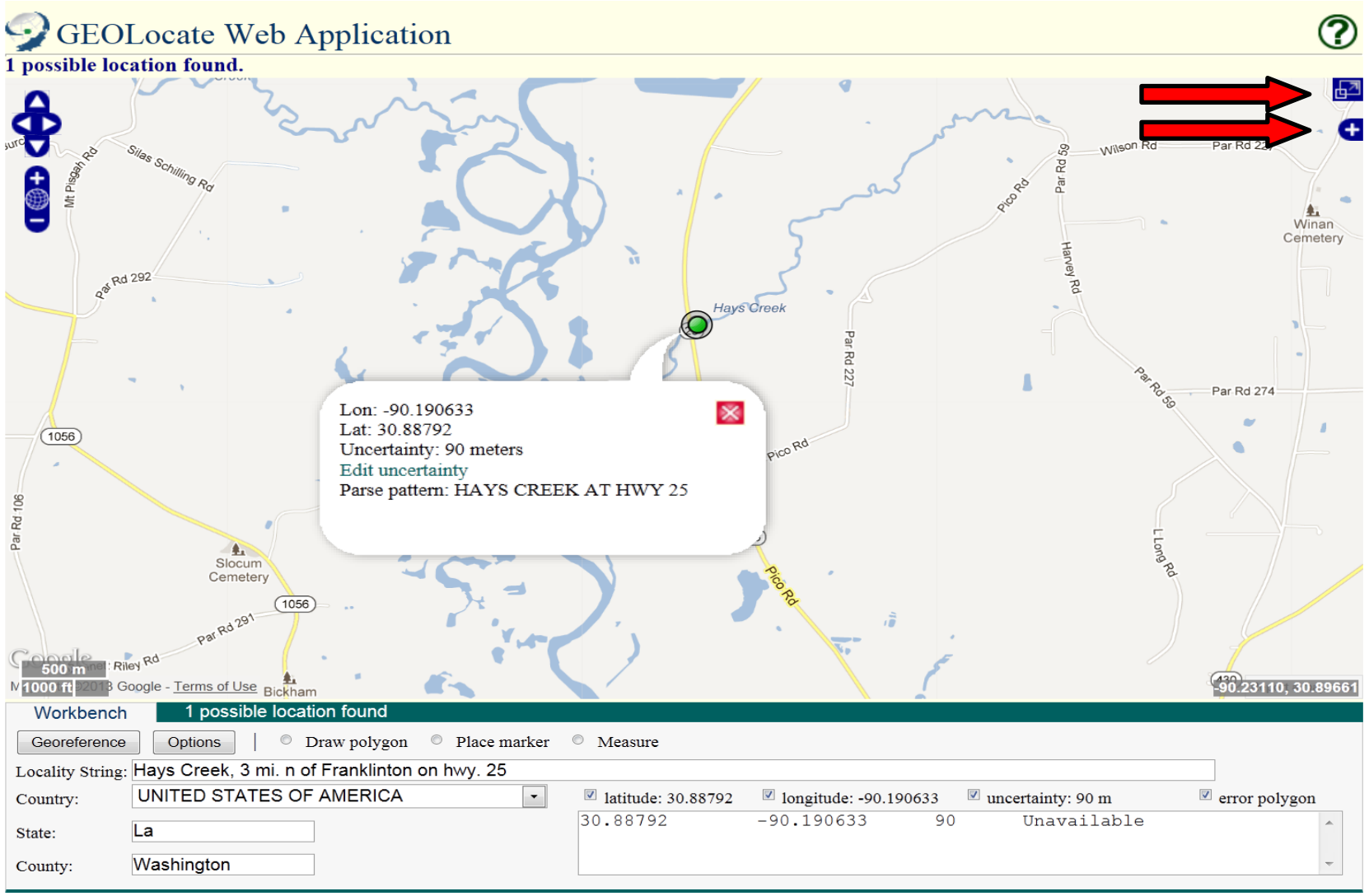
Demonstration: How to use GEOLocate

GEOLocate Workflow



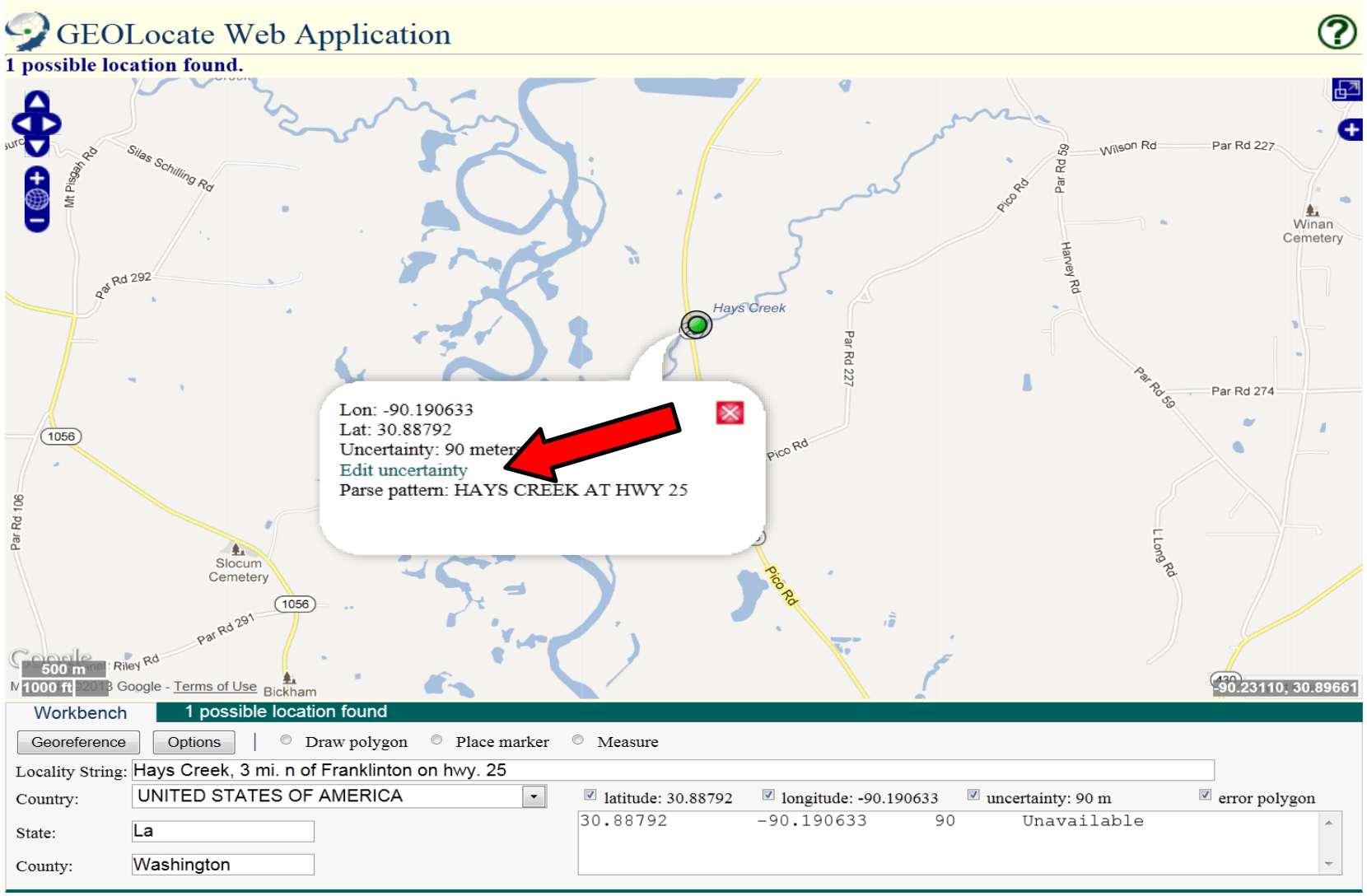
