Fundamentals of Georeferencing Using GEOLocate

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

latitude: 30.88797
longitude: -89.83601
uncertainty radius: 48m
uncertainty polygon: 30.88823,-89.83641, 30.88815,-89.83634, 30.88808,-89.83622...
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 web services.

Desktop Application
The original standalone desktop application.

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

Copyright© 2014

Tulane University
NSF
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...
1 possible location found.

**Locality String:** TALLAPOOSA RIVER 5.0 mi. N of Buchanan, Hwy 27

**Country:** UNITED STATES OF AMERICA

**State:** Georgia

**County:** Haralson
Algorithm Performance

Versus Known U. S. Localities

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>% Found</th>
<th>Mean Dist. Off</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEOLocate</td>
<td>95%</td>
<td>6.1 km</td>
<td>2.1 km</td>
</tr>
</tbody>
</table>

Map showing a circular buffer around Riverton with a radius of 6 km.
## Algorithm Performance

### Versus Known Australian Localities

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>% Found</th>
<th>Mean Dist. Off</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>GEO Locate</td>
<td>86 → 97%</td>
<td>796 → 218 km</td>
<td>154 → 48 km</td>
</tr>
</tbody>
</table>

![Map showing distances](image-url)
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
Named Creek 7 miles North of Named Place at Hwy 1

Match Water Body

Detect Hwy/River Crossing

Do Error Polygon

Displace Polygon

Restrict to Lowest Adm. Unit

Language: English
<table>
<thead>
<tr>
<th>Location</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lawrence, KS</td>
<td>15 mi N of Lawrence, KS no displacement</td>
</tr>
<tr>
<td>15 mi N of Lawrence, KS</td>
<td>15 mi N of Lawrence, KS displaced polygon</td>
</tr>
</tbody>
</table>

Generating Polygons:
Visualization: Base Layers
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](#)
Evaluating Results

Beaver Creek; LA; USA

16 possible locations found.
Batch Georeferencing

<table>
<thead>
<tr>
<th>Locality</th>
<th>Country</th>
<th>StateProvince</th>
<th>County</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Corrected</th>
<th>precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chambers Spring Road 2.5 km S of Hwy 412, 8.0 km E of Siloam Springs, T17N, R33W,</td>
<td>USA</td>
<td>Arkansas</td>
<td>Benton</td>
<td>36.186027</td>
<td>-94.451005</td>
<td>no</td>
<td>High(89)</td>
</tr>
<tr>
<td>Osage Creek, 1.0 mile N on gravel road to bridge crossing; gravel road jcts with</td>
<td>USA</td>
<td>Arkansas</td>
<td>Benton</td>
<td>36.189077</td>
<td>-94.395375</td>
<td>no</td>
<td>High(87)</td>
</tr>
<tr>
<td>Yocum Creek, near Oak Grove (Pass 11a), Sec. 30</td>
<td>USA</td>
<td>Arkansas</td>
<td>Carroll</td>
<td>36.454896</td>
<td>-93.322008</td>
<td>no</td>
<td>Low(35)</td>
</tr>
<tr>
<td>Village Creek State Park, S of driving range, Sec. 6</td>
<td>USA</td>
<td>Arkansas</td>
<td>Cross</td>
<td>35.16111</td>
<td>-90.70833</td>
<td>no</td>
<td>Low(39)</td>
</tr>
<tr>
<td>Sugar Creek, Hwy 163 at Bay Village, Sec. 4</td>
<td>USA</td>
<td>Arkansas</td>
<td>Cross</td>
<td>35.44909</td>
<td>-90.67533</td>
<td>no</td>
<td>High(100)</td>
</tr>
<tr>
<td>Buck Creek, 8.0 miles SE Corydon</td>
<td>USA</td>
<td>Indiana</td>
<td>Harrison</td>
<td>38.155118</td>
<td>-86.014724</td>
<td>no</td>
<td>High(88)</td>
</tr>
<tr>
<td>E Branch Mill Creek, Hessdale Road, 4.0 km S of Allendorph, Sec. 36</td>
<td>USA</td>
<td>Kansas</td>
<td>Wabaunsee</td>
<td>39.003564</td>
<td>-96.277745</td>
<td>no</td>
<td>High(88)</td>
</tr>
<tr>
<td>Blissdale Creek, Hillside National Wildlife Refuge, 500 m SW of Blissdale on Blis</td>
<td>USA</td>
<td>Mississippi</td>
<td>Holmes</td>
<td>33.083754</td>
<td>-90.224033</td>
<td>no</td>
<td>High(84)</td>
</tr>
</tbody>
</table>
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

