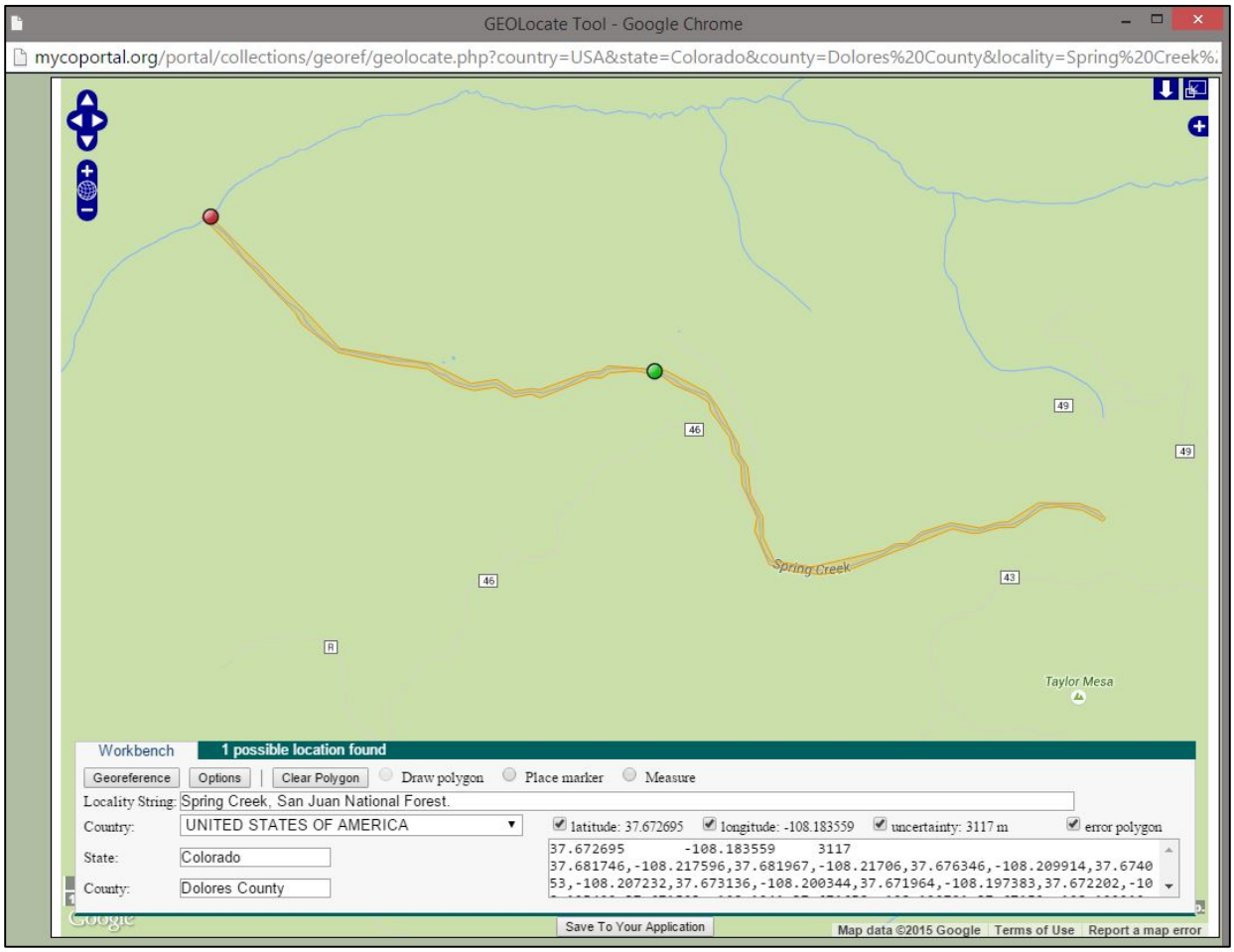
Georeferencing: The Polygon Method

Mike Yost
MaCC Project Assistant



Introduction

Greetings fellow georeferencers!