Demonstration: How to use GEOLocate

GEOLocate Workflow



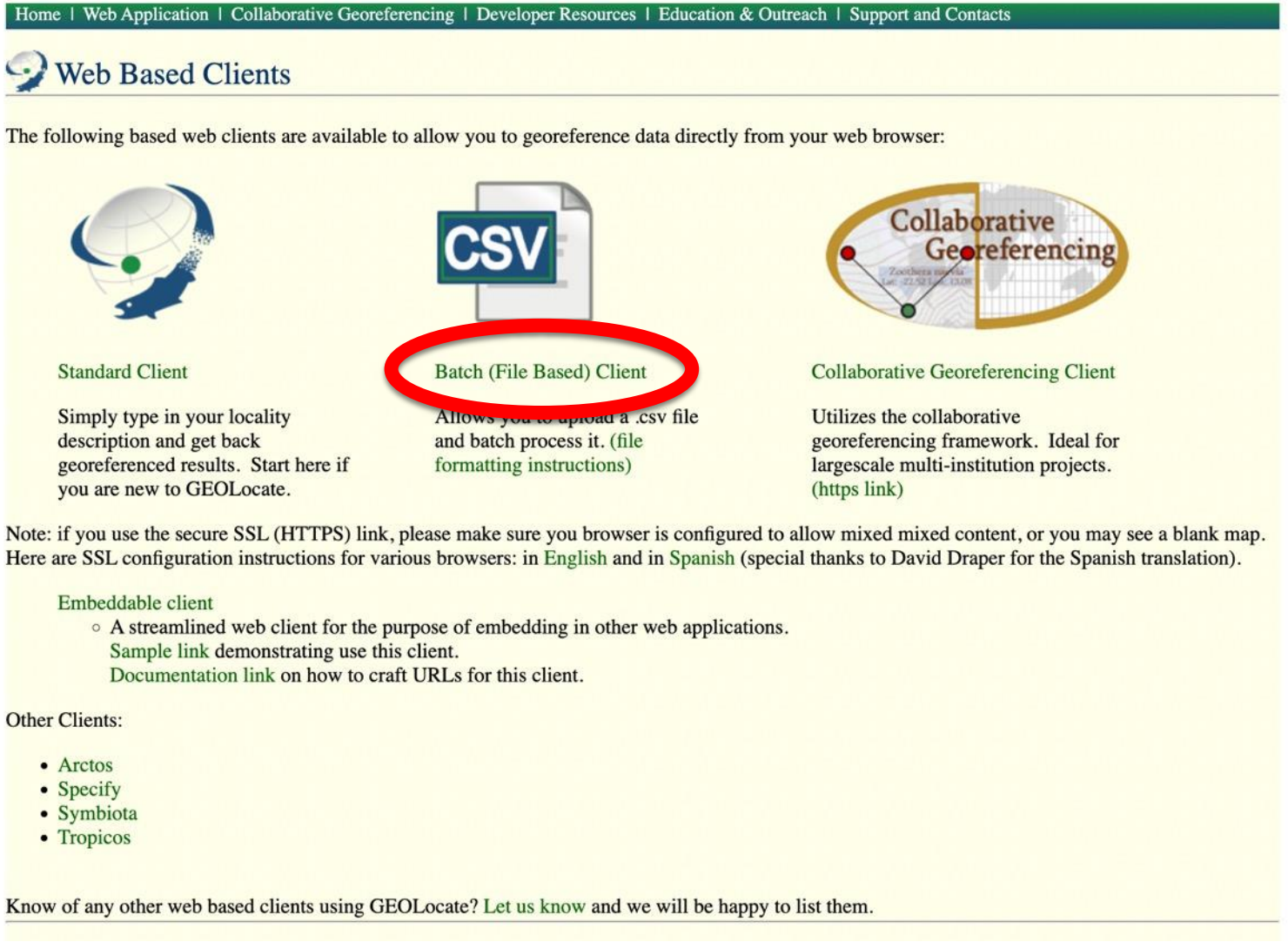
Demonstration: How to use GEOLocate

GEOLocate Workflow



