Using GEOLocate for Collaborative Georeferencing

By Eva M. Biedron and Nicholas A. Famoso

Table of Contents

University of Oregon
Museum of Natural and Cultural History

3 December 2016

This is a collection of workflows for using the GEOLocate's online collaborative georeferencing portal developed by Nelson Rios <nrios@tulane.edu> at the Tulane University Biodiversity Institute. Product produced for the EPICC TCN.

Create a profile

1. Navigate to <http://www.museum.tulane.edu/coge>.
2. Click 'Sign Up'
3. Fill in your information and click 'create account'

NOTE:
- You will need to be approved before your account is active and can access a community. It may take up to a day to be approved. You will be able to be added to a community before approval.

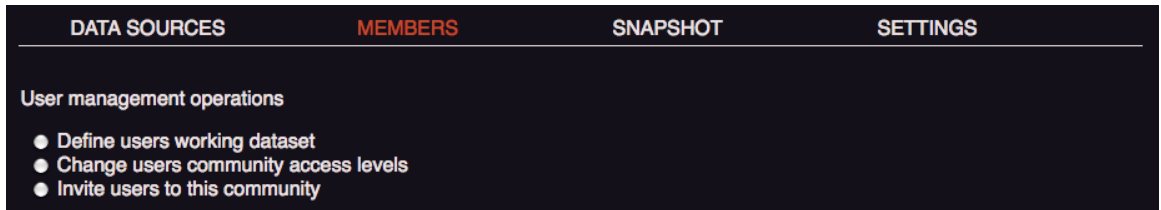**Creating a community**

1. Navigate to <http://www.museum.tulane.edu/coge>. Log in. You will be directed to your 'Dashboard'.
2. Click on the 'Create' tab. You will be redirected to a community creation page.
3. Name your community, update the description, set georeferencing options, and examine the privacy settings. At least 'Restricted' is recommended.
   a. It has been suggested that if most of your locality data are cities or counties, then you should leave "Do Error Polygon" checked.
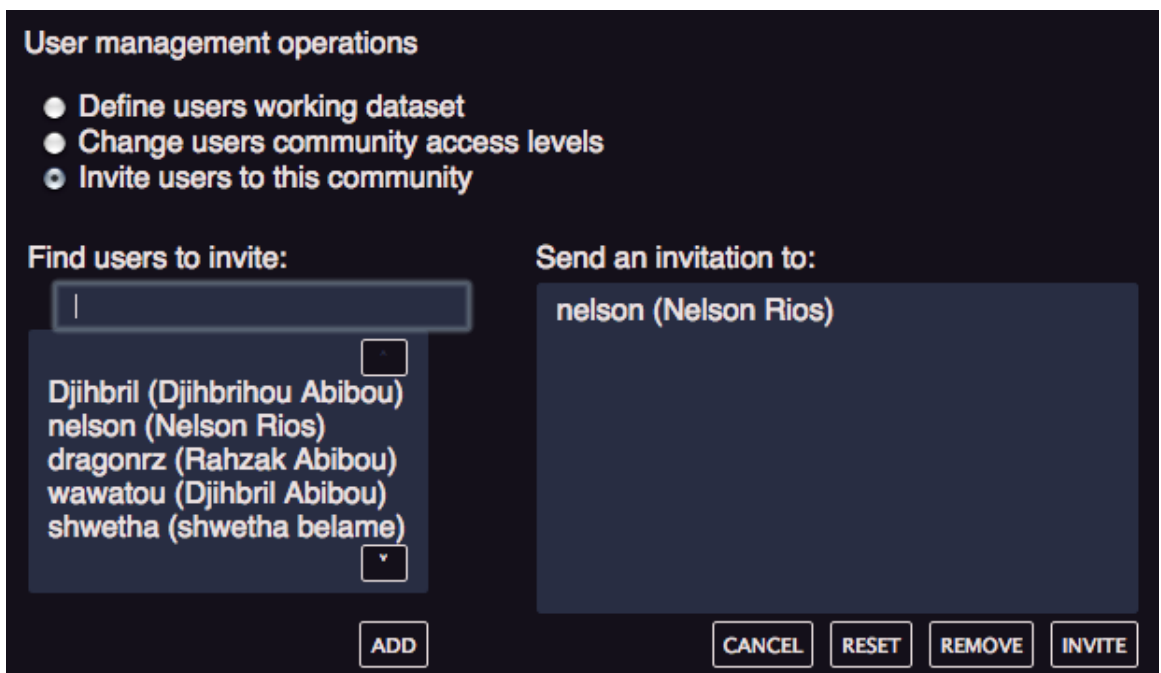4. Click the 'Create Community' button.