Mike Yost here, Macrofungi Collection Consortium (MaCC) Project Assistant. In my previous blog, I outlined how to batch georeference specimens using the point-radius method.

For this blog, I'm going to demonstrate how I use the polygon method to georeferencing specimens, utilizing the embedded GEOLocate tool in Symbiota.

> **NOTE:** If you need a quick refresher on how to batch georeference specimens, you can read through my previous blog on iDigBio: "Steps in Georeferencing Specimen Locality Data - Community Examples"

Though it is currently standard protocol at most institutions to georeference specimens using the point-radius method, often the radius that is generated encompasses additional landmarks such as nearby rivers, mountains, or roads that should not be included within the uncertainty.

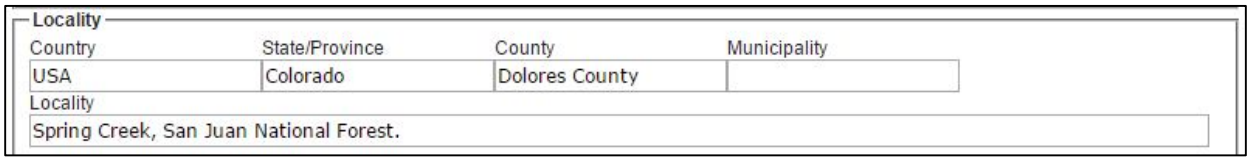
Polygons increase the utility of georeferencing data associated with a location by further reducing the area of uncertainty around a collection site.

For example, a case study conducted in-part by the University of California in 2011 found that using the polygon method (compared to the point-radius method) reduced the uncertainty of several localities at Yosemite National Park by up to 99.5 percent.

To demonstrate how this reduction in uncertainty area is accomplished, let's start with a quick overview of the polygon method.

