iDigBio
Integrated Digitized Biocollections

ENM Crash Course: Georeferencing

Edmonton, Alberta
July 26th, 2015

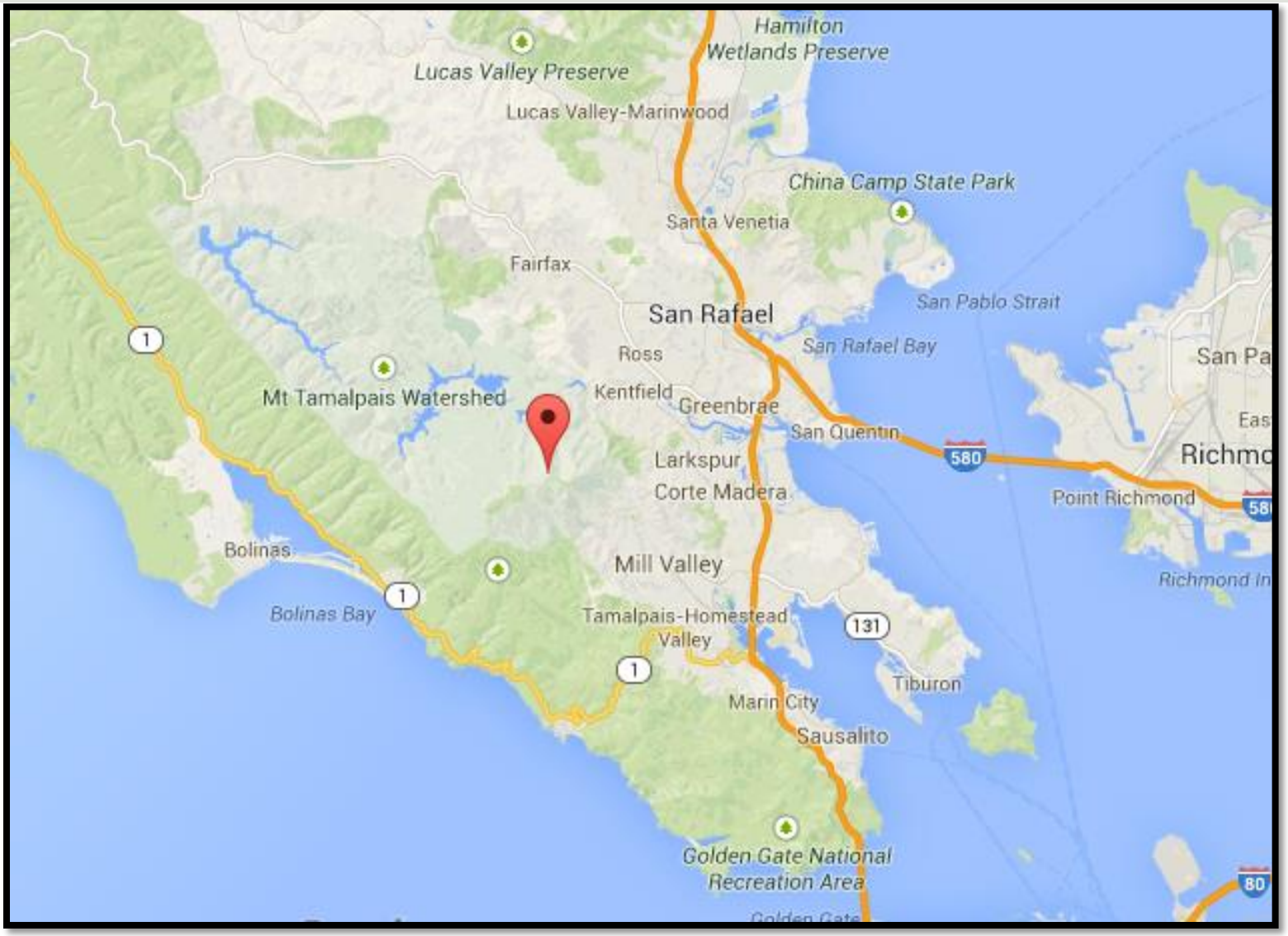
UF UNIVERSITY of FLORIDA
FLORIDA STATE UNIVERSITY 1851
FLORIDA MUSEUM OF NATURAL HISTORY
NSF

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. All images used with permission or are free from copyright.

What is a georeference?

A numerical description of a place that can be mapped.

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



37.930, -122.587

Why Georeference?

- Correct geographic and specimen identification data = dependable occurrence record.

- Occurrence data validates the importance of biological collections, especially to non-taxonomists.
  - Distribution of populations and species ranges
  - Phylogeography
  - **Niche modeling**
  - Conservation planning and biodiversity management

- Provides uncertainty data, which allows data to be evaluated with regards to its fitness for research application and resulting quality of output.

Precision vs. Accuracy Disambiguated

Precision is…

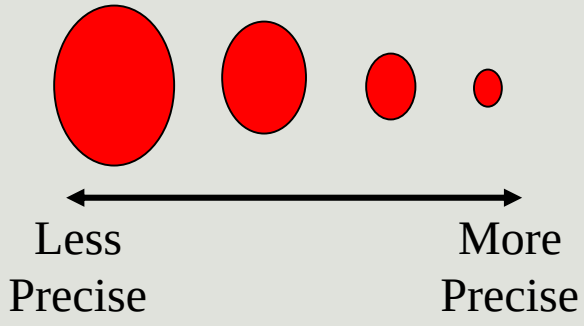
…the level of detail contained in or described by the data.

Example:

42, precise

42.1, more precise

42.01, even more precise



Precision can help to minimize uncertainty.