NOTE:
- Have the project lead create the community. They will have Head Administrative powers that outmatch the powers of those they designate Administrators themselves.
- You can change the settings at anytime through the settings tab for the community.

Adding community members

1. Navigate to <http://www.museum.tulane.edu/coge>. Log in. You will be directed to your 'Dashboard'.
2. In the 'Dashboard', under the 'My Communities' tab, click your community to highlight it, and click the 'Go to community' button. You will be directed to your community page.
3. Click on the 'Members' tab. You will be directed to the member management interface.
4. If you have set the privacy setting to 'Private' during community creation, follow the steps below:
   a. Under 'User Management Operations', the Administrator will click the third option 'Invite users to this community'.

b. Under 'Find users to invite', search for a specific prospective community member's name. Click their name to select it. Click the add button.

c. Repeat Step 4.b until all prospective community members have been added to the 'Send an invitation to' list.

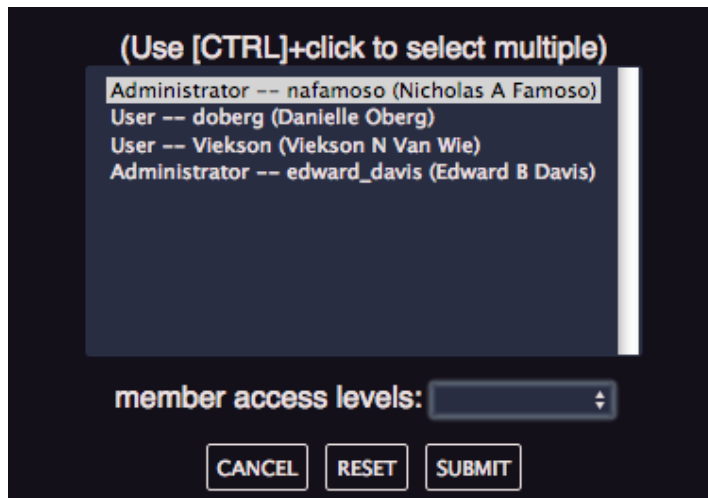d. Click the 'Invite' button. An invitation will be sent to those members you invited.

5. If you have set the privacy setting to 'Restricted' during community creation, you may follow the 'add member' procedure listed above in Step 4 or members may request inclusion into the group as follows below:

a. From the 'Dashboard', the prospective member will click on the 'Join' tab.

b. Scroll until you see the community you wish to join. Click on it to expand the entry.

c. Scroll to the bottom of the expanded entry and in the right-hand corner will be a button labeled 'Join'. Clicking it will send a request to join.

d. The Administrator will have to approve the request to join.

NOTE:
- You can also create an easy signup link to send an auto-join invite to users through the community settings tab.

Changing community member privileges

1. Navigate to <http://www.museum.tulane.edu/coge>. Log in. You will be directed to your 'Dashboard'.
2. In the 'Dashboard', under the 'My Communities' tab, click your community to highlight it, and click the 'Go to community' button. You will be directed to your community page.
3. Click on the 'Members' tab. You will be directed to the member management interface.
4. Under 'User Management Operations', the Administrator will click the second option 'Change users community access levels'.
5. Select one or more members.
    a. Yes, you can change user access levels in bulk!
6. Select the desired member access level from the drop down list. Hit 'Submit'.

NOTE:
- Members can only change the member access levels of members with lower access levels then themselves.
- "User" is the basic level of access. They can only log into the portal and see the data. Just for members who are strictly georeferencing.
- "Reviewer" can look at all data in the community with some bulk viewing tools.
- "Administrator" has the same level of access as the owner. You can see the statistics for each member and for the data source, change people's access level, add new data, etc.

Creating data

In order to add data, you will need to create a CSV file with all of the localities you want to georeference. The column names, or fields, do not have to match Darwin Core (DwC) standards exactly, but there are fields that you need to have. GEOLocate can support specimen data in the collaborative portal, but you do not need these data if you are just looking at localities. The fields need to be there but you can place some dummy data in there to make the portal happy.