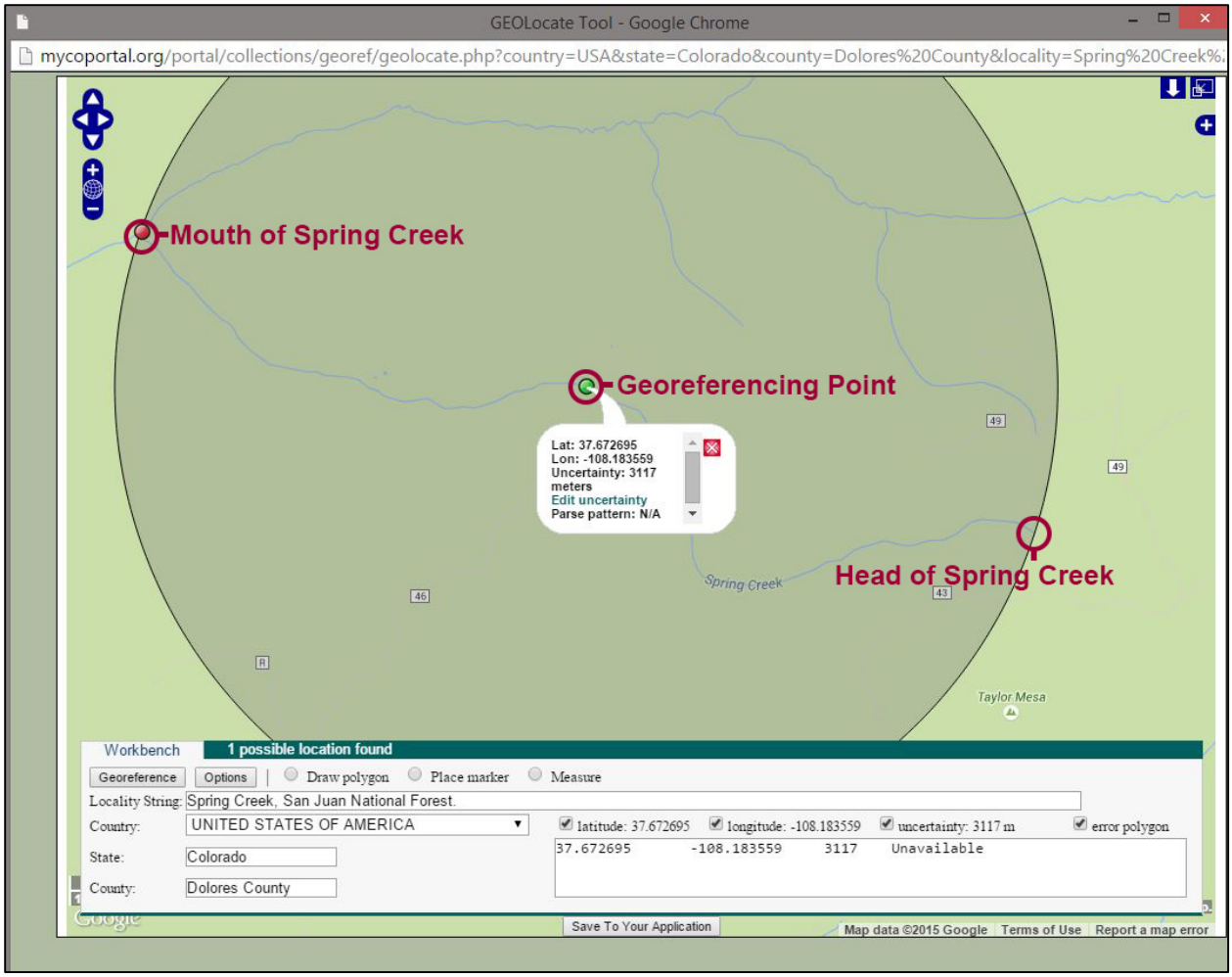
Polygon Method: A Brief Summary

In the example below, I have a specimen that was collected along Spring Creek in Colorado's San Juan National Forest.



The current protocol, following the point-radius method, is to place a georeferencing point at the approximate