

The image shows a collection of fossil specimens housed in several cardboard boxes. The boxes are arranged on a wooden surface. One box in the foreground is open, revealing several dark, flat, irregularly shaped fossil fragments. A small, light-colored card with a red square and the number '25' is placed on top of the fossils. To the right, another box contains a larger, more complex fossil specimen. In the background, more boxes are visible, some containing smaller fragments. The overall scene is a workspace for paleontological collections.

# Georeferencing Paleontology Collections

Una Farrell, University of Kansas

Jessica Utrup, Yale Peabody Museum

iDigBio Paleodigitization Workshop, Sept. 24th -25th 2013

**What is a georeference?**



# What is a georeference?

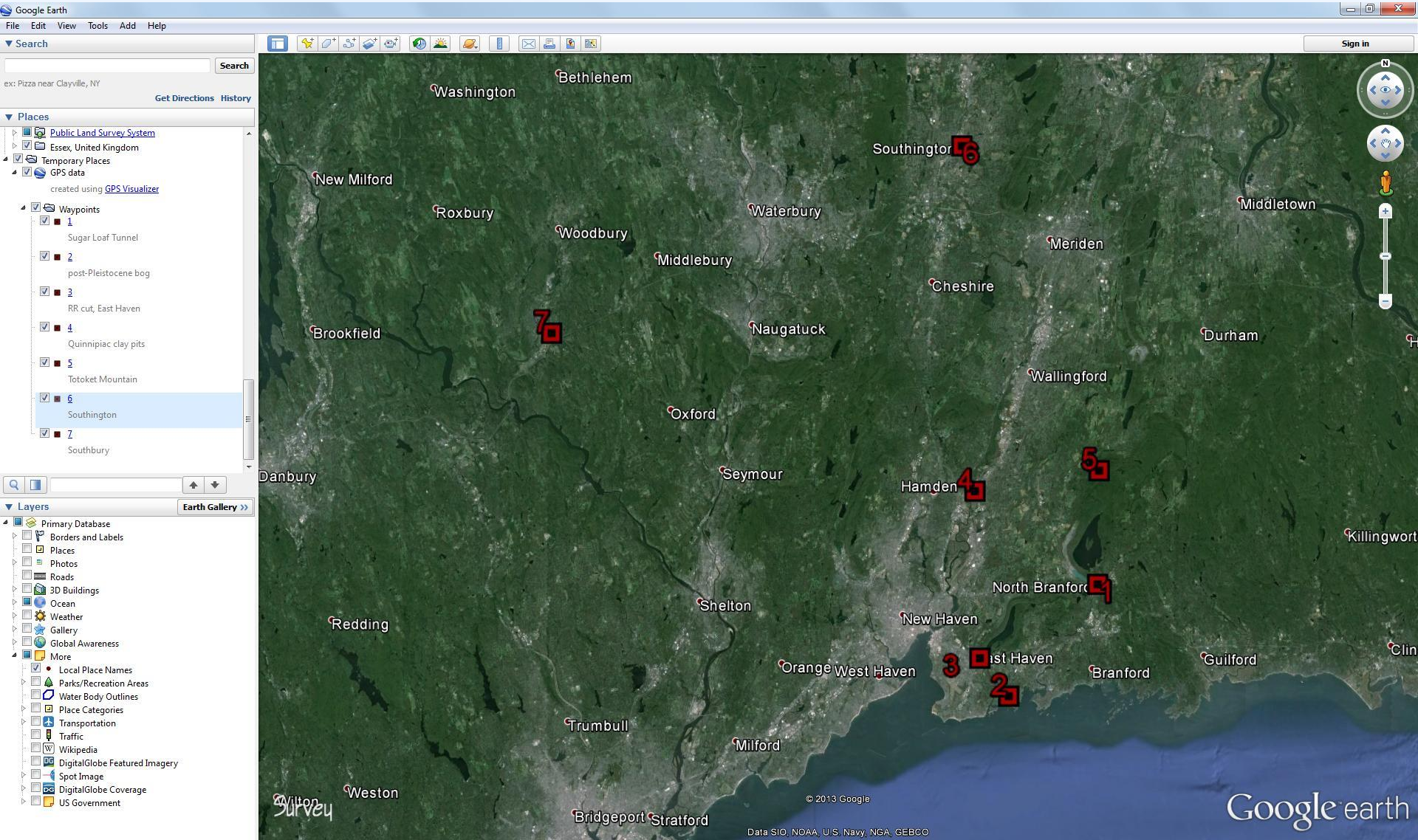
A numerical description of a place that can be mapped.



# What we have: locality descriptions

ID	Taxon	Locality description	Stratigraphy
1	<i>Anchisauripus hitchcocki</i>	232 ft in from west portal, Sugar Loaf Tunnel. North Branford, CT.	Triassic. "Posterior shales."
2	<i>Littorina palliata</i>	6-8 ft below surface of post-Pleistocene bog, 0.25 mi S10E of intersection of Shore Dr. & Killams Point Rd. Branford, CT.	Quaternary. Holocene.
3	<i>Clathropteris meniscoides</i>	NH RR cut, East Haven, CT.	Late Triassic. Newark Supergroup. Shuttle Meadow Fm.
4	<i>Rangifer tarandus</i>	11 ft below surface, Quinnipiac clay pits. North Haven, CT.	Quaternary. Pleistocene.
5	<i>Diplurus longicaudatus</i>	W slope of Totoket Mountain, south part of Durham, CT.	Newark Supergroup.
6	<i>Fucusopsis angulatus</i>	Southington, CT.	Triassic.
7	<i>Pomperaugoxylon connecticutense</i>	David Blersch farm, on N slope of Horse Fence Hill. Southbury, CT.	Late Triassic. Newark Supergroup. South Britain Fm. Pierce Hollow Mbr.

# What we want: localities we can map



# Why georeference?

- Correct geographic and specimen identification data = dependable **occurrence data**
- Occurrence data validates the importance of paleontological and biological collections
  - Distribution of populations and species ranges
  - Phylogeography
  - Niche modelling
  - 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.