Required fields:
- CatalogNumber
  - UO has been using the locality number here since we are just using localities. However, this field should be unique for each locality.
  - If you do have specimen data, you will be able to see which specimens are associated with each locality.
- ScientificName
  - Note: You can put "NA" or "No Name" in here, or something similar, if you are just doing localities.
- Locality
  - This is the locality string that GEOLocate will use to figure out where your locality is from. Make sure to put the Township Range and Section (TRS) in here (formatted like, T12S R17E Sec9) if you want to use those data in finding your locality.
  - Notes on TRS
    - In the event of a state with multiple meridians, GEOLocate would either return more than one result or the result from the meridian with the largest area, if a county is not specified.

Basically required field if you want GEOLocate to work properly. Those in bold are needed if you want GEOLocate to find your locality:
- **Country**
- **StateProvince**
- **County**
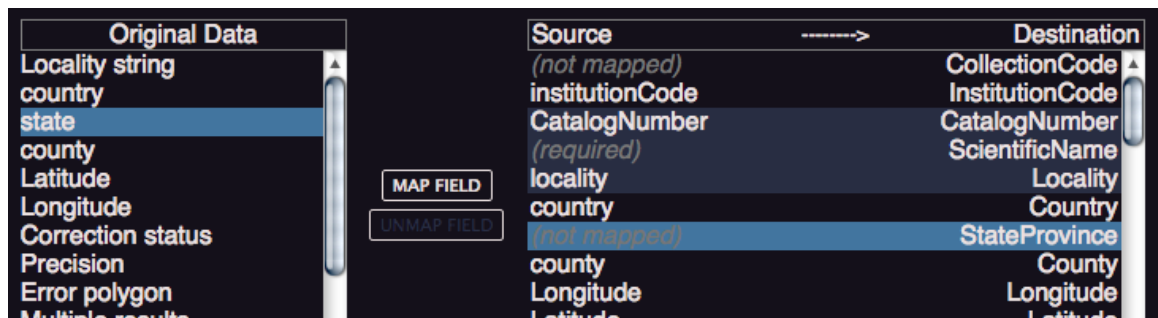- CollectionCode
- InstitutionCode

You can add any other fields you want if you thing they will be helpful.

Uploading the data

1. Navigate to <http://www.museum.tulane.edu/coge>. Log in. You will be directed to your 'Dashboard'.
2. In the 'Dashboard', under the 'My Communities' tab, click your community to highlight it, and click the 'Go to community' button. You will be directed to your community page.

3. In your community page, under the 'Data Sources' tab, click 'Add new community data source via csv files'. A 'New CSV data source addition' window will open.
4. Add a data source identifier (*e.g.*, EPICC TCN Data) and a description (if desired). Make sure 'File encoding' is set to Unicode (UTF-8).
    a. UTF-8 works for most North American data, but if you have localities from countries with languages that use diacritical marks or a different alphabet, you should set this to Windows 1252. For EPICC, this means anything from Central and South America.
5. In the lower right hand corner of the window, click 'Select file'. A file-select window will open if you have Flash installed in your browser.
6. Choose the data file (saved in .csv format) you want to upload. Click 'Open'. A progress bar will open and fill as the file uploads (takes ~5 seconds). You will be automatically directed to a 'mapping' interface.
7. You will be directed to map the columns in your spreadsheet to standard columns in the GeoLocate CoGe interface. Under the 'Source' column (on the center-right), it will be indicated if the column is '(required)' to be mapped or not required ('(not mapped)'). Many columns will remain unmapped.