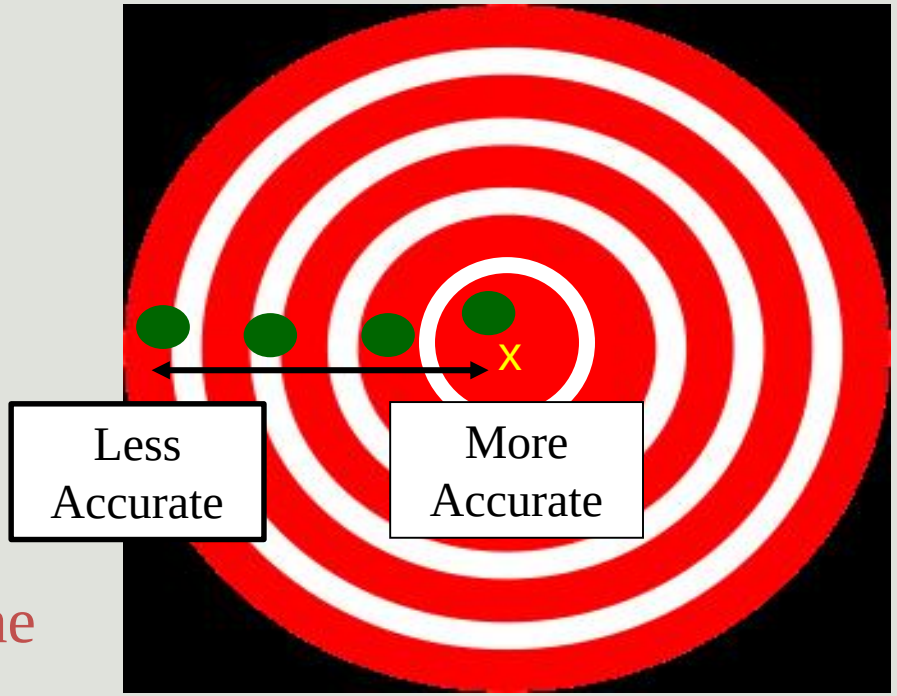
Precision vs. Accuracy Disambiguated

Accuracy is…

…a measure of how close a given value is to the true value.*

Example: Truth = 42

41.999 = more precise, less accurate



*We may never actually know the true value in georeferencing, but we do our best to reproduce the location of the true location.

Precision vs. Accuracy Disambiguated

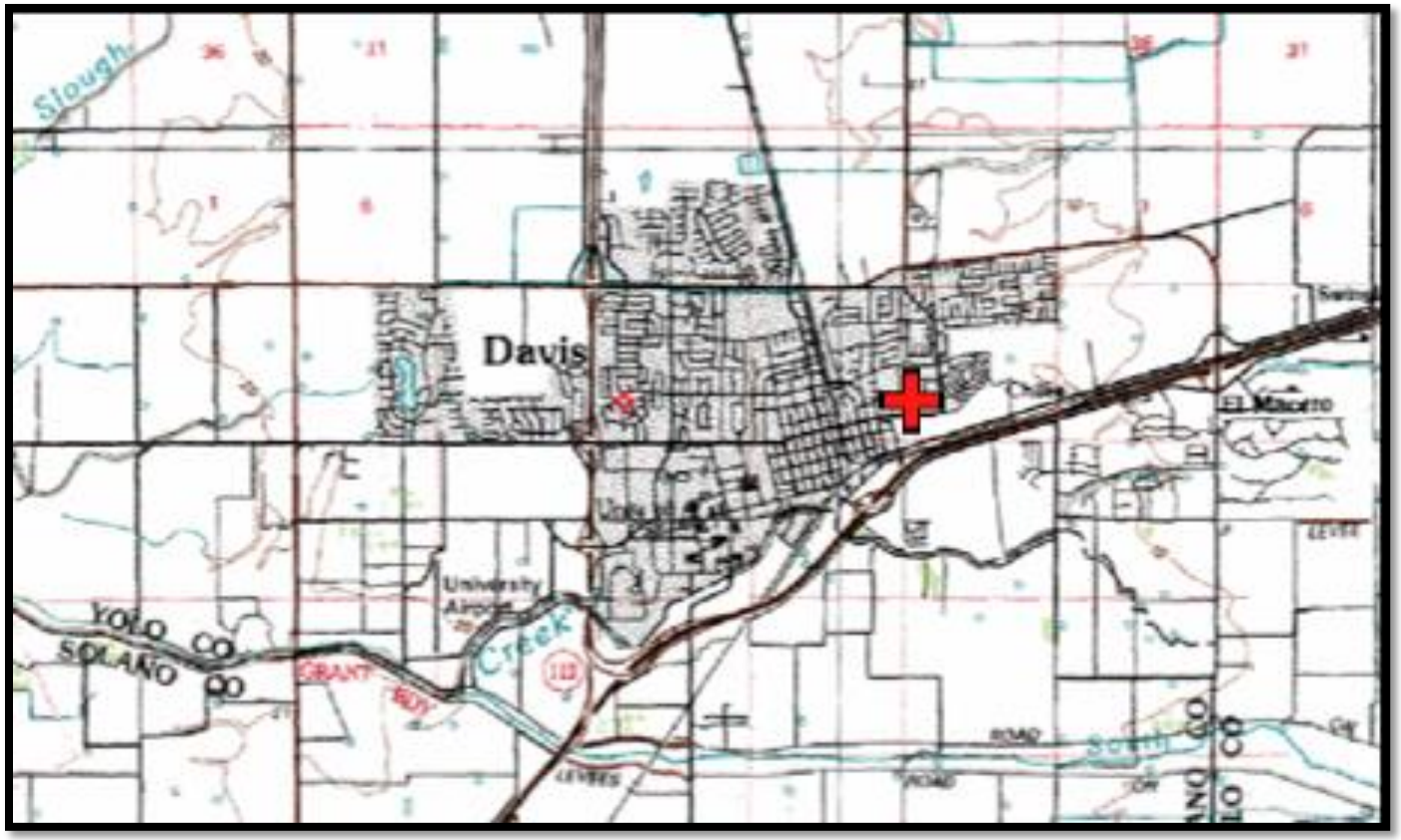
Final Answer?

The sum of what we know…

about what we don't know…

is the **uncertainty**.