# The scale of the georeferencing problem for natural history collections...

~ $2.5 \times 10^9$  records

~6 records per locality\*

~14 localities georeferenced  
per hour\*

~15,500 years

\* based on the MaNIS Project



# Darwin Core Location Terms

- higherGeography
- waterbody, island, islandGroup
- continent, country, countryCode, stateProvince, county, municipality
- locality
- minimumElevationInMeters, maximumElevationInMeters, minimumDepthInMeters, maximumDepthInMeters





# Darwin Core Georeference Terms

- decimalLatitude, decimalLongitude
- geodeticDatum
- coordinateUncertaintyInMeters
- georeferencedBy, georeferenceProtocol
- georeferenceSources
- georeferenceVerificationStatus
- georeferenceRemarks
- coordinatePrecision
- pointRadiusSpatialFit
- footprintWKT, footprintSRS, footprintSpatialFit



# What is a georeference?

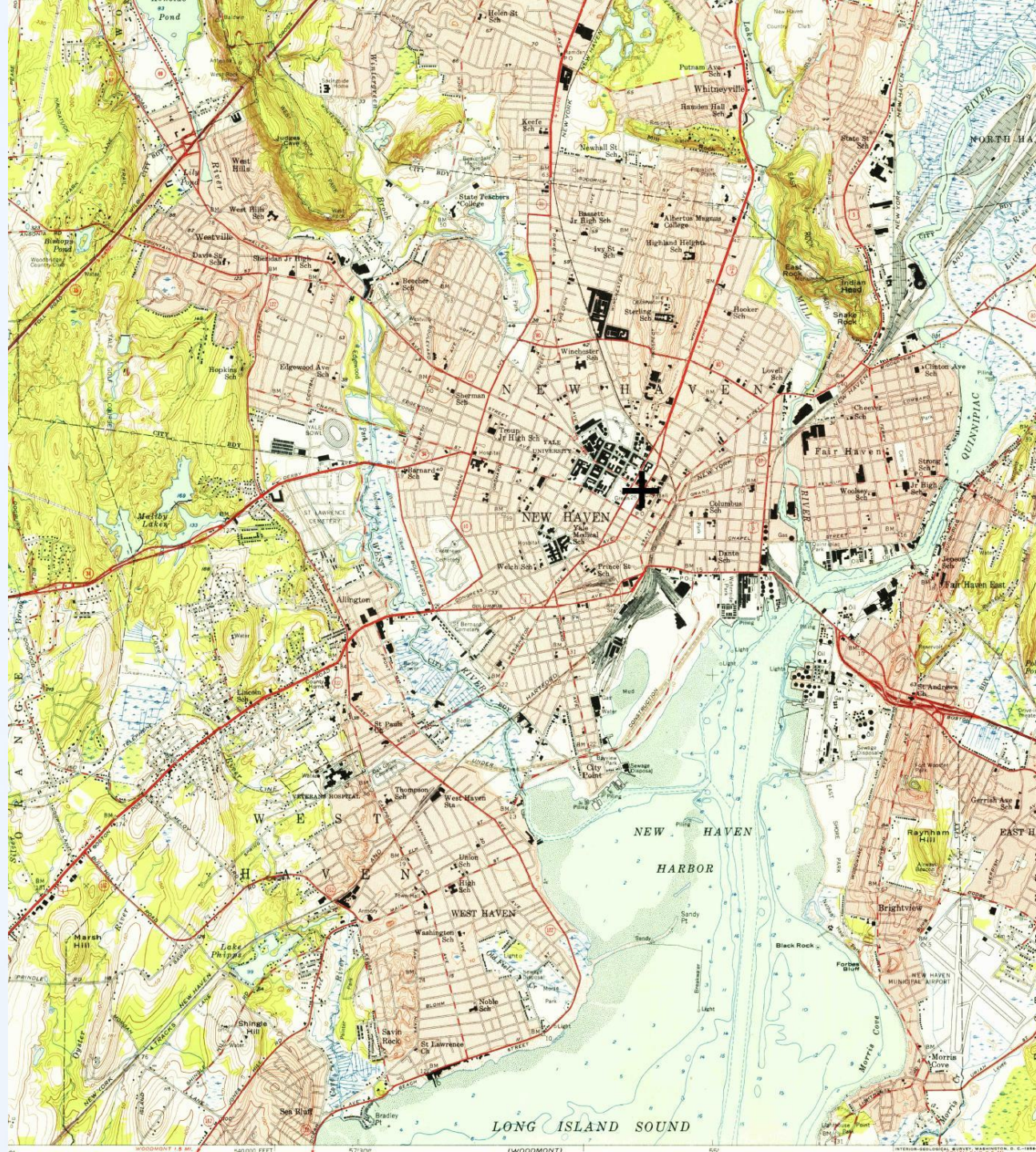
A numerical description of a place that can be mapped.



**Location:**  
New Haven,  
Connecticut

**Coordinates:**  
41.30861, -72.92468  
(NAD27)

**Method:** Point



# What is a *acceptable* georeference?

A numerical description of a place that can be mapped,

*and that describes the **spatial extent** of a locality and its **associated uncertainty**.*