To map columns, highlight the 'Original Data' column (on the left) and the 'Destination' column (on the far right) and click the 'Map Field' button. The columns are now mapped.
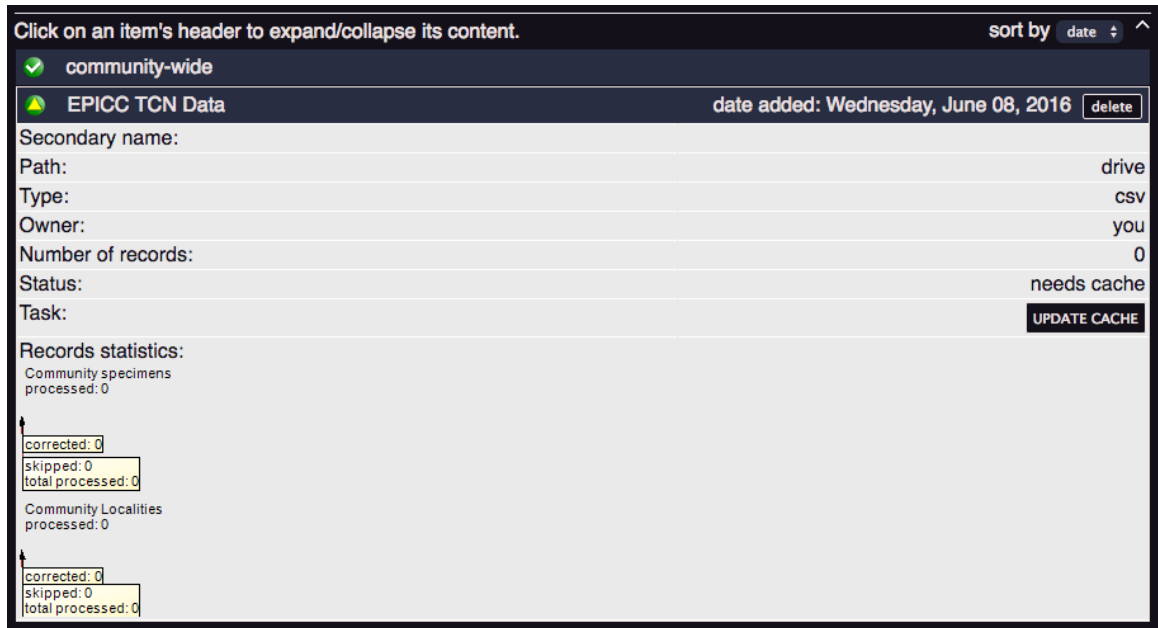

NOTE:
- Destination 'Locality' does not have a character limit, but Destination 'Notes' has a 255 character limits (which includes spaces)
- Map Original Data 'Locality string' to the Destination 'Locality' NOT Destination 'Notes'

When you have mapped your columns to your hearts desire, click the 'Save Data Source' button. The window will close, revealing the 'Data Sources' tab again.

8. Your newly-uploaded data set will appear with a 'triangle in a circle' icon. Click on the data set once to expand it. Click the 'Update Cache' button.

5

This may take several minutes.

NOTE:
- GEOLocate is automatically georeferencing your localities after you click the 'Update Cache' button. This happens in the background, so you do not need to have your browser open. This is also the point where GEOLocate is assigning similarities between records and separating locality from specimen data.
- At this point you may receive errors, warnings, or notifications, like
  - You mapped one Original Data to two Destination columns
  - Your data may have been truncated (because you put 255+ character strings into the Destination 'Notes' column, for example).
  - Your data may have false duplicates (because two field numbers may be the same or you may have replaced multiple missing field numbers with 'na' in your data sheet, for example).
- You will also be told how many of your localities are duplicates. You can check and see why GEOLocate thought they were duplicates.
- Fix the errors in your spreadsheet and repeat the upload process from Step 3. Repeat as necessary until no records are skipped.

9. Click the 'Save' button. Again, this may take several minutes, but it will probably take several less minutes than the 'Update Cache' step.
10. High-five! You did it! Now you can begin georeferencing.

Assigning specific data sources to a community member

It is recommended that you add multiple data sources to your community and assign a member of your community to each of them. This helps prevent members from georeferencing the same locality multiple times.

1. Navigate to <http://www.museum.tulane.edu/coge>. Log in. You will be directed to your 'Dashboard'.
2. In the 'Dashboard', under the 'My Communities' tab, click your community to highlight it, and click the 'Go to community' button. You will be directed to your community page.
3. Click on the 'Members' tab. You will be directed to the member management interface.
4. Under 'User Management Operations', the Administrator will click the first option 'Define users working dataset'.
5. Select the dataset you want for that member.

What if I only have one data source?

1. Navigate to <http://www.museum.tulane.edu/coge>. Log in. You will be directed to your 'Dashboard'.
2. In the 'Dashboard', under the 'My Communities' tab, click your community to highlight it, and click the 'Go to community' button. You will be directed to your community page.
3. Click on the 'Members' tab. You will be directed to the member management interface.
4. Under 'User Management Operations', the Administrator will click the first option 'Define users working dataset'.
5. Select the dataset, Country, and County, as well as the member you would like to assign the subset of the dataset.

NOTE:
- You can technically assign two people to the same chunk of data. However, the point of assigning the data to separate people is to allow A) multiple people to work on the data at the same time (like a georeferencing party! Or, you know, training.) and B) for continuity within each member's work. If you do assign two people to the same chunk of data, they cannot work at the same time or the collaborative georeferencing portal will give them the same localities and they will repeat the work (in my personal experience).

6. Click the 'Submit Changes' button.

NOTE:
- You will want to do this when you have one data source that you want to split up among multiple users. UO has been doing it this way and assigning specific counties to each member, and updating the assigned counties as they are completed.