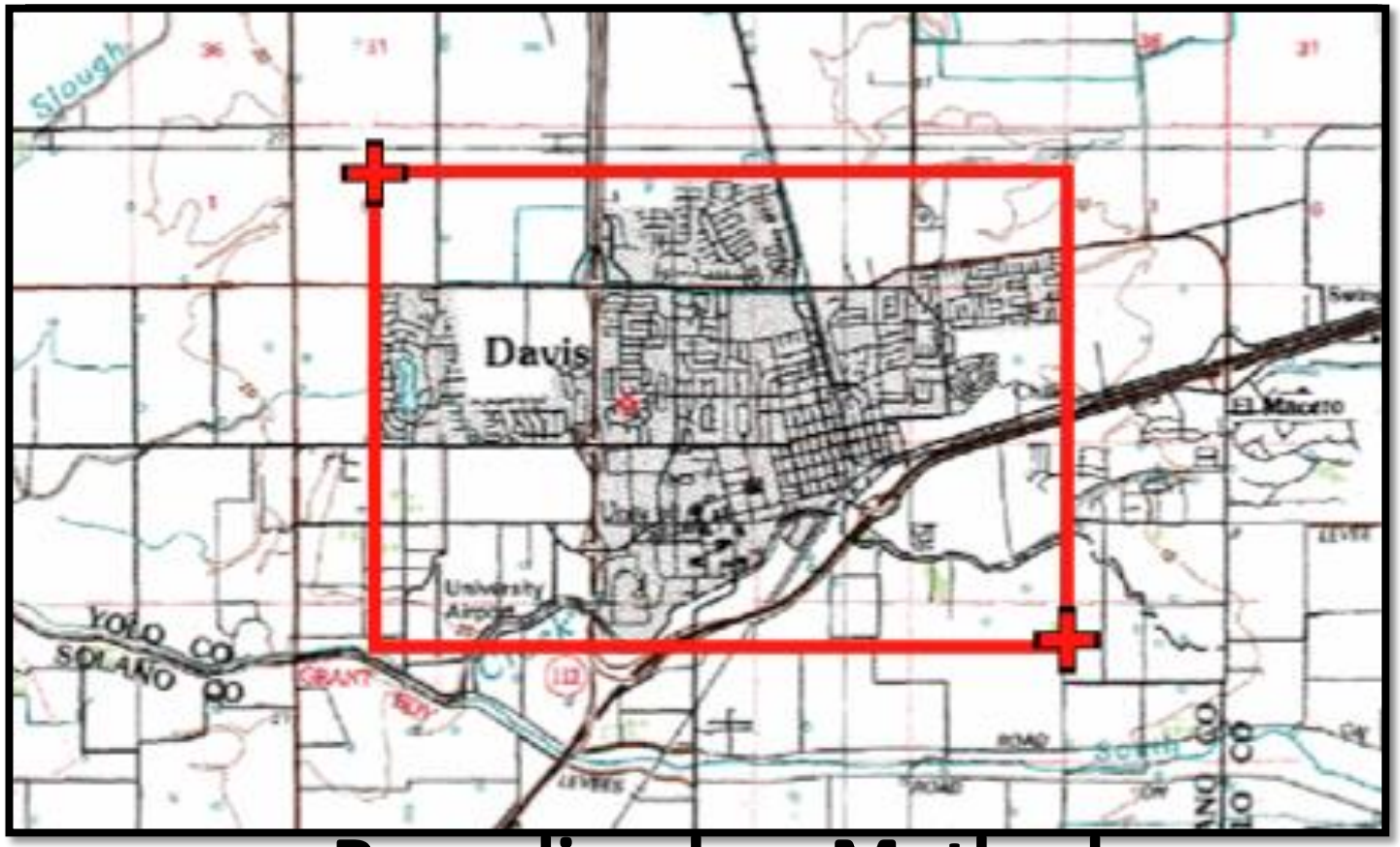
“Davis, Yolo County, CA”



Point Method

Coordinates: 38.5463, -121.7425

“Davis, Yolo County, CA”



Bounding-box Method

Coordinates: 38.5486, -121.7542

38.545, -121.7394

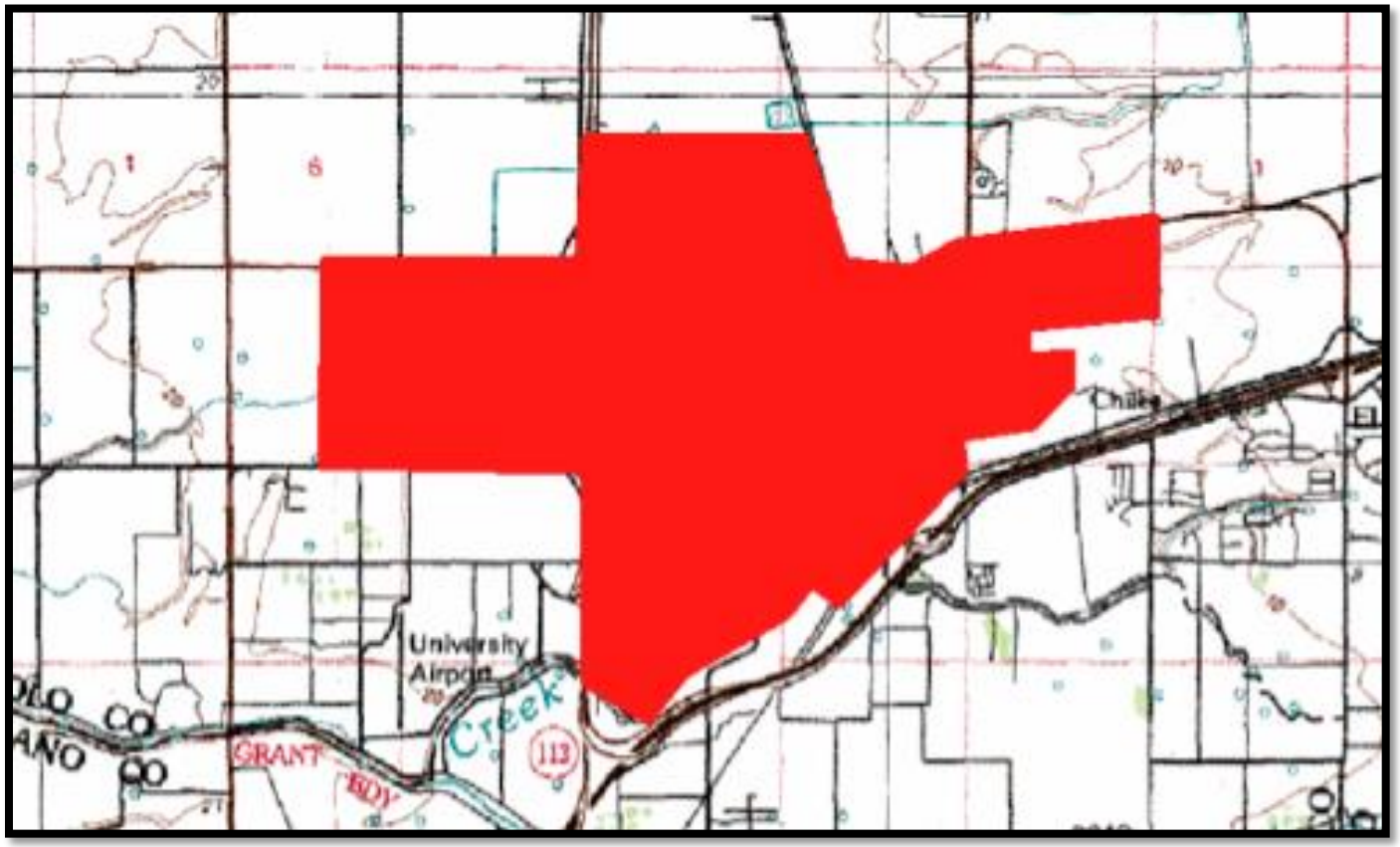
“Davis, Yolo County, CA”