**Location:**

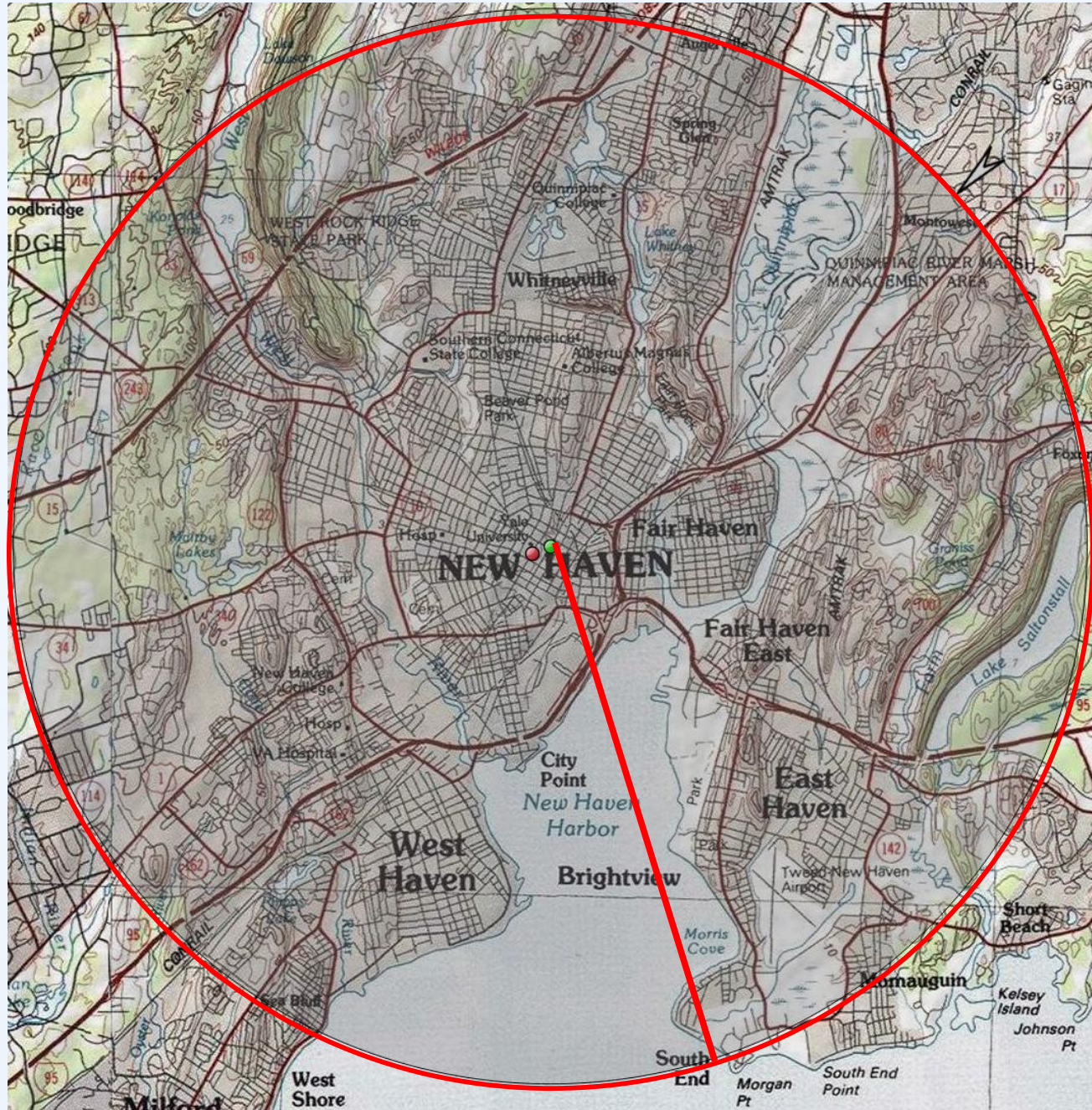
New Haven,  
Connecticut

**Coordinates:**

41.30861, -72.92468  
(NAD27)

**Uncertainty:** 7642m

**Method:** Point-radius



# 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 uncertainty  
*as well as possible.*



**Location:**  
New Haven,  
Connecticut

**Coordinates:**  
41.30861, -72.92468  
(NAD27)







# Protocol: MaNIS/HerpNET/ORNIS (MHO)

## Guidelines

### GBIF Best Practices

- Uses **point-radius** representation of georeferences
- Circle encompasses **all sources of uncertainty** about the location
- Formalizes assumptions, algorithms, and documentation standards that promote **reproducible** results
- Methods are **universally applicable**



# GEOREFERENCING QUICK REFERENCE GUIDE

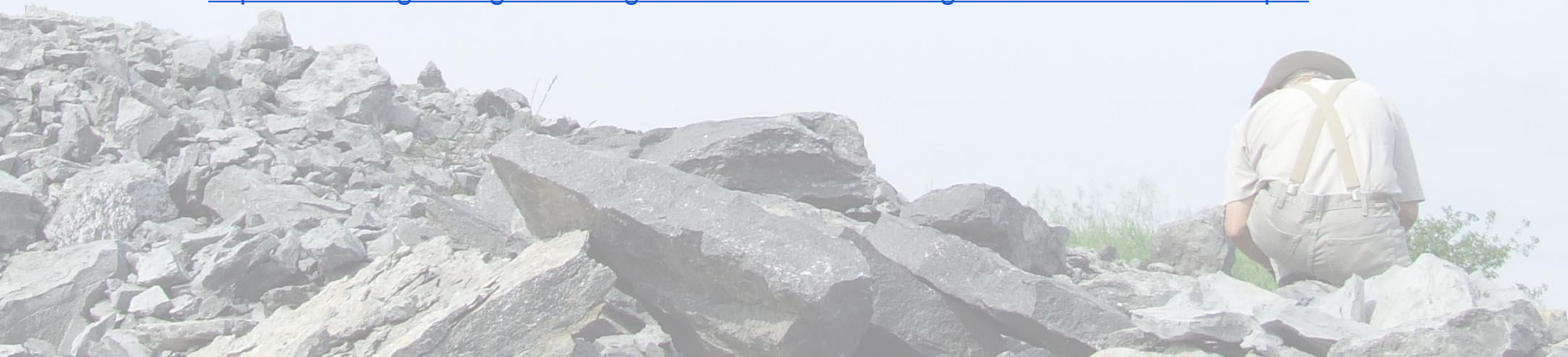
Version: 2012-10-02

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].**

<https://www.idigbio.org/wiki/images/1/1e/GeoreferencingQuickReferenceGuide.pdf>



# Locality Types

Named Place: Towns, farms, quarries, street address, junctions, rivers...

Offset from a named place: Distance along a path (e.g. along a highway), distance at a heading (e.g. miles N), orthogonal offset (e.g. miles N and E)

Coordinates: GPS, verbatim lat-long, grid system

Problems/More than one possibility



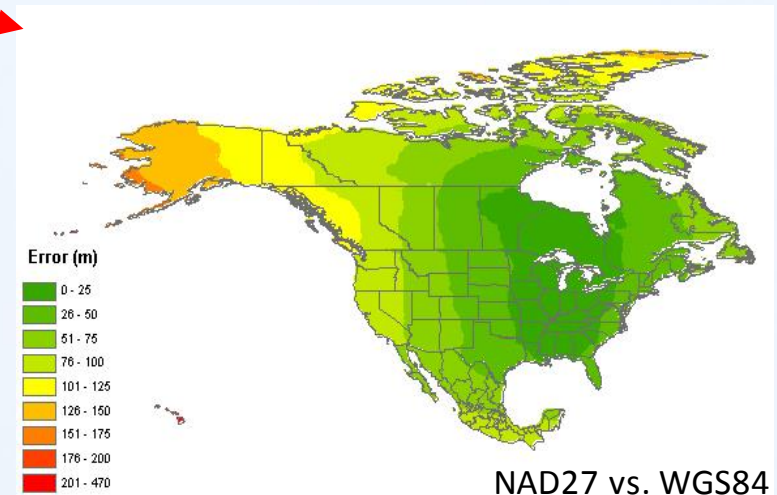
# Sources of Uncertainty

20° 30' N 112° 36' W

- Coordinate uncertainty
- Map Scale
- The extent of the locality
- GPS accuracy
- Unknown datum
- Imprecision in direction measurements (N vs. NNE)
- Imprecision in distance measurements (1km vs. 1.1km)



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



# Georeferencing Calculator

English (l... ▾)

Calculation Type

Coordinates and error - enter the Lat/Long for the named place or starting point ▾

Locality Type

Distance along orthogonal directions (e.g., 2 mi E and 3 mi N of Bakersfield) ▾

**Step 3) Enter all of the parameters for the locality.**

Coordinate Source

gazetteer ▾

North or South Offset Distance

.. ▾

Coordinate System

decimal degrees ▾

East or West Offset Distance

.. ▾

Latitude

Extent of Named Place

Longitude

Measurement Error

Datum

datum not recorded ▾

Distance Units

... ▾

Coordinate Precision

nearest degree ▾

Distance Precision

1 km ▾

Decimal Latitude

Decimal Longitude

Maximum Error Distance

**Calculate**

**Promote**

Distance Converter:

... ▾

=

... ▾

Scale Converter:

... ▾

1:24000 ▾

=

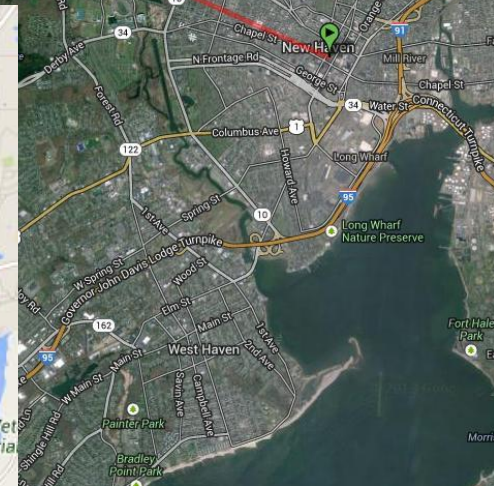
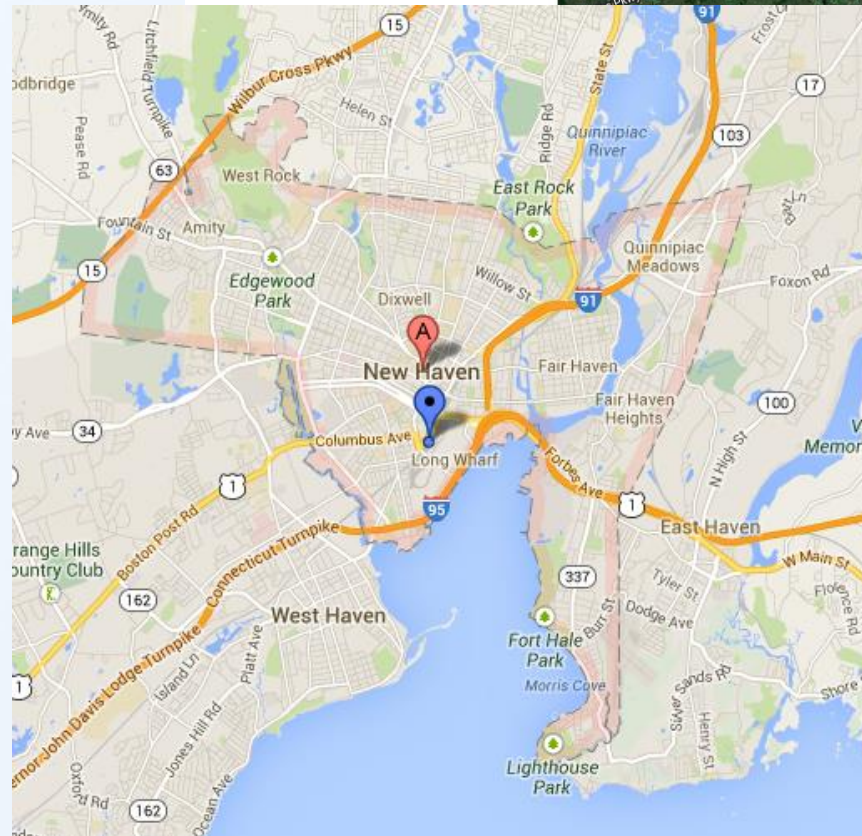
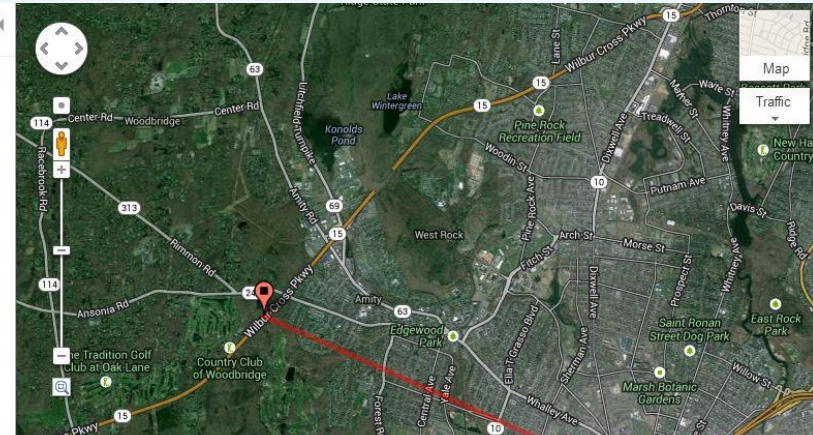
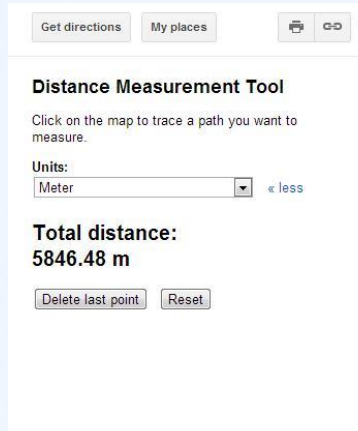
... ▾

Version 20130205en

copyright (c) 2001-2013 Regents of the University of California

# Google maps

- Measuring tools, lat-long markers in **Map Labs**
- Satellite vs. map view - old quarries may be visible
- Good for mapping addresses
- Polygons for determining extent