midpoint of Spring Creek between the mouth of the creek and the head of the creek, ensuring the uncertainty radius encompasses the entire length of the creek.

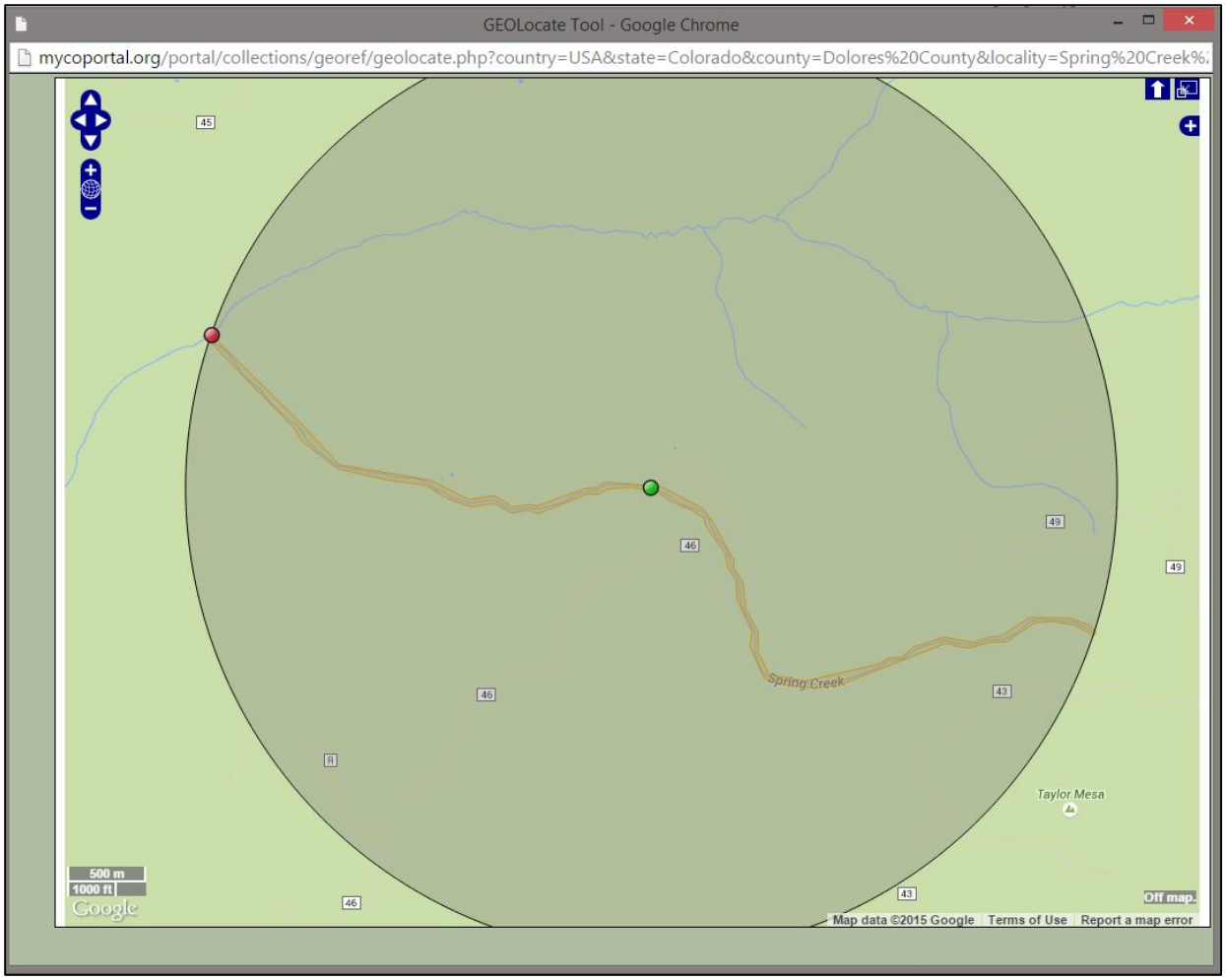
Following that standard, I georeference the specimen at Spring Creek with the following results:



As you can see, the georeferencing point has been placed at the approximate midpoint of Spring Creek, with the uncertainty radius extending to the mouth of the creek to the northwest (at Stoner Creek) and to the creek's head to the southeast.

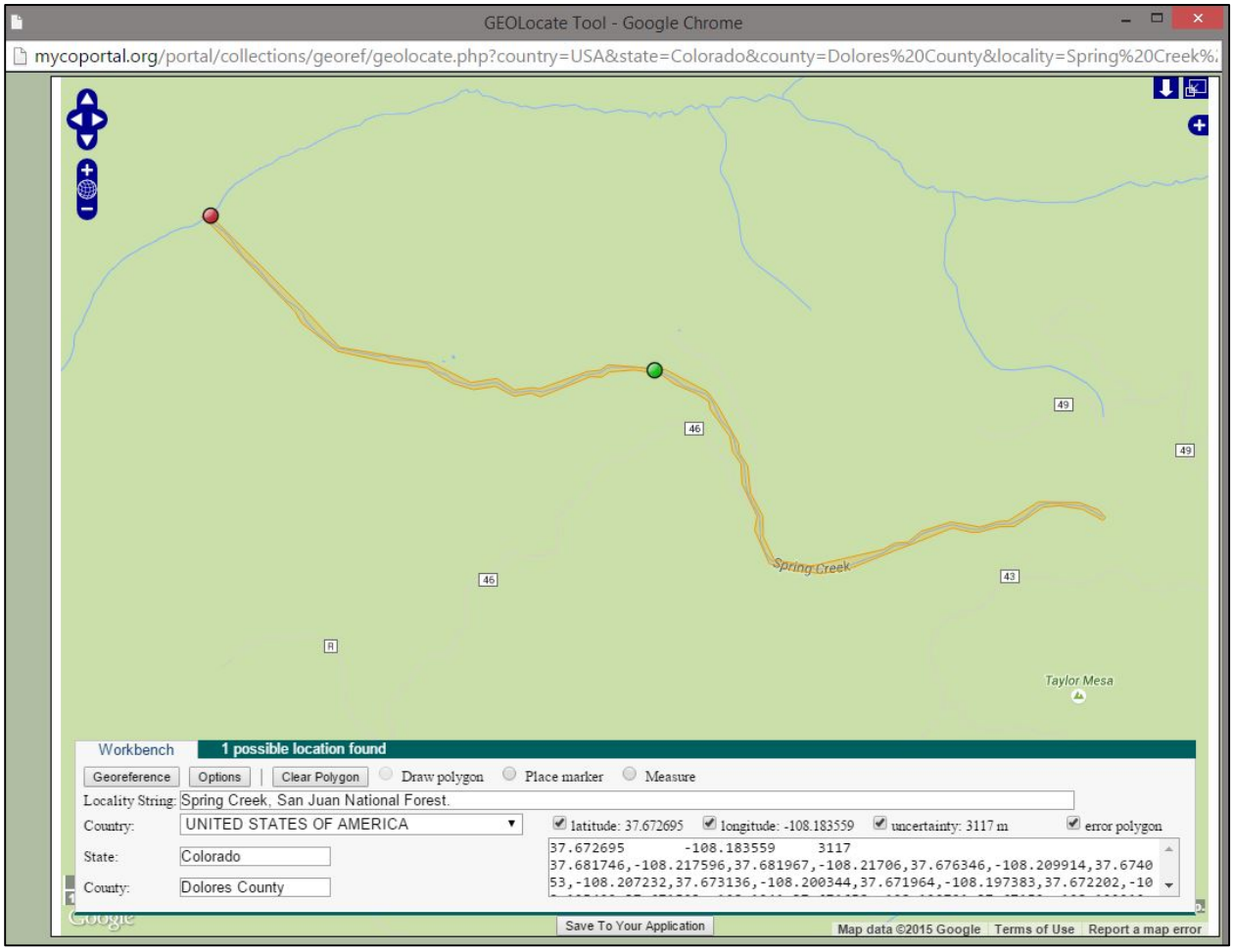
However, I also have a large area of uncertainty around the creek that not only encompasses a section of Stoner Creek, but several additional creeks to the north as well as Forest Road 46 to the south.

To limit the area of uncertainty, I want to draw a polygon around the entire perimeter of the creek. I will outline how this is done in the next section, but for now, let's look at the end result (without the Workbench displayed at the bottom of the map).



As you can see, in addition to the uncertainty radius, I now have a polygon (outlined in yellow) encompassing the entirety of Spring Creek.

To get a better look at the polygon, I remove the uncertainty radius from the map (I'll demonstrate how this is done in the next section).



Notice that the lat/long coordinate pairs located in the bottom right corner of the Workbench (more on this later). If I were to copy and paste all of those coordinates, I would have the following result:

Lat: 37.672695
Long: -108.183559
Uncertainty: 3117
Polygon: *37.681746,-108.217596,37.681967,-108.21706,37.676346,-108.209914,37.674053,-108.207232,37.673136,-108.200344,37.671964,-108.197383,37.672202,-108.195409,37.671523,-108.1941,37.671658,-108.192791,37.67159,-108.192212,37.67261,-108.188671,37.672694,-108.187684,37.673017,-108.187105,37.672728,-108.18277,37.671404,-108.179616,37.670571,-108.178865,37.668329,-108.177299,37.66782,-108.177256,37.667005,-108.176612,37.665425,-108.176419,37.664032,-108.175475,37.662809,-108.175453,37.661433,-108.174617,37.661059,-108.171634,37.661076,-108.171827,37.662062,-108.16642,37.662435,-108.165604,37.662486,-108.164617,37.663132,-108.163609,37.663692,-108.161377,37.66359,-108.160583,37.663387,-108.159188,37.663709,-108.157858,37.66376,-108.157064,37.664881,-108.154811,37.664847,-108.151743,37.664406,-108.150284,37.663947,-108.149769,37.663794,-108.149983,37.66444,-108.151464,37.664661,-108.153245,37.664576,-108.154446,37.663421,-108.156614,37.663404,-108.157794,37.663098,-108.15921,37.663064,-108.159746,37.663336,-108.16127,37.662435,-*