Point-Radius Method

Coordinates: 38.5468 -121.7469

Maximum Uncertainty: 8325 m

“Davis, Yolo County, CA”



Shape Method

What is an *ideal* georeference?

A numerical description of a place that can be mapped

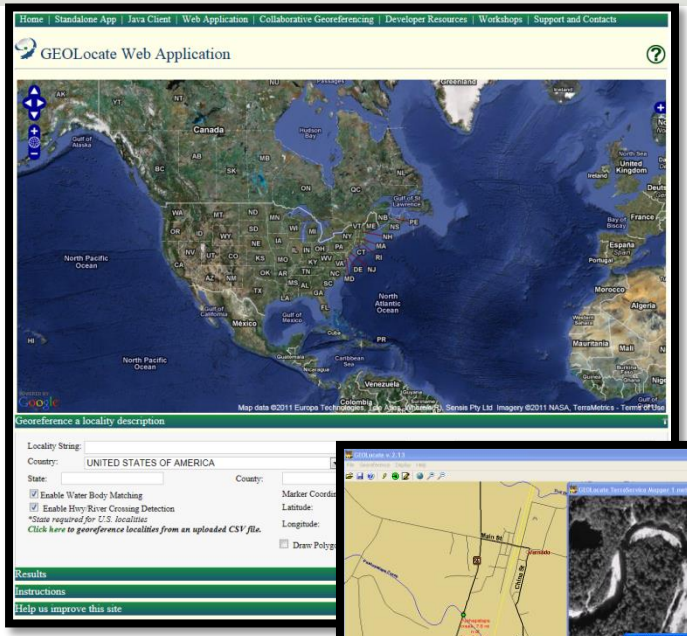
and that describes the spatial extent of a locality and its associated uncertainties

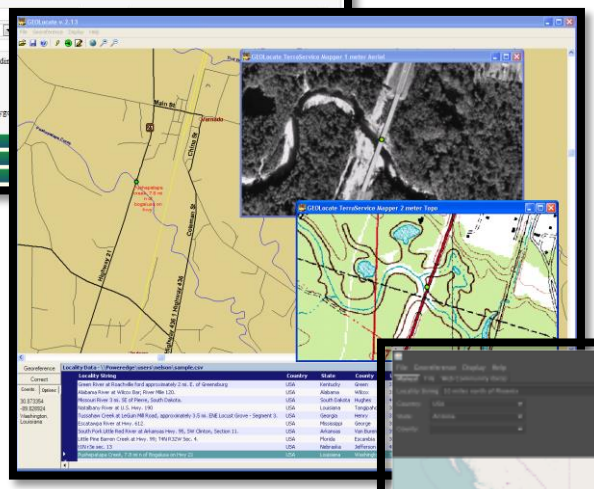
as well as possible

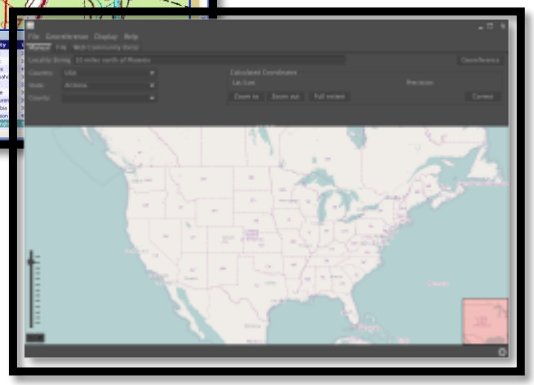
How to do it?