<table>
<thead>
<tr>
<th>Community: FishNet 2</th>
<th>History</th>
<th>Review</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Luggar Landing, West Pearl River, Starlding; United States; Louisiana;</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>FMNH14004 Fundulus chrysotus</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **1 possible location found**

- **History**
  - Luggar Landing, West Pearl River near Starlding; United States; Louisiana;
  - West Pearl River, Starlding, Luggar Landing, Deer Island; United States; Louisiana;
  - Luggar Landing, West Pearl River, near Starlding; United States; Louisiana;

- **Review**
  - Clear Polygon
  - Draw polygon
  - Place

- **Similar Records(5)**
  - Luggar Landing, West Pearl River near Starlding; United States; Louisiana;
  - West Pearl River, Starlding, Luggar Landing, Deer Island; United States; Louisiana;
  - Luggar Landing, West Pearl River, near Starlding; United States; Louisiana;
Organizational Units

Community

Data sources

Community

Data sources

Community

Data sources

Community

Users
Data source management operations

- Add new community data source via CSV files

Click on an item’s header to expand/collapse its content.

<table>
<thead>
<tr>
<th>Community-wide</th>
<th>you</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community Owner:</td>
<td>you</td>
</tr>
<tr>
<td>Community description:</td>
<td>Collaborative georeferencing of data from FishNet 2</td>
</tr>
<tr>
<td>Number of community data sources:</td>
<td>168</td>
</tr>
<tr>
<td>Records statistics:</td>
<td></td>
</tr>
<tr>
<td>Specimens: 1,302,853</td>
<td>Corrected: 1,122,421</td>
</tr>
<tr>
<td>Skipped: 122,175</td>
<td>Total processed: 1,244,596</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Iowa_with_county</th>
<th>date added: Monday, January 07, 2013</th>
<th>delete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maine_with_county</td>
<td>date added: Monday, January 07, 2013</td>
<td>delete</td>
</tr>
<tr>
<td>Minnesota_with_county</td>
<td>date added: Monday, January 07, 2013</td>
<td>delete</td>
</tr>
<tr>
<td>Montana_with_county</td>
<td>date added: Monday, January 07, 2013</td>
<td>delete</td>
</tr>
<tr>
<td>Nevada_with_county</td>
<td>date added: Monday, January 07, 2013</td>
<td>delete</td>
</tr>
</tbody>
</table>
Georeferencing Communities

<table>
<thead>
<tr>
<th>Communities List</th>
<th>Create Communities</th>
<th>Join Communities</th>
<th>Account Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fill out the form below to create a new community.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fields marked * are required.</td>
<td></td>
<td>Community name: New Community</td>
<td></td>
</tr>
<tr>
<td>Description: Community creation Demo.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Georeferencing options:</td>
<td></td>
<td>Water body matching</td>
<td>Look for hwyriver crossing</td>
</tr>
<tr>
<td>Privacy setting:</td>
<td></td>
<td>public: any user can join your community.</td>
<td>restricted: users send you requests to join your community.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>private: only users you invite can join your community.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>RESET CREATE COMMUNITY</td>
<td></td>
</tr>
</tbody>
</table>

Create Communities

Data Sources

<table>
<thead>
<tr>
<th>Data Sources</th>
<th>Members</th>
<th>Com Center</th>
<th>Settings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data source management operations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Add new community data source via DiGIR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Add new community data source via CSV files</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Add New Users & Link Out

Use this easy signup link to send a auto-join invite to users:

| Data source management operations | | | |
| Add new community data source via DiGIR | | | |
| Add new community data source via CSV files | | | |

Enable hyperlinks on data source records:
base link format: "http://domain.com?<key>=" where <key> is a parameter key of your choice.

| Data source management operations | | | |
| Add new community data source via DiGIR | | | |
| Add new community data source via CSV files | | | |

GENERATE EASY SIGNUP LINK
ADD SPECIMEN HYPERLINK
RESET SAVE
Monitoring & Managing Progress
1 possible location found.