Google  
Earth

Measuring tools

Outcrops,  
roadcuts, quarries



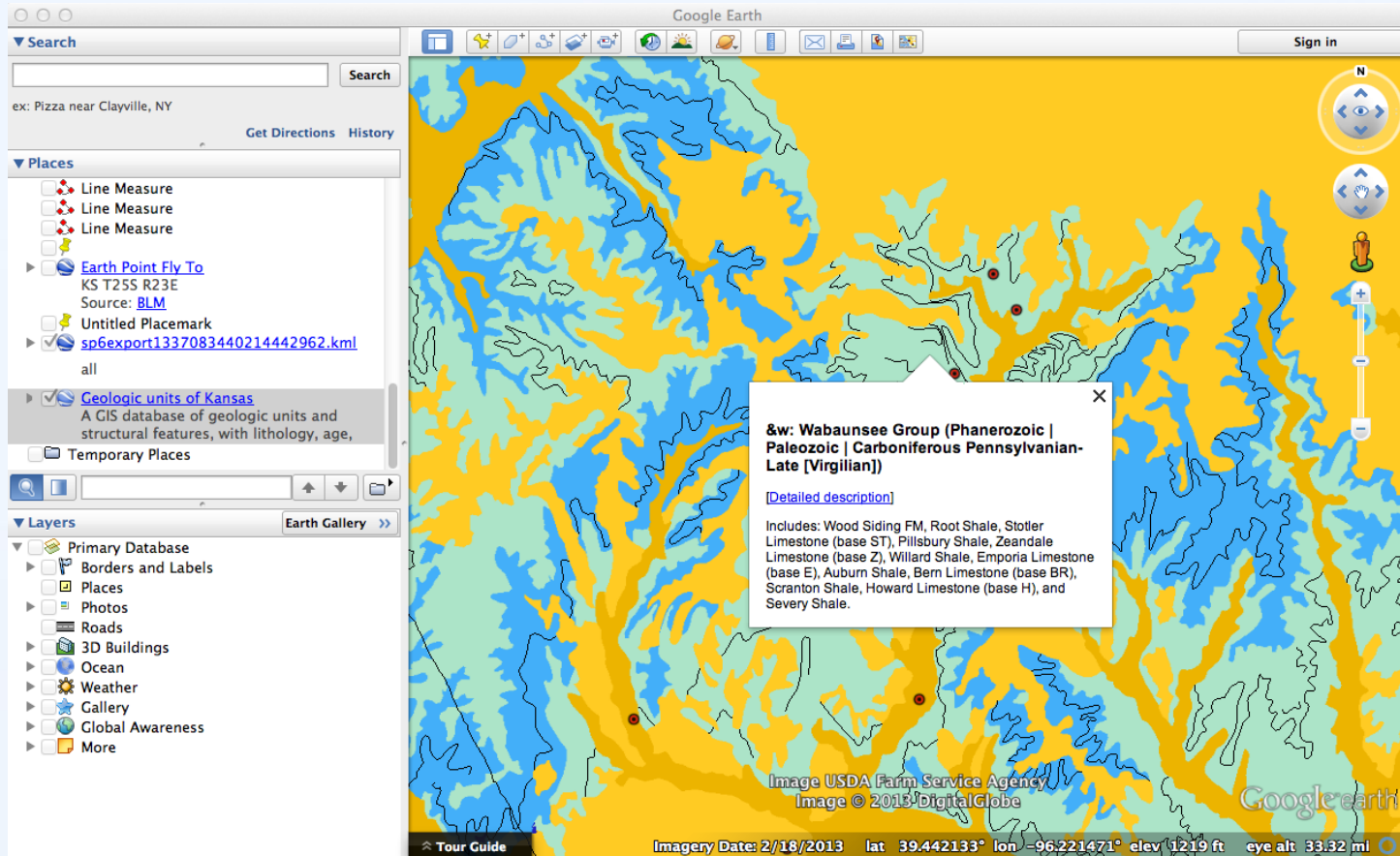


Measuring tools

Outcrops,  
roadcuts, quarries

Geology

<http://mrdata.usgs.gov/geology/state/>







Measuring tools

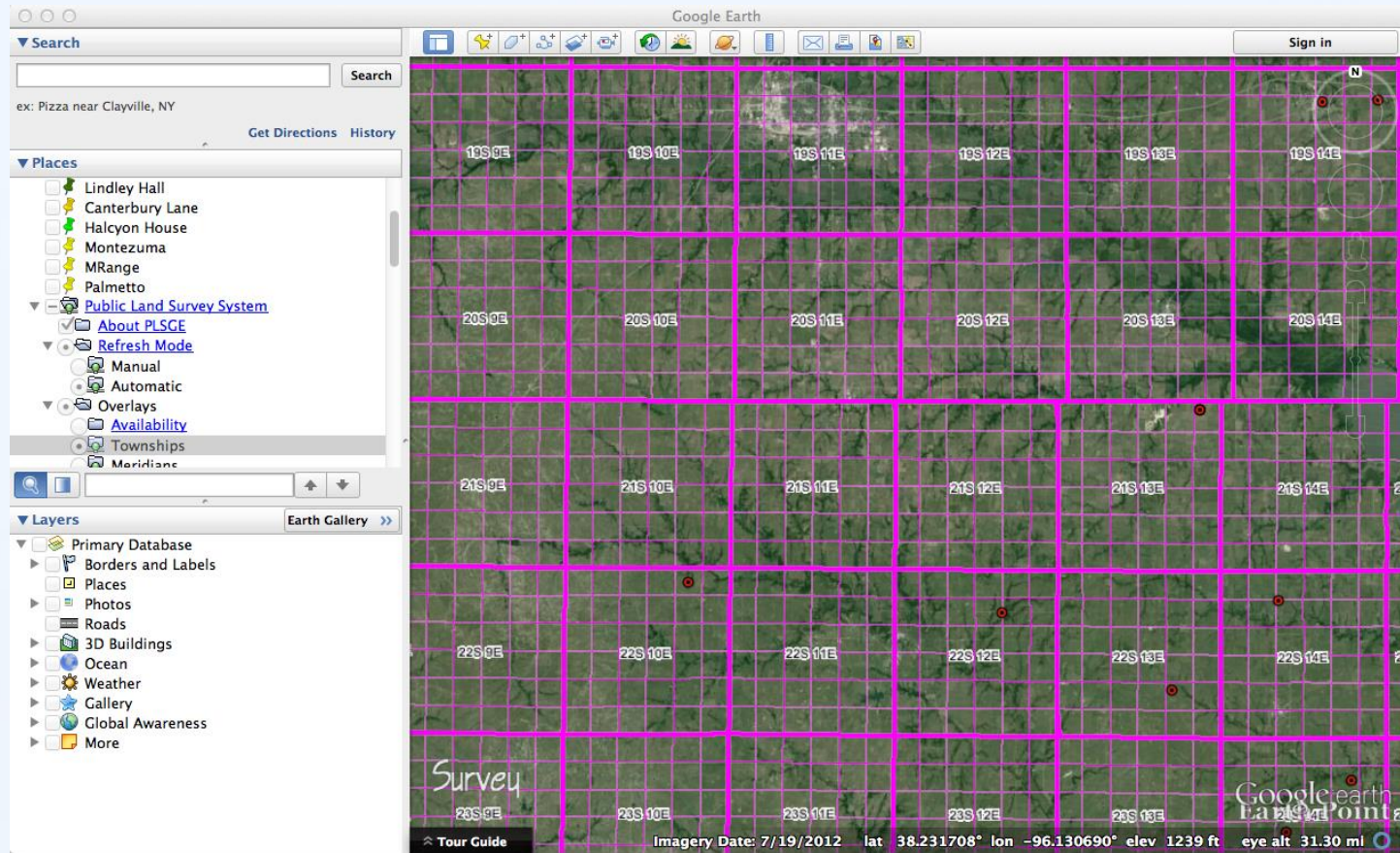
Outcrops,  
roadcuts, quarries

Geology

<http://mrdata.usgs.gov/geology/state/>

Township-Range-  
Section

<http://www.metzgerwillard.us/plss/plss.html>





automated georeferencing

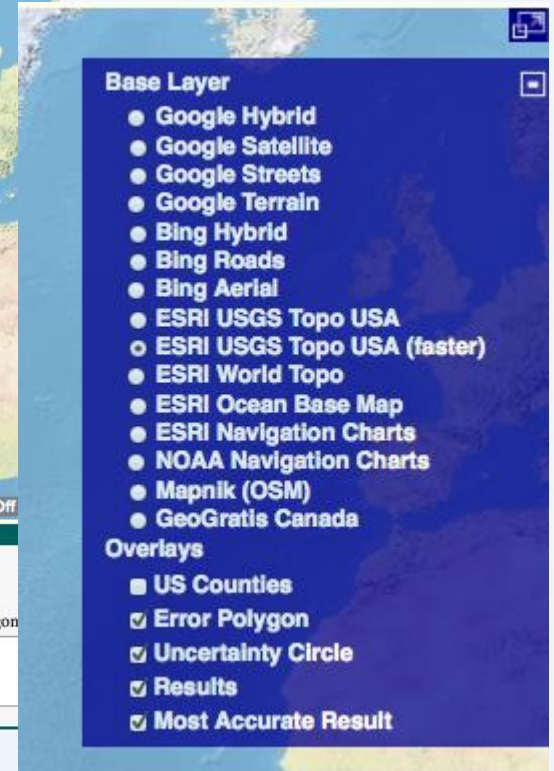
uncertainty determination

Home | Standalone App | Web Application | Collaborative Georeferencing | Developer Resources | Workshops | Support and Contacts