Georeferencing a locality

1. Navigate to
   <http://www.museum.tulane.edu/geolocate/web/webcomgeoref.aspx>. Log in.
   Select your community. Click the 'Continue' button.
   a. Note that there is an "https://" (more secure) version if you are worried
      about your password being stolen.
   b. The "https://" version does not have as many map layers.
2. On the lower-left hand button menu, select 'Next Record'.
3. One record will appear along with any identical or similar records. If you wish to
   correct the record AND its identical matches or similar records, you must check
   the boxes next to them as well.
   a. To add a comment to a record, click 'Add comments'. A pop-up will open
      for your comments.
   b. To skip a record, click 'Skip selected' and comment why you skipped it.
4. The green dot that appears is the best guess as to where your locality is.
   a. This is not always right.
   b. If you already have Lat/Lon data in your data source, this point will be the
      green dot. When you click on the green dot, the parse pattern will say it is
      "Legacy result."
5. You can drag the green dot on the map around to move the location or you can
   manually change the point using the "Manual Edit" in the lower right corner.
6. There are also three bubbles on top of your workbench for drawing the error
   polygon, placing the green marker, and measuring.

7. Clicking on the green dot will open a small information bubble, in which there is an option to 'Edit uncertainty'. Click the text and a grey arrow will appear on the green dot. Drag the arrow to expand or contract the uncertainty radius.
8. You may also double-click on the title of the locality and it will bring up the "Locality re-georeferenceing interface." You can alter the locality description, county, state/province, and county. There are also some georeferencing options that you can change. This is a good option if GEOLocate has put the locality in a nonsensical location or if it could not georeference it at all.
    a. This is an important step if you find that a critical name of a locality or near-by place is misspelled.
    b. You can also enter the TRS of a locality or edit it here so that it fits the appropriate format (*e.g.*, T10S R3E Sec3). Make sure the TRS in in the locality string.
9. When you have edited the location and uncertainty radius, hit the 'Correct' button on the bottom left-hand menu. The locality or checked localities have been georeferenced.
10. Repeat from Step 2.

NOTE:

- You can also skip a locality if you click the 'Skip Selected' button. UO use this option for localities with legacy lat/lon data that have already been georeferenced using a map with localities plotted on them or those that have already been georeferenced using GEOLocate by another institution.
- You can see your georeferencing history using the 'History' tab at the top of the work bench. You can use this option to find a locality you already did to help you figure out a new locality (the point will be in yellow). You can also edit this record if you made a mistake originally; you will need to use the "move to workbench" button on the lower right to do this.
- The 'Review' tab is available only for Reviewers and Administrators. From here, you can review a wide variety of localities that have already been done to see how users have done. You can pick the data source, the user, county, state, country, and specifically look at corrections or skips. You can also change the number of localities that will be mapped (1, 100, or 1000). Points will show up as red dots. Click on a locality dot. You can then double click on the description to see who corrected it, when they did it, as well as several other pieces of data. If you want to fix it, click the "move to workbench" button on the lower right.
- When you are looking at your map, there are some options in the upper right corner. The top button will expand the map to fill your screen. The plus button will give you base maps and map overlay options, like a geologic map!

Evaluating progress

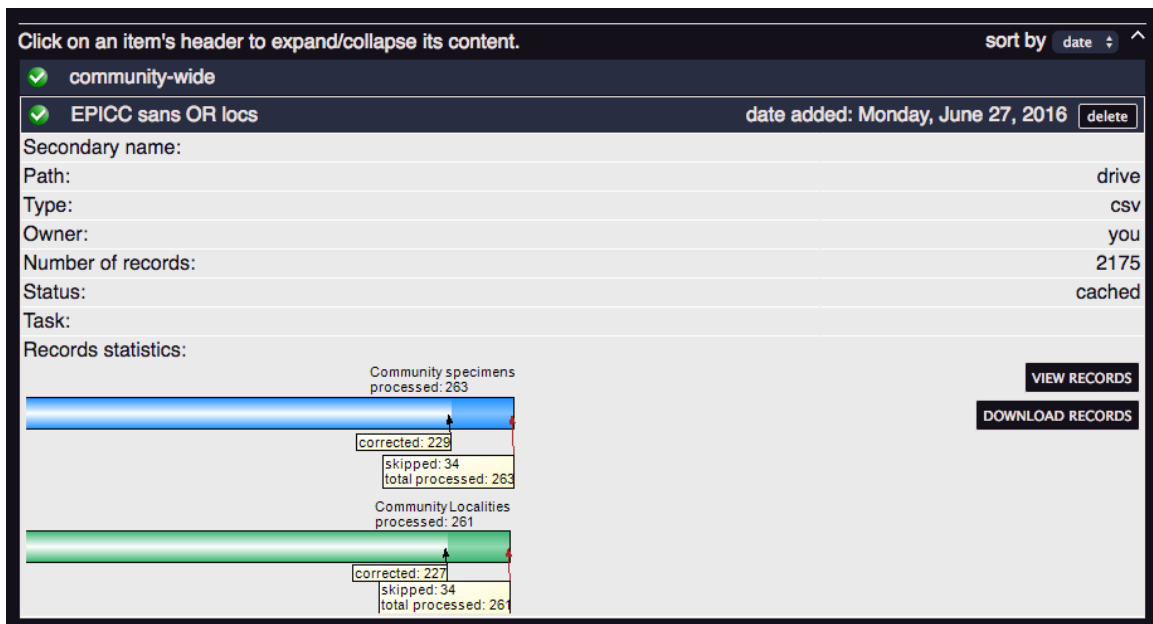There are several ways to evaluate community-wide and member progress.