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

- Internet Resources







GEOLocate

http://www.museum.tulane.edu/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

↓

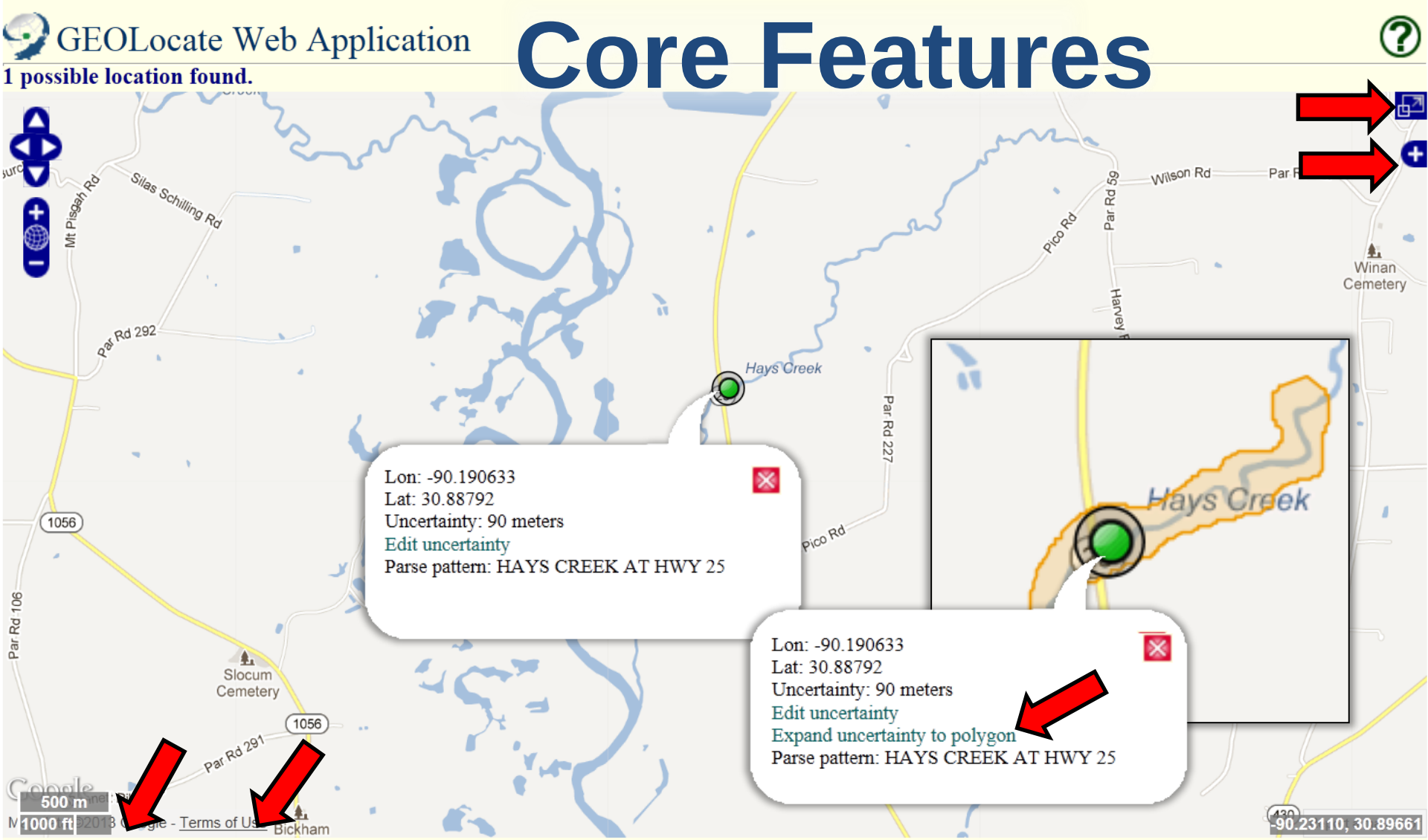
Visualize, verify & adjust output coordinates & uncertainties