**Luggar Landing, West Pearl River, Starliding; United States; Louisiana;**

**FMNH14004 Fundulus chrysotus**
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 PLUMBES
- UWFCUW 002715 ONCORHYNCHUS CLARKII

Calculated Coordinates
Lat: 48.296377
Lon: -112.253363
U. Radius: 5214 m

Logout | logged in as nelson

Map data ©2014 Google | Terms of Use | Report a map error
Lake Winnipesaukee, Fish Cove near mouth of T 108, 3.5 mi ESE of Meredith, Merrimack Watershed, elev. 50.
Utilizing and Reviewing Past Work

ON WORKBENCH:
Lake Winnipesaukee in Blackey Cove & outlet Lake Kanasatka, Merrimack Watershed, 0.75 mi E of Center Harbor, elev 504; United States; New Hampshire; Carroll;
Interoperability

• Webservices
  – SOAP
  – JSON
  – GeoJSON

• Web Client

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>
<table>
<thead>
<tr>
<th>Loc</th>
<th>Locality/String</th>
<th>Country</th>
<th>Country</th>
<th>State</th>
<th>County</th>
<th>Lat</th>
<th>Lon</th>
<th>Latitude uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Cable Station at Cape York</td>
<td>Australia</td>
<td>Queensland</td>
<td></td>
<td></td>
<td>-13.8385</td>
<td>142.53167</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Aberdeen Lake</td>
<td>Canada</td>
<td>Queensland</td>
<td></td>
<td></td>
<td>64.48667</td>
<td>-99.00000</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Pioneer Park near Davis</td>
<td>USA</td>
<td>CA</td>
<td>Yolo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Conroe</td>
<td>USA</td>
<td>TX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Twin Creek at US Hwy 24</td>
<td>USA</td>
<td>CO</td>
<td>Teller</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Turkey Creek at Goodwin Road.</td>
<td>USA</td>
<td>AL</td>
<td>Jefferson</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Borden Creek at Forest Road 224</td>
<td>USA</td>
<td>AL</td>
<td>Lawrence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Moon Lake</td>
<td>USA</td>
<td>MS</td>
<td>Coahoma</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Red River along right bank at River Mile 105; Station 1.</td>
<td>USA</td>
<td>LA</td>
<td>Rapides</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Bienville National Forest on US Hwy 20</td>
<td>USA</td>
<td>MS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Zion National Park</td>
<td>USA</td>
<td>UT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Whitehall state wildlife management area</td>
<td>USA</td>
<td>NE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Lincoln Beach</td>
<td>USA</td>
<td>OR</td>
<td>Lincoln</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Green River near neatisville</td>
<td>USA</td>
<td>KY</td>
<td>Adair</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>New Milan Reservoir, 1.75 miles north of milan</td>
<td>USA</td>
<td>MO</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Homochitto River, 1.5 mi. S of Buda, Hwy. 98,</td>
<td>USA</td>
<td>MS</td>
<td>Franklin</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Homochitto River 5 mi. E of Union Church, MS Hwy. 550.</td>
<td>USA</td>
<td>MS</td>
<td>Lincoln</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Mamala Bay</td>
<td>USA</td>
<td>HI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Bay of bengal</td>
<td>India</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Henley On Thames</td>
<td>United Kingdom</td>
<td>Oxfordshire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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=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 geocode 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+
Arctos
## Arizona State University Vascular Plant Herbarium

### Return Count: 39

- Pima County, Growing on sanitary landfill, Tumamoc Hill, Tucson [1]
- Pima County, N slope above laboratory ground, Tumamoc Hill [1]
- Pima County, North slope above laboratory grounds, Tumamoc Hill, Tucson [1]
- Pima County, North slope above laboratory grounds, Tumamoc Hill, Tucson [1]
- **Pima County, S side of Tumamoc Hill on Desert Laboratory property, TRS: T14S R13E S15 [1]**
  
<table>
<thead>
<tr>
<th>Deg. Min.</th>
<th>Sec.</th>
<th>Decimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>111.80</td>
<td>45</td>
<td>0.50</td>
</tr>
<tr>
<td>31.71</td>
<td>30</td>
<td>0.50</td>
</tr>
</tbody>
</table>

### Sources:

egbot 2012-04-01

### Verification Status:
reviewed - high confidence

---

**Statistics**

- Records to be Georeferenced: 8317
- Total Number: 39723
- Total Percentage: 21.6%
<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Locality</td>
<td></td>
<td>Country</td>
<td>Link to GEOLocate Web App</td>
</tr>
<tr>
<td>2</td>
<td>Bogalusa</td>
<td>USA</td>
<td>LA</td>
<td>Washington</td>
</tr>
</tbody>
</table>