Demonstration: How to use GEOLocate

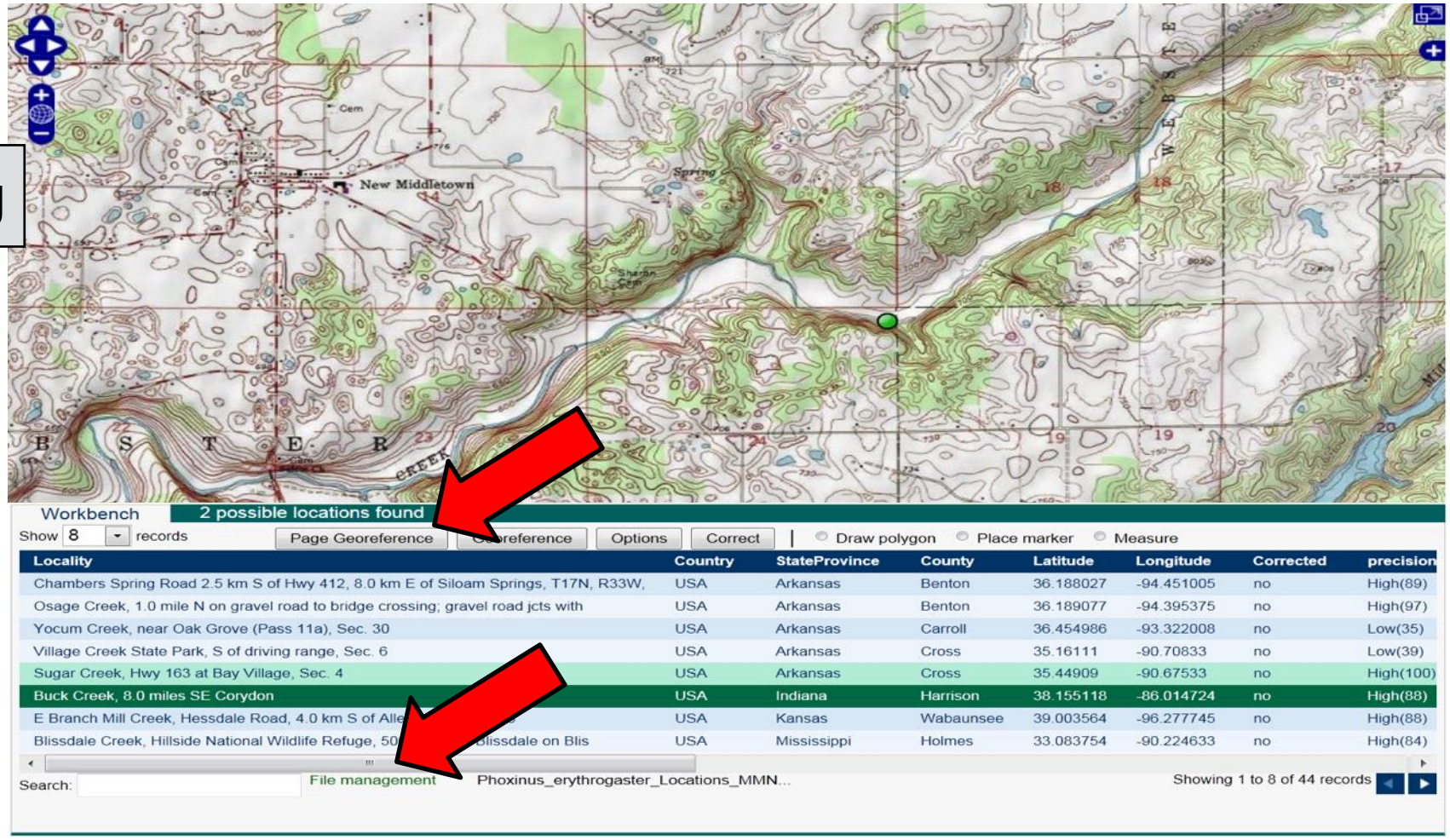
GEOLocate Workflow



Demonstration: How to use GEOLocate

GEOLocate Workflow

Batch Georeferencing



Time to Exercise!

- ✓ Instruction sheet
- ✓ Data file
- ✓ Go to geo-locate.org

Time to Exercise!