108.164102,37.662028,-108.165476,37.66089,-108.169059,37.660635,-108.17114,37.660805,-108.173115,37.661229,-108.175003,37.663217,-108.176097,37.664032,-108.175904,37.665951,-108.177041,37.66714,-108.177127,37.667854,-108.177749,37.668771,-108.177964,37.670334,-108.179402,37.67086,-108.179552,37.671336,-108.180517,37.671845,-108.182062,37.672474,-108.183178,37.672711,-108.185109,37.672593,-108.186976,37.672287,-108.187556,37.672304,-108.188435,37.671115,-108.192169,37.671336,-108.192877,37.671149,-108.194122,37.671743,-108.195538,37.67154,-108.19719,37.671913,-108.198714,37.672202,-108.1994,37.67283,-108.200645,37.673119,-108.203348,37.673815,-108.207361,37.675497,-108.210086,37.676176,-108.21058,37.681848,-108.21751,37.681746,-108.217596

Wow! That's a lot of lat/long coordinates! Each pair of coordinates (*37.681746 and -108.217596, for example)* represents a single lat/long plotpoint I used to draw the polygon around Spring Creek.

Before I outline my polygon process, I want to review a few limitations and cautions when using the polygon method to georeference specimens.

Caveats

Visualization: There is no way in Symbiota (*as of yet*) to display or *see* the original polygon drawn in GEOLocate after the data has been saved to the Portal. In other words, the polygon I drew for Spring Creek in the previous example cannot be seen again once the data points have been uploaded to the Portal's server.

Adjustments: As a result, there is no way to reopen the polygon I drew using the built-in GEOLocate tool so that I can make adjustments to the polygon. It's similar to how there's currently no way to *see* the uncertainty radius I establish when using the point-radius method after the data has been saved to the Portal.

> **NOTE:** Bear in mind that there are other programs (such as Quantum GIS) which *can* use the data points to draw the polygon on a map. *In addition, there are plans to add this feature to the built-in GEOLocate tool in Symbiota.*

Efficacy vs. Specificity: As I've stated before, the main advantage of polygons is the reduction of uncertainty associated with that specimen, creating a more accurate representation of the location site. However, this process is more time consuming than the point-radius method. *Be sure to check with your project manager regarding which method should be utilized at your institution.*

Supplemental Data: The polygon method *does not* supplant the point-radius method when georeferencing specimens. Drawing a polygon around a locality is *an additional step* in georeferencing specimens that simply increases the utility of the georeferencing data.

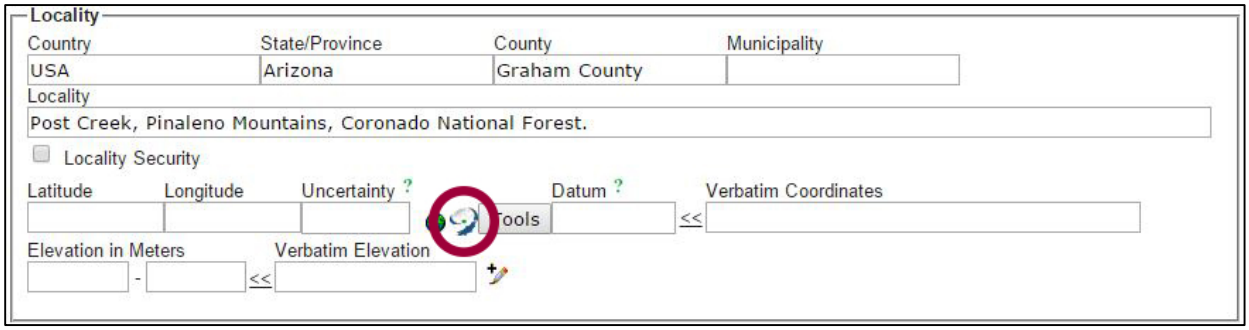
With that said, let's walk through each step in the polygon-method process, using an example of a specimen collected at a different creek.