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

uncertainty radius: **48m**

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

Core Features



Workbench | 1 possible location found

Georeference | Options | Draw polygon | Place marker | Measure

Locality String: Hays Creek, 3 mi. n of Franklinton on hwy. 25

Country: UNITED STATES OF AMERICA

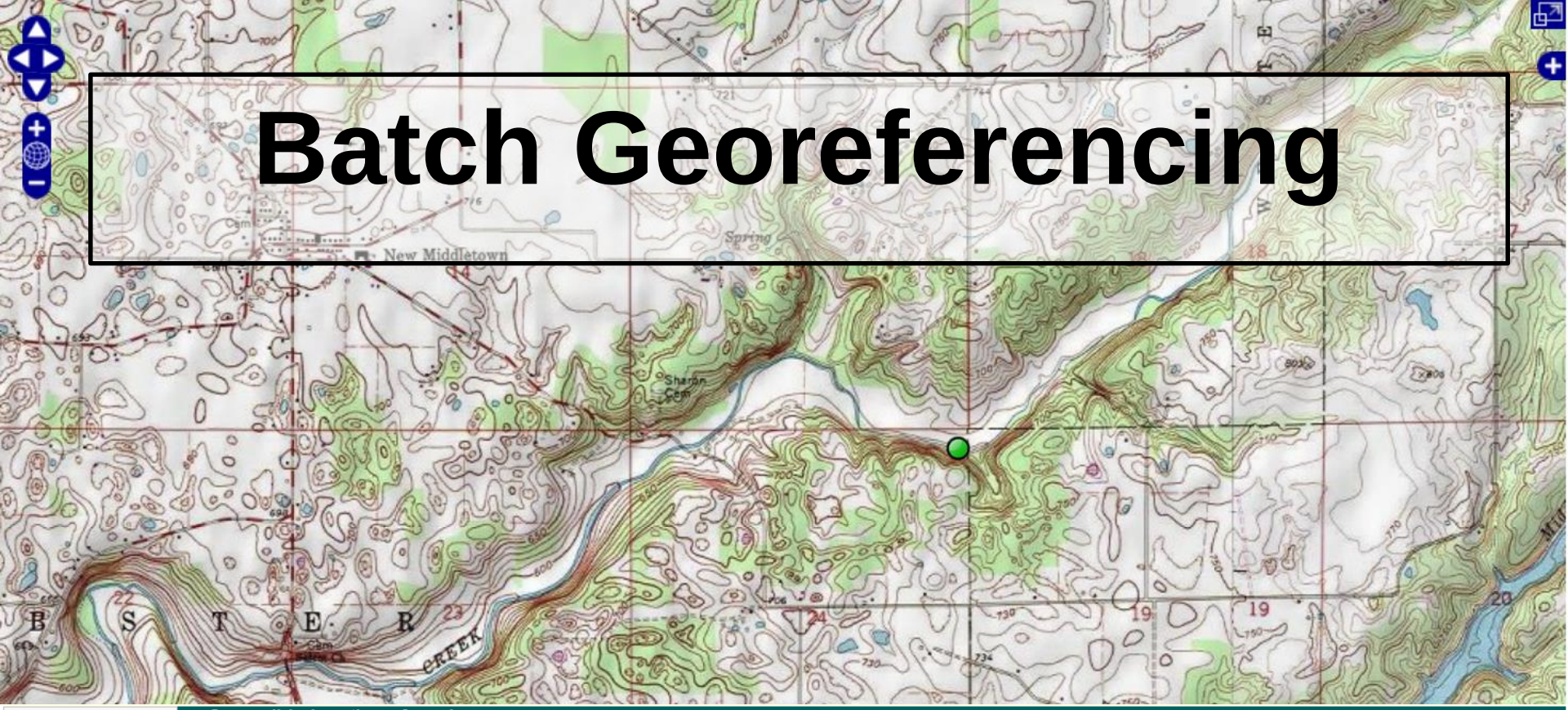
State: La

County: Washington

latitude: 30.88792 | longitude: -90.190633 | uncertainty: 90 m | error polygon

30.88792 -90.190633 90 Unavailable

Batch Georeferencing



Workbench | 2 possible locations found

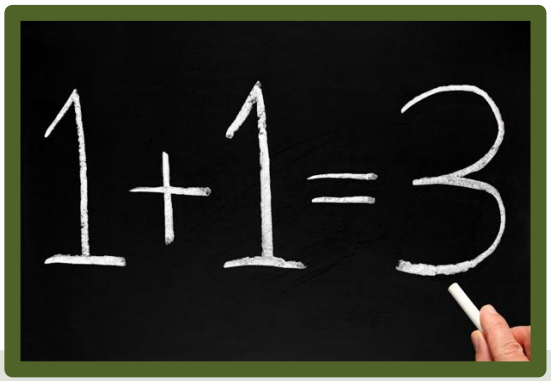
Show 8 records | Page Georeference | Georeference | Options | Correct | Draw polygon | Place marker | Measure

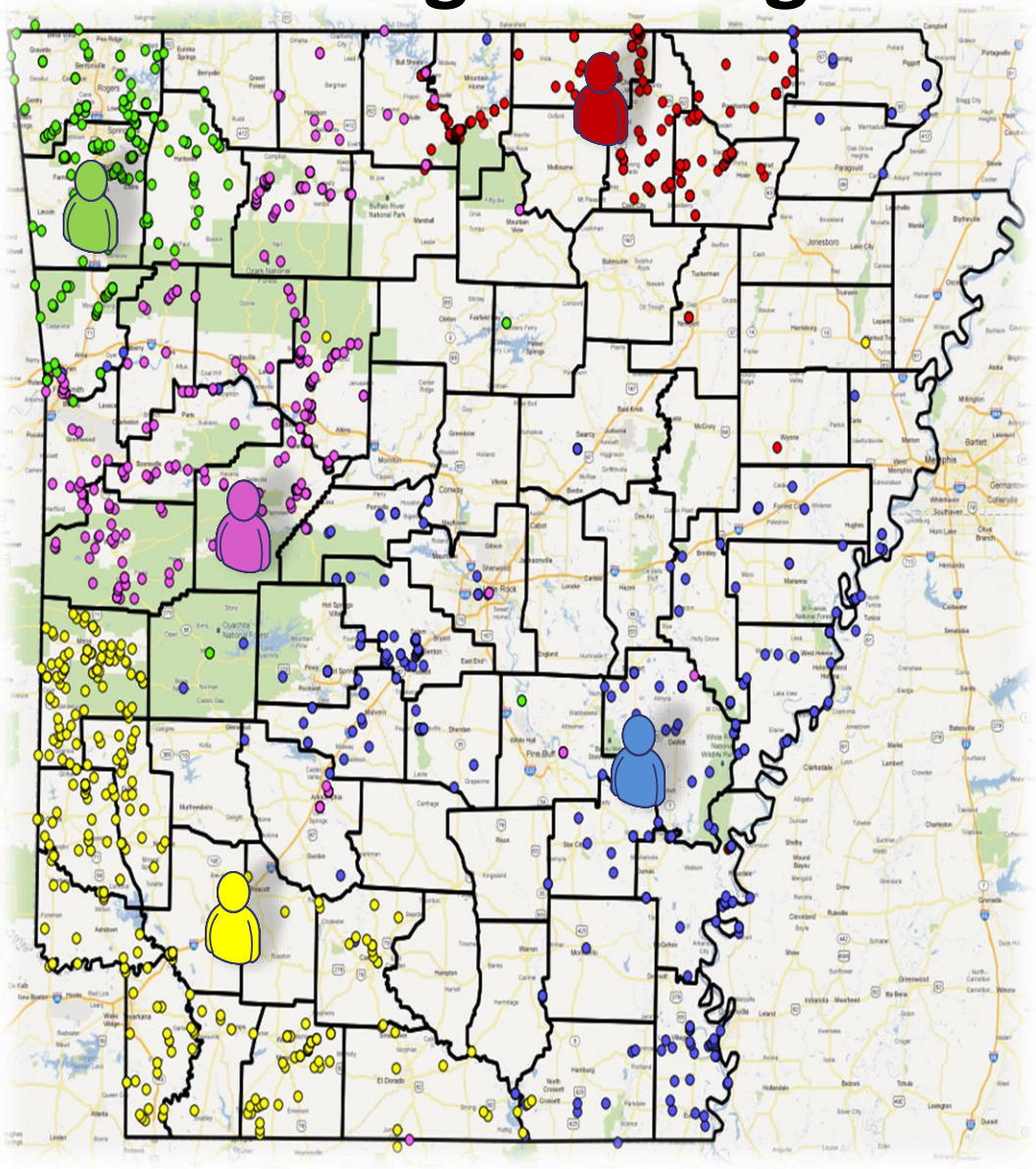
Locality	Country	StateProvince	County	Latitude	Longitude	Corrected	precision
Chambers Spring Road 2.5 km S of Hwy 412, 8.0 km E of Siloam Springs, T17N, R33W,	USA	Arkansas	Benton	36.188027	-94.451005	no	High(89)
Osage Creek, 1.0 mile N on gravel road to bridge crossing; gravel road jcts with	USA	Arkansas	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: Sharing Data