GEOLocate Web Application



geographic visualization



interoperability

kml export

batch processing

collaborative georeferencing

verification & correction

multi-lingual

# Typical GEOLocate

## Workflow

1 Data Entry & Preparation

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

2 Automated Processing

*Georeferencing Algorithm*

3 Manual Verification

*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...*

# Visualization: Base Layers



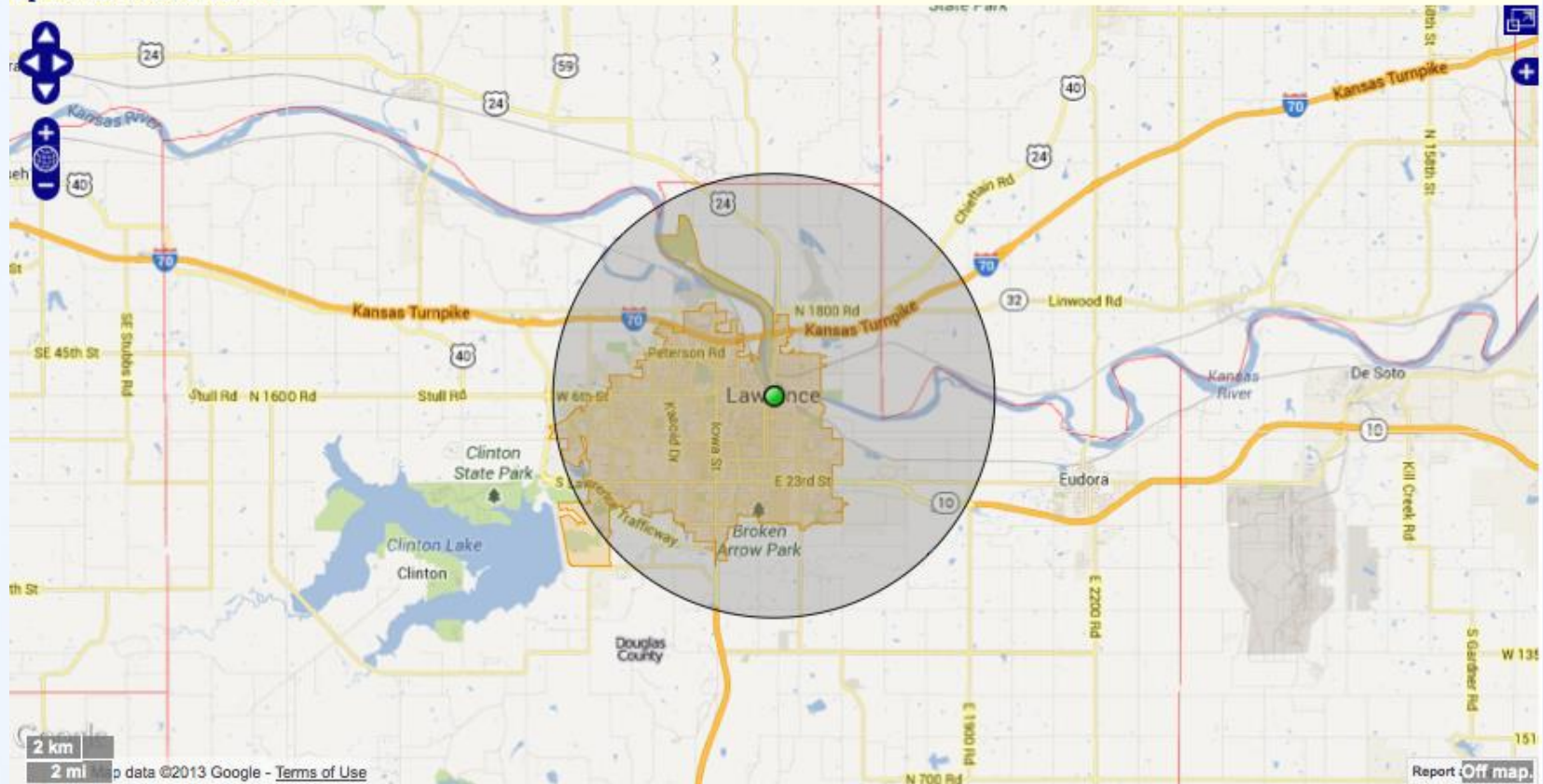
# Lawrence, Kansas

Home | Standalone App | Web Application | Collaborative Georeferencing | Developer Resources | Workshops | Support and Contacts

 **GEOLocate Web Application**



1 possible location found.



Workbench **1 possible location found**

Georeference | Options |   Draw polygon  Place marker  Measure

Locality String: Lawrence

Country: UNITED STATES OF AMERICA

State: KS

County:

latitude: 38.97167  longitude: -95.235  uncertainty: 8424 m  error polygon

```
38.97167 -95.235 8424
38.9295730196,-95.2791629241,38.9296920196,-95.2791759241,38.929921019
6,-95.2791969241,38.9302050196,-95.2792059241,38.9306130196,-95.279203
9241,38.9311370196,-95.2792009241,38.9311770196,-95.2792009241,38.93113
```