Method 1 (the data source):

1. Navigate to <http://www.museum.tulane.edu/coge>. Log in. You will be directed to your 'Dashboard'.
2. In the 'Dashboard', under the 'My Communities' tab, click your community to highlight it, and click the 'Go to community' button. You will be directed to your community page.
3. Click on the 'Data Sources' tab and click on your dataset. The entry will expand and show information about the number of localities georeferenced.

NOTE:

- You can also do this in the community-wide tab if you have more than one data source.

Method 2 (a user's progress):

1. Navigate to <http://www.museum.tulane.edu/coge>. Log in. You will be directed to your 'Dashboard'.
2. In the 'Dashboard', under the 'My Communities' tab, click your community to highlight it, and click the 'Go to community' button. You will be directed to your community page.
3. Click on the 'Members' tab and after it redirects you to the member management page, click on a particular member. The member's individual progress will be displayed.

Method 3 (the user's progress while georeferencing):

1. Navigate to
   <http://www.museum.tulane.edu/geolocate/web/webcomgeoref.aspx>. Log in.
   Select your community. Click the 'Continue' button.
2. On the bottom left-hand button menu, selecting 'Correction Report' will display
   a gray pop-up box detailing the member's progress with their assigned work.

Downloading data

1. Go to the collaborative georeferencing community page
2. Click on the "Data Source" tab
3. Click on the data source you want to download
4. Click on the "Download Records" button
5. Check the fields you want to download
6. Choose the format you want for the file
   a. CSV
      i. Not the default, so make sure to change it if you want it.
      ii. Keeps all of the commas and quotation marks
   b. Tab
      i. Makes a simple file
      ii. Strips out embedded tabs
7. You can pick just the localities or just the specimens
8. You can also tell it to download the error polygons if you have them
   a. Pick the format
      i. WKT
         1. This is a standard
         2. Works well with GIS software
      ii. GLC
9. Decide how much of the data you want to download.
   a. You can decide if you want skipped records, corrections, and unprocessed records.
   b. You can also pick all the work that's been done, or just the recent work
      i. All work will give you all corrections, so if you went and updated a record because you placed it in the wrong place, you will get two corrections for this.
      ii. Recent work (A.K.A., the latest verification) will just give you the most recent updated point for each locality.
10. Click "Download" button
11. You now have a file that you can ingest into your database!
   a. Figuring this step out is complicated. Specify is working on a way to make this easy. More on this step to come!

NOTE:
- You can also download a KML file, so that you can see your work in Google Earth! It is all the same steps as above except you pick the KML format rather than delimited text option. All the associated metadata will be there.

External resources for georeferencing

- Nelson Rios has a wonderful webinar for using this portal recorded at <https://vimeo.com/album/2163673/video/96842881>
- The UO Museum of Natural and Cultural History uses Trello <https://trello.com> to manage this project

- For converting TRS to lat/lon, we use
  <http://www.earthpoint.us/TownshipsSearchByDescription.aspx>
- For a geologic map of the USA <http://mrdata.usgs.gov/geology/state/>
- To find the locations of named places
  <http://geonames.usgs.gov/apex/f?p=136:1:0:::::>
- To see historical maps <http://historicalmaps.arcgis.com/usgs/>
- Other information from iDigBio
  <https://www.idigbio.org/wiki/index.php/Georeferencing>