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





Google Maps

- **Search & directions**
  - Free text search, depends on **zoom level**
  - Directions for travelling by car, bike, public transport and foot
  - Data compiled from different sources
- **Maps**
  - Views: map, satellite (= aerial photography), terrain, Google Street View + various layers of information
  - Data compiled from different sources (indicated at the bottom of each map)

http://maps.google.com

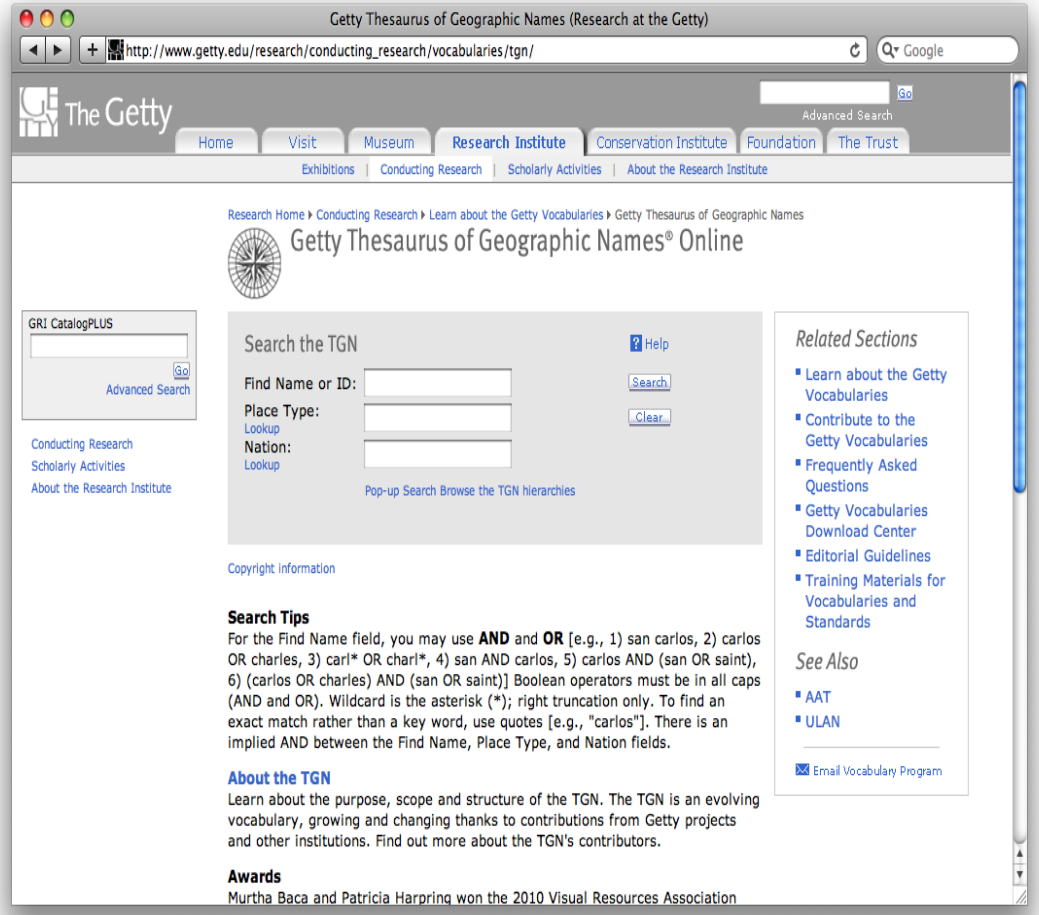
Falling Rain

- Worldwide gazetteer for **cities** and **towns**
- Great for **hard to find localities**, especially outside US
  - Browse to find locality (**no search**)
  - Provides hierarchy, alternative names, topo maps, altitude, weather information and location of nearby towns in nautical miles (nm)
  - Example: Qaryeh-ye Gol`alam, Velayat-e Lowgar, AF

http://www.fallingrain.com

Getty Thesaurus of Geographic Names (TGN)