✓ Instruction sheet

Georeferencing Exercises

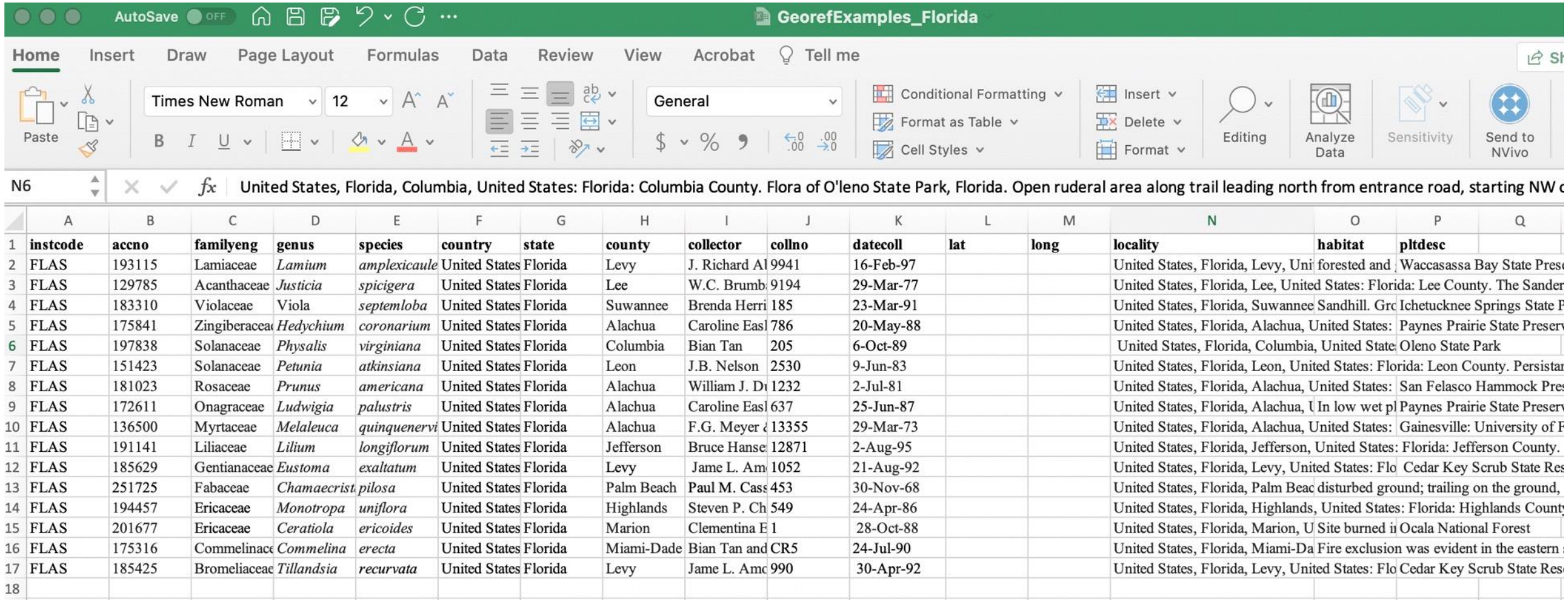
Resources:

GEOLocate – Web application: http://www.geo-locate.org/default.html
Google Maps: https://www.google.com/maps
Falling Rain: http://www.fallingrain.com
Getty Thesaurus of Geographic Names (TGN): http://bit.ly/Getty-TGN

1. Use the standard GEOLocate client to identify the first three localities in the GeorefExamples_Florida.xls file.
   a. Enter the locality string, country, state, and county information from the Excel sheet.
   b. Click "Georeference."
   c. Inspect the "Possible Locations" by clicking on the "XX possible locations found" where XX is the number of locations GEOLocate identified.
   d. Use an alternative resource to double check the locality. Try Google Maps.
   e. Adjust the point location as you see fit. The green point is the active one.
   f. Click the green point on the map, then click "Edit uncertainty". Adjust the uncertainty radius by moving the grey arrow.
   g. Return to the "Workbench" and record the latitude, longitude, and uncertainty.
   h. If the uncertainty is >1000, it then discards the points.

Time to Exercise!