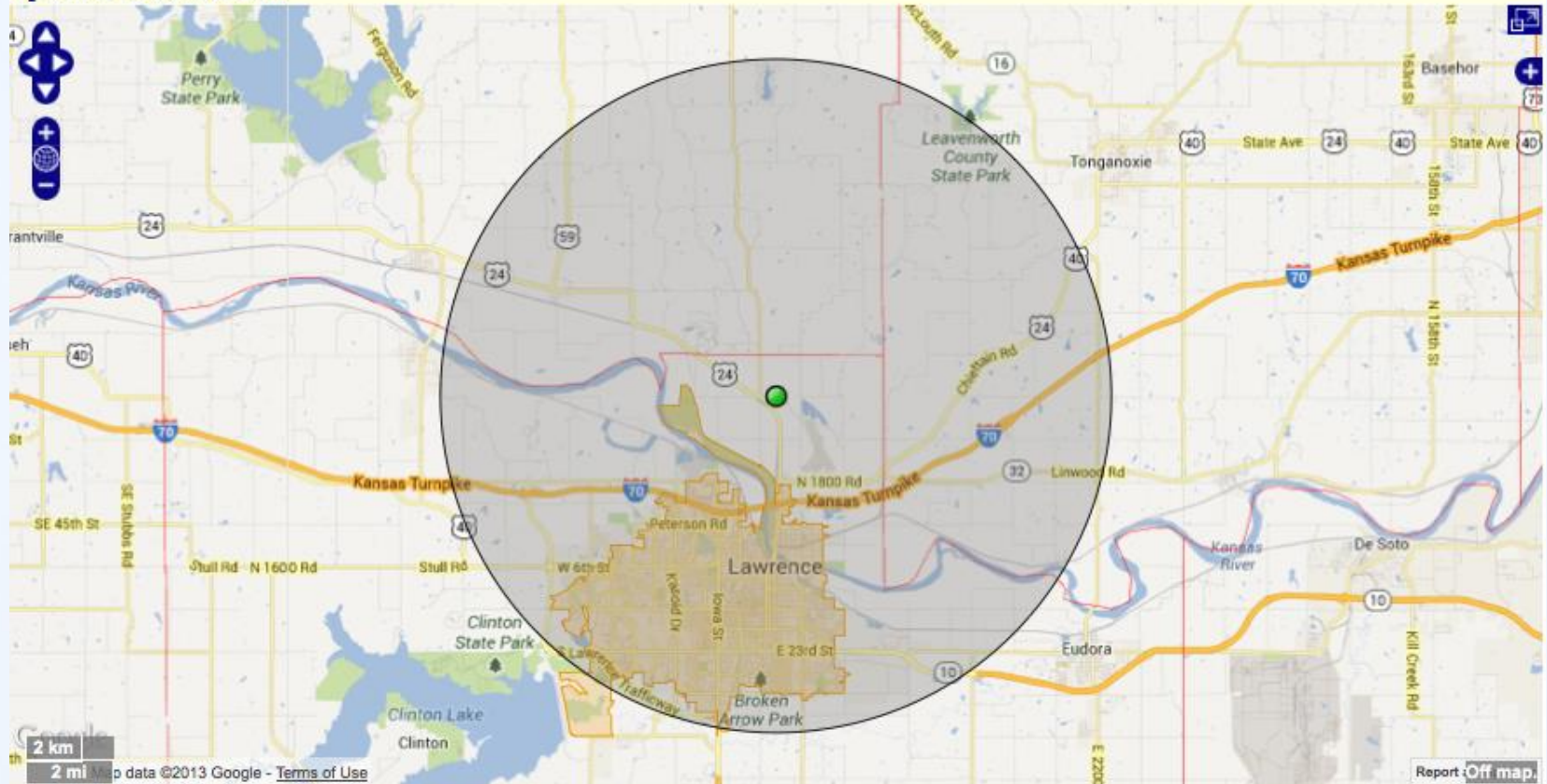
# 4 miles N of Lawrence, Kansas

Home | Standalone App | Web Application | Collaborative Georeferencing | Developer Resources | Workshops | Support and Contacts

 **GEOLocate Web Application**



1 possible location found.



Workbench

1 possible location found

Georeference

Options

Clear Polygon

Draw polygon

Place marker

Measure

Locality String: 4 miles N of Lawrence

Country: UNITED STATES OF AMERICA

State: KS

County:

latitude: 39.029656  longitude: -95.235  uncertainty: 12757 m  error polygon

```
39.029656 -95.235 12757
38.9295730196,-95.2791629241,38.9296920196,-95.2791759241,38.9299210196,-95.2791969241,38.9302050196,-95.2792059241,38.9306130196,-95.2792039241,38.9311370196,-95.2792009241,38.9311770196,-95.2792009241,38.9311370196,-95.2791629241,38.9295730196
```

# Common problems faced by paleo collections:

- Old towns which no longer exist.
- Railroad cuts which have belonged to seven different railroad companies over the years.
- Quarries which have changed names or closed.
- Localities based on “Mr. Smith’s” farm or other locally familiar terms (locally familiar at the time of collection and often not on published maps).



# GNIS - USGS Geographic Names Information System

- some historical names
- lots of feature names:
  - dams
  - topographic features
  - water bodies

Query Form For The United States And Its Territories

Feature Name:

Feature ID:

Exact Match  Exclude Variants

Feature Class:   
[Press Ctrl to select more than one.](#) [Definitions](#)

State:

County:

Elevation\*:  -

Feet  Meters

<http://geonames.usgs.gov/pls/gnispublic/>





# Getty Thesaurus of Geographic Names™

- good for historical names

## Getty Thesaurus of Geographic Names® Online

Search the TGN ? Help

Find Name or ID:

Place Type:

[Lookup](#)

Nation:

[Lookup](#)

[Pop-up Search](#)      [Browse the TGN hierarchies](#)

<http://www.getty.edu/research/tools/vocabularies/tgn/index.html>





science for a changing world

# Map Store

- downloadable pdfs of maps
- historical maps of different scales
- modern with satellite imagery
- geopdf toolbar for coordinates

## Map Locator & Downloader

[DOI Disclaimer on Google Maps API](#)

Don't see the Map Locator & Downloader? [Help](#) | Having trouble? Call: 1-888-ASK-USGS (1-888-275-8747, Select 0

Search:

Address or Place

Go

[\[Search Help\]](#)

or

The screenshot shows the USGS Map Store interface. At the top, there is a search bar with the text "new haven, ct" and a "Go" button. To the right of the search bar are links for "Address or Place", "Search Help", and "or". Below the search bar is a map of New Haven, CT, with a red pin indicating the location. A dropdown menu is open, showing a list of map products for download. The products are listed in a table with columns for "BUY", "Size", "Date", "View", and "DOWNLOAD".