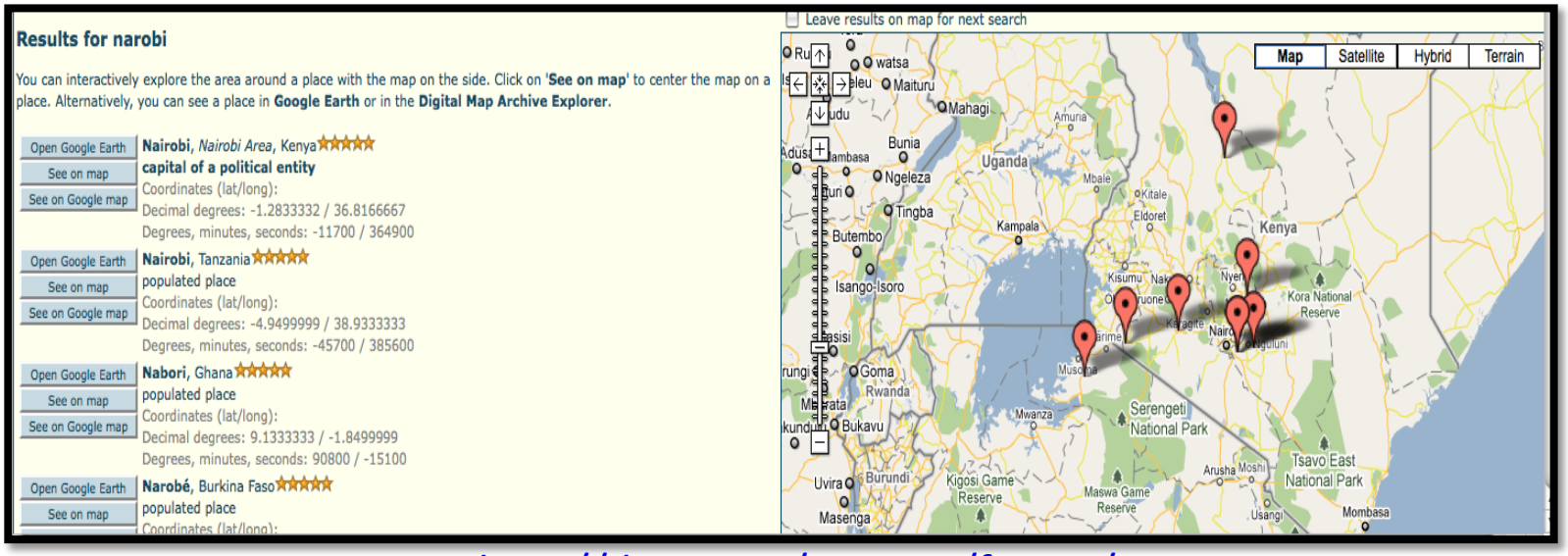
- Worldwide gazetteer by The Getty
- Useful for finding **alternative** and **old names**
  - Feature types
  - Geographical hierarchy
  - Degrees minutes or **no coordinates!** Use recent name and search in Google Maps
- Example: New Amsterdam, US



http://bit.ly/Getty-TGN

FuzzyG – JRC Fuzzy Gazetteer

- Worldwide gazetteer designed for **bad spelling**.
- Useful for finding alternative, doubtful spelling and old names
  - Feature types, by continent
  - Degrees minutes or no coordinates
  - Use recent name and then search in Google Maps
- Example: Narobi instead of Nairobi, Africa



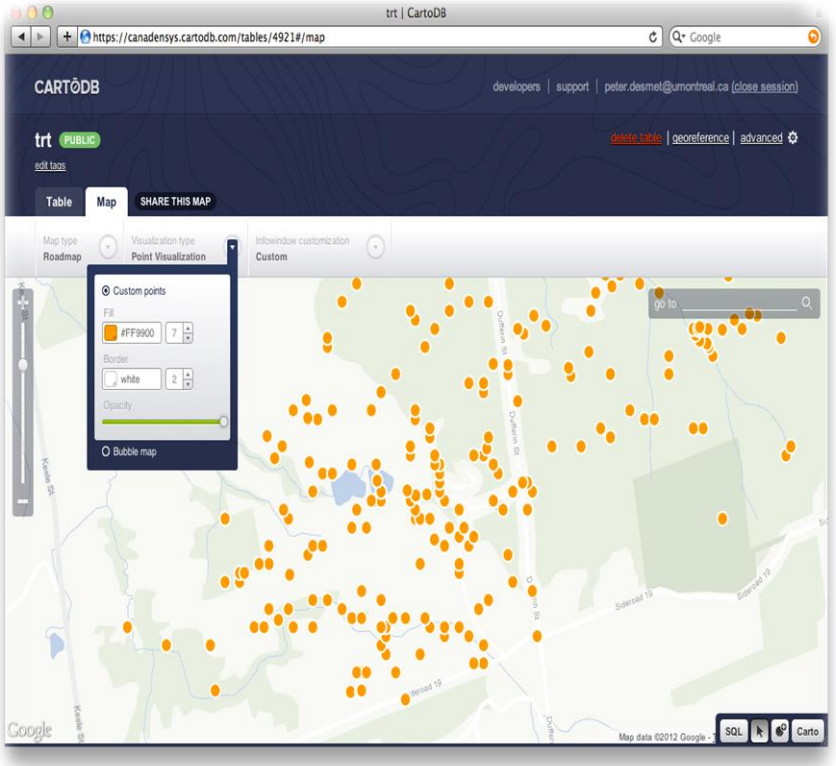
http://dma.jrc.it/services/fuzzyg/

Other tools - Visualizing data

- GPS Visualizer
  - Use to translate a file with coordinates into kml or a picture
- CartoDB
  - Online geospatial database by Vizzuality
  - Useful for visualizing your data
  - Drag and drop CSV upload
  - Easy customization of your map
  - Share and embed your map
  - Powerful development tools
  - Free account = 5 tables / 5MB of data

GPS Visualizer

http://www.gpsvisualizer.com/



http://cartodb.com

GEOREFERENCING QUICK REFERENCE GUIDE

Version: 2012-10-08

John Wieczorek, David Bloom, Heather Constable, Janet Fang, Michelle Koo, Carol Spencer, Kristina Yamamoto

This is a practical guide for georeferencing using the point-radius method [1, 2, 3] using the Georeferencing Calculator [4, 5], maps, gazetteers, and other resources from which coordinates and spatial boundaries for places can be found. This guide is an update of "Georeferencing for Dummies" [6], and explains the recommended calculation procedure for localities encountered in the georeferencing process.

Georeferences using the methods in this guide will be maximally useful if as much information as possible is captured about and during the georeferencing process in the following fields defined in the Darwin Core standard [7]. For additional community discussion and recommendations, see the Darwin Core Project wiki [8].

Darwin Core Georeferencing terms:

- **decimalLatitude**, **decimalLongitude**, **geodeticDatum** – the combination of these three fields provide the reference for the center of the point-radius representation of the georeference.
- **coordinateUncertaintyInMeters** – The horizontal distance (in meters) from the given **decimalLatitude** and **decimalLongitude** describing the smallest circle containing the whole of the Location. Leave the value empty if the uncertainty is unknown, cannot

http://www.herpnet.org/herpnet/documents/GeoreferencingQuickGuide.pdf

Links with links

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

http://www.canadensys.net/georeferencing

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