✓ Data file



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	**instcode**	**accno**	**familyeng**	**genus**	**species**	**country**	**state**	**county**	**collector**	**collno**	**datecoll**	**lat**	**long**	**locality**	**habitat**	**pltdesc**	
2	FLAS	193115	Lamiaceae	*Lamium*	*amplexicaule*	United States	Florida	Levy	J. Richard Al	9941	16-Feb-97			United States, Florida, Levy, Uni	forested and	Waccasassa Bay State Pres	
3	FLAS	129785	Acanthaceae	*Justicia*	*spicigera*	United States	Florida	Lee	W.C. Brumb	9194	29-Mar-77			United States, Florida, Lee, United States: Florida: Lee County. The Sander			
4	FLAS	183310	Violaceae	Viola	*septemloba*	United States	Florida	Suwannee	Brenda Herri	185	23-Mar-91			United States, Florida, Suwannee	Sandhill. Gro	Ichetucknee Springs State P	
5	FLAS	175841	Zingiberaceae	*Hedychium*	*coronarium*	United States	Florida	Alachua	Caroline Easl	786	20-May-88			United States, Florida, Alachua, United States:		Paynes Prairie State Preserv	
6	FLAS	197838	Solanaceae	*Physalis*	*virginiana*	United States	Florida	Columbia	Bian Tan	205	6-Oct-89			United States, Florida, Columbia, United State		Oleno State Park	
7	FLAS	151423	Solanaceae	*Petunia*	*atkinsiana*	United States	Florida	Leon	J.B. Nelson	2530	9-Jun-83			United States, Florida, Leon, United States: Florida: Leon County. Persistar			
8	FLAS	181023	Rosaceae	*Prunus*	*americana*	United States	Florida	Alachua	William J. Du	1232	2-Jul-81			United States, Florida, Alachua, United States:		San Felasco Hammock Pres	
9	FLAS	172611	Onagraceae	*Ludwigia*	*palustris*	United States	Florida	Alachua	Caroline Easl	637	25-Jun-87			United States, Florida, Alachua, U	In low wet pl	Paynes Prairie State Preserv	
10	FLAS	136500	Myrtaceae	*Melaleuca*	*quinquenervi*	United States	Florida	Alachua	F.G. Meyer &	13355	29-Mar-73			United States, Florida, Alachua, United States:		Gainesville: University of F	
11	FLAS	191141	Liliaceae	*Lilium*	*longiflorum*	United States	Florida	Jefferson	Bruce Hanse	12871	2-Aug-95			United States, Florida, Jefferson, United States: Florida: Jefferson County.			
12	FLAS	185629	Gentianaceae	*Eustoma*	*exaltatum*	United States	Florida	Levy	Jame L. Am	1052	21-Aug-92			United States, Florida, Levy, United States: Flo		Cedar Key Scrub State Res	
13	FLAS	251725	Fabaceae	*Chamaecrist*	*pilosa*	United States	Florida	Palm Beach	Paul M. Cass	453	30-Nov-68			United States, Florida, Palm Beac	disturbed ground; trailing on the ground,		
14	FLAS	194457	Ericaceae	*Monotropa*	*uniflora*	United States	Florida	Highlands	Steven P. Ch	549	24-Apr-86			United States, Florida, Highlands, United States: Florida: Highlands County			
15	FLAS	201677	Ericaceae	*Ceratiola*	*ericoides*	United States	Florida	Marion	Clementina E	1	28-Oct-88			United States, Florida, Marion, U	Site burned i	Ocala National Forest	
16	FLAS	175316	Commelinace	*Commelina*	*erecta*	United States	Florida	Miami-Dade	Bian Tan and	CR5	24-Jul-90			United States, Florida, Miami-Da	Fire exclusion was evident in the eastern		
17	FLAS	185425	Bromeliaceae	*Tillandsia*	*recurvata*	United States	Florida	Levy	Jame L. Amc	990	30-Apr-92			United States, Florida, Levy, United States: Flo		Cedar Key Scrub State Res	
18																	

Time to Exercise!

Georeferencing Exercises

1. Use the standard GEOLocate client to identify the first three localities in the GeorefExamples_Florida.xls file.

a. Enter the locality string, country, state, and county information from the Excel sheet.
b. Click “Georeference.”
c. Inspect the “Possible Locations” by clicking on the “XX possible locations found” where XX is the number of locations GEOLocate identified.
d. Use an alternative resource to double check the locality. Try Google Maps.
e. Adjust the point location as you see fit. The green point is the active one.
f. Click the green point on the map, then click “Edit uncertainty”. Adjust the uncertainty radius by moving the grey arrow.
g. Return to the “Workbench” and record the latitude, longitude, and uncertainty.
h. If the uncertainty is >1000, it then discards the points.

Time to Exercise!

- ✓ Check results…

Time to Exercise!