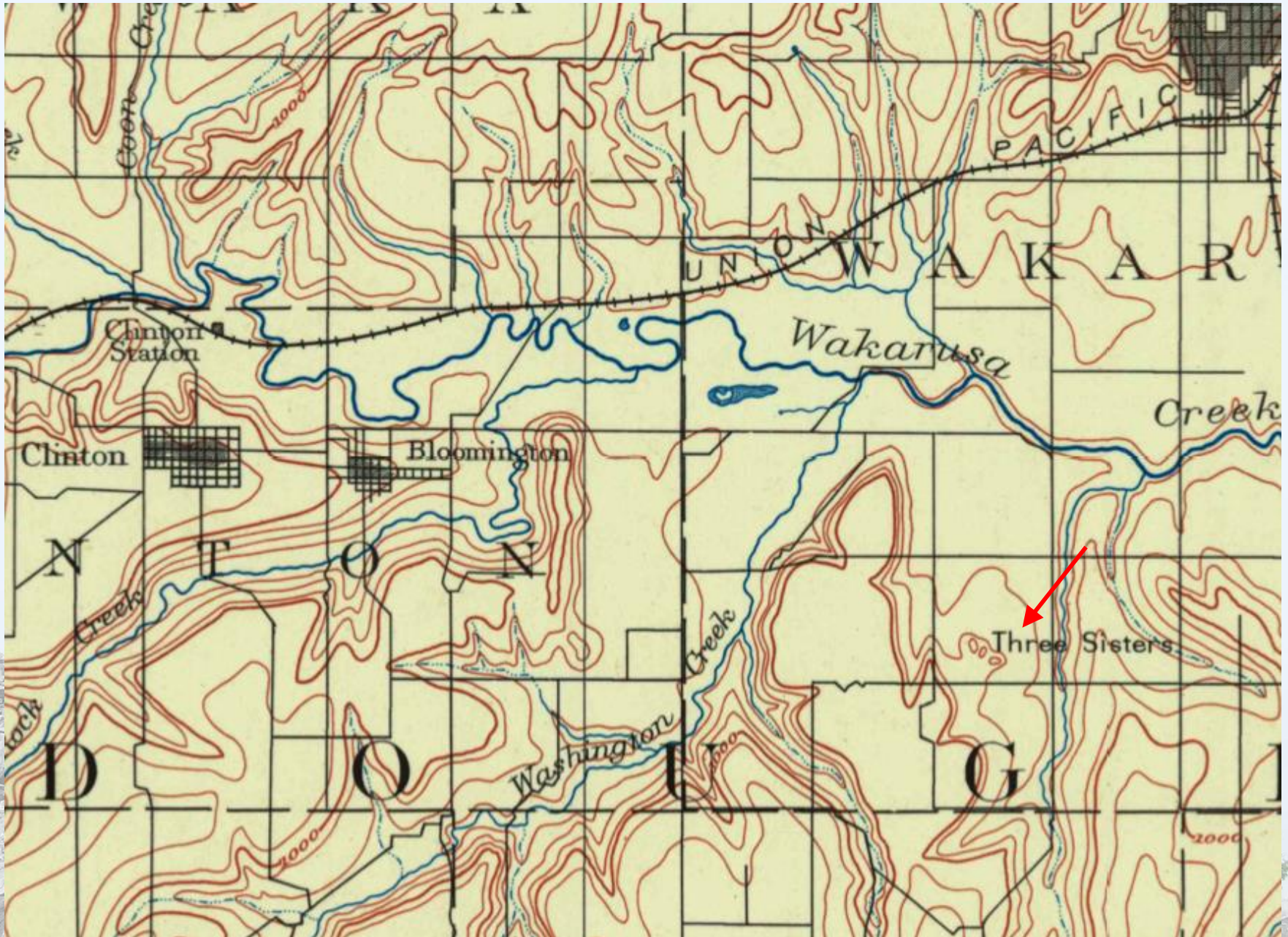
BUY	Size	Date	View	DOWNLOAD
<a href="#">New Haven US Topo</a>	7.5X7.5	2012	<a href="#">view</a>	<a href="#">23.8MB</a>
<a href="#">New Haven</a>	7.5X7.5	1967	<a href="#">view</a>	<a href="#">17.2MB</a>
<a href="#">New Haven</a>	7.5X7.5	1967	<a href="#">view</a>	<a href="#">19.4MB</a>
<a href="#">New Haven</a>	7.5X7.5	1967	<a href="#">view</a>	<a href="#">20.6MB</a>
<a href="#">New Haven</a>	7.5X7.5	1967	<a href="#">view</a>	<a href="#">18.9MB</a>
<a href="#">New Haven</a>	7.5X7.5	1967	<a href="#">view</a>	<a href="#">16.9MB</a>

At the bottom of the map, there is a coordinate box showing the coordinates: 41-29.959N, 72-42.657W. Below the coordinates is the text "USNG 18T XL 9105 9671". At the very bottom of the page, there is a link: <http://tinyurl.com/USGSMapStore>

<http://tinyurl.com/USGSMapStore>

Three Sisters, SW of Lawrence, KS?

USGS Topo map, Lawrence Quadrangle, 1885



# To find out more....

iDigBio Georeferencing working group:

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



For powerpoints, notes, resources:

[https://www.idigbio.org/wiki/index.php/GWG\\_Second\\_Train\\_the\\_Trainers\\_Workshop](https://www.idigbio.org/wiki/index.php/GWG_Second_Train_the_Trainers_Workshop)

For videos of talks:

<https://vimeo.com/album/2163673>



For links to Guide to Best Practices, Georeferencing calculator, Georeferencing Quick Guide:

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



Google Maps: [maps.google.com](https://maps.google.com)

Google Earth: <http://www.google.com/earth/index.html>

GEOLocate: <http://www.museum.tulane.edu/geolocate/>

ACME mapper: <http://mapper.acme.com/>

USGS map store, for historical map PDFs: <http://tinyurl.com/USGSMapStore>

Getty Thesaurus, useful for Old & alt. names: [bit.ly/Getty-TGN](http://bit.ly/Getty-TGN)

Geonames: <http://geonames.usgs.gov/pls/gnispublic/>

LatLong crosshairs for Google Maps: [www.canadensys.net/latlong-crosshairs](http://www.canadensys.net/latlong-crosshairs)

Hard to find localities: [www.fallingrain.com](http://www.fallingrain.com)

PLSS/TRS: [www.earthpoint.us](http://www.earthpoint.us), <http://www.metzgerwillard.us/plss/plss.html>

UTM map: [www.dmap.co.uk/utmworld.htm](http://www.dmap.co.uk/utmworld.htm)

UTM calculator: <http://www.earthpoint.us/Convert.aspx>



<http://www.openstreetmap.org/>: free, editable map of the world

<http://global.mapit.mysociety.org/>: can give you administrative boundaries if you have a lat-long (could be useful for verifying localities outside the US)

<http://www.oldmapsonline.org/> and <http://www.davidrumsey.com/>: historical maps

<http://www.gelib.com/ng-topo.htm>: topo map overlay for Google Earth

<http://mrddata.usgs.gov/geology/state/>: geology layers for Google Earth

<http://www.nrcan.gc.ca/earth-sciences/geography-boundary/geographical-name/11680>: natural resources of Canada

[http://www.lib.utexas.edu/maps/map\\_sites/cities\\_sites.html](http://www.lib.utexas.edu/maps/map_sites/cities_sites.html): Univ. of Texas at Austin Perry Castenada Map Library

<http://www.gadm.org/>: Global Administrative Boundaries Database

<http://www.geody.com/>: Google Earth wiki

<http://www.naturalearthdata.com/>: for making maps

<http://geo.data.gov/geoportal/catalog/main/home.page>: various geographical datasets

<http://maps.nls.uk/>: historical maps of Scotland/Great Britain/Ireland and Belgium

<http://cidades-brasil-ibge-google-earth.softonic.com.br/download>: cities in Brazil for google

<http://www.topomapper.com/>: compare topo and google maps side by side.

<http://www.bl.uk/maps/>, <http://www.bl.uk/maps/georeferencingmap.html>: British Library crowdsource project for historical maps

<http://www.uwgb.edu/dutchs/usefuldata/ConvertUTMNoOZ.HTM>: UTM and other useful resources