2. Use the batch GEOLocate client to upload the localities in the GeorefExamples.xls file.

a. Copy and paste the appropriate information from the GeorefExamples.xls file into your own GEOLocateBatchFormat.csv.

i. http://www.geo-locate.org/standalone/tutorial.html

ii. Do not label the columns (your first row = first sample)

iii. **Make sure to save as a .csv**

iv. The majority of the columns will be empty

b. Go to the batch GEOLocate client and upload the formatted csv file

c. "Page Georeference" will georeference all eight localities available at once. "Georeference" will do one at a time.

d. Select a locality and go through Steps 1c to 1g. Once you are pleased with the locality and uncertainty, click "Correct" to note that you have gone through this georeference.

e. Work through the remaining localities.

f. If you do not finish a batch georeferencing, you can click on "File Management" at the bottom of the screen to receive a retrieval code. This will allow you to re-access this file whenever you wish without the need to download and upload.

g. If you do finish a batch georeferencing, you can click on "File Management" and then "Export" to download the finished georeferenced file.

Time to Exercise!

2. Use the batch GEOLocate client to upload the localities in the GeorefExamples.xls file.

i. http://www.geo-locate.org/standalone/tutorial.html

How to import data from Microsoft Excel into GEOLocate

Create a blank spreadsheet within Excel. Add the following to the first row of columns in your spreadsheet:

	Column 1	Column 2	Column 3	Column 4	Column 5	Column 6	Column 7	Column 8	Column 9	Column 10
Row1	locality string	country	state	county	latitude	longitude	correction status	precision	error polygon	multiple results

After the 10th column, you may add the names of other fields in your dataset that you would like to include (field number, localityID, remarks etc.) but you must be certain that each entry in the first row is unique.

Copy and paste your data from its original source to the appropriate columns starting at row 2 (alternatively, you could modify an existing spreadsheet to conform to the above format).

Select *Save As*... from the file menu in Excel. Enter a filename for your file, select CSV *(comma delimited) (*.csv)* from the drop down box labeled *Save as type*, and click the save button. Exit Microsoft Excel.

You now have a file that can be used with GEOLocate.

From the GEOLocate file menu, select Import -> CSV, and select the file you just created to import.

***NOTE: Our web based file client now supports the generation of uncertainty radii.**
If you wish to generate and record uncertainty radii for your records using this client, your spreadsheet must also include an additional column to store that data. In this case you would need at least 11 columns for your data instead of 10. Using uncertainty radii is optional, so you can ignore this, if you choose not to utilize that option.

Time to Exercise!

2. Use the batch GEOLocate client to upload the localities in the GeorefExamples.xls file.

i. http://www.geo-locate.org/standalone/tutorial.html

Time to Exercise!

2. Use the batch GEOLocate client to upload the localities in the GeorefExamples.xls file.

a. Copy and paste the appropriate information from the GeorefExamples.xls file into your own GEOLocateBatchFormat.csv.

i. http://www.geo-locate.org/standalone/tutorial.html

ii. Do not label the columns (your first row = first sample)

iii. **Make sure to save as a .csv**

iv. The majority of the columns will be empty

b. Go to the batch GEOLocate client and upload the formatted csv file

c. "Page Georeference" will georeference all eight localities available at once. "Georeference" will do one at a time.

d. Select a locality and go through Steps 1c to 1g. Once you are pleased with the locality and uncertainty, click "Correct" to note that you have gone through this georeference.

e. Work through the remaining localities.

f. If you do not finish a batch georeferencing, you can click on "File Management" at the bottom of the screen to receive a retrieval code. This will allow you to re-access this file whenever you wish without the need to download and upload.

g. If you do finish a batch georeferencing, you can click on "File Management" and then "Export" to download the finished georeferenced file.

Time to Exercise!

- ✓ Check results…

Session Summary, Q&A and Discussion

- ✓ Understand the importance of georeferencing for biodiversity analyses
- ✓ Understand the sources of uncertainty in a georeference
- ✓ Be familiar with GEOLocate
- ✓ Use GEOLocate to georeference localities

Links with links...

iDigBio
Integrated Digitized Biocollections

https://www.idigbio.org/wiki/index.php/Georeferencing

http://herpnet.org/Gazetteer/GeorefResources.htm

Session Summary, Q&A and Discussion

Any questions??

Please use the chat for questions!