Post Creek, Coronado National Forest, Arizona

There are five steps I follow to accurately georeference a specimen using a polygon, starting with the establishment of the specimen's lat/long coordinates and uncertainty radius.

Step #1: Point-radius Groundwork

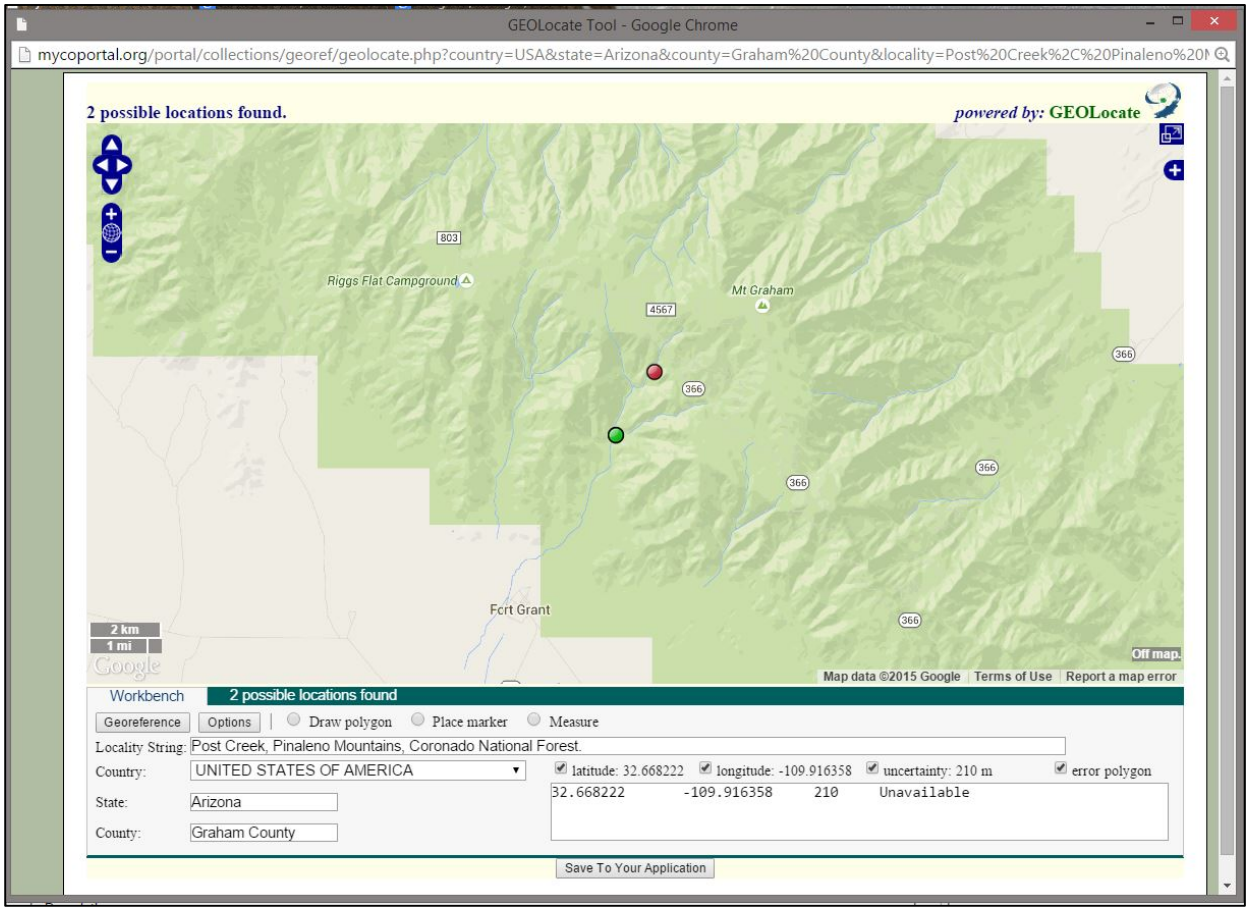
Below, I have a specimen that was collected along Post Creek in Arizona's Coronado National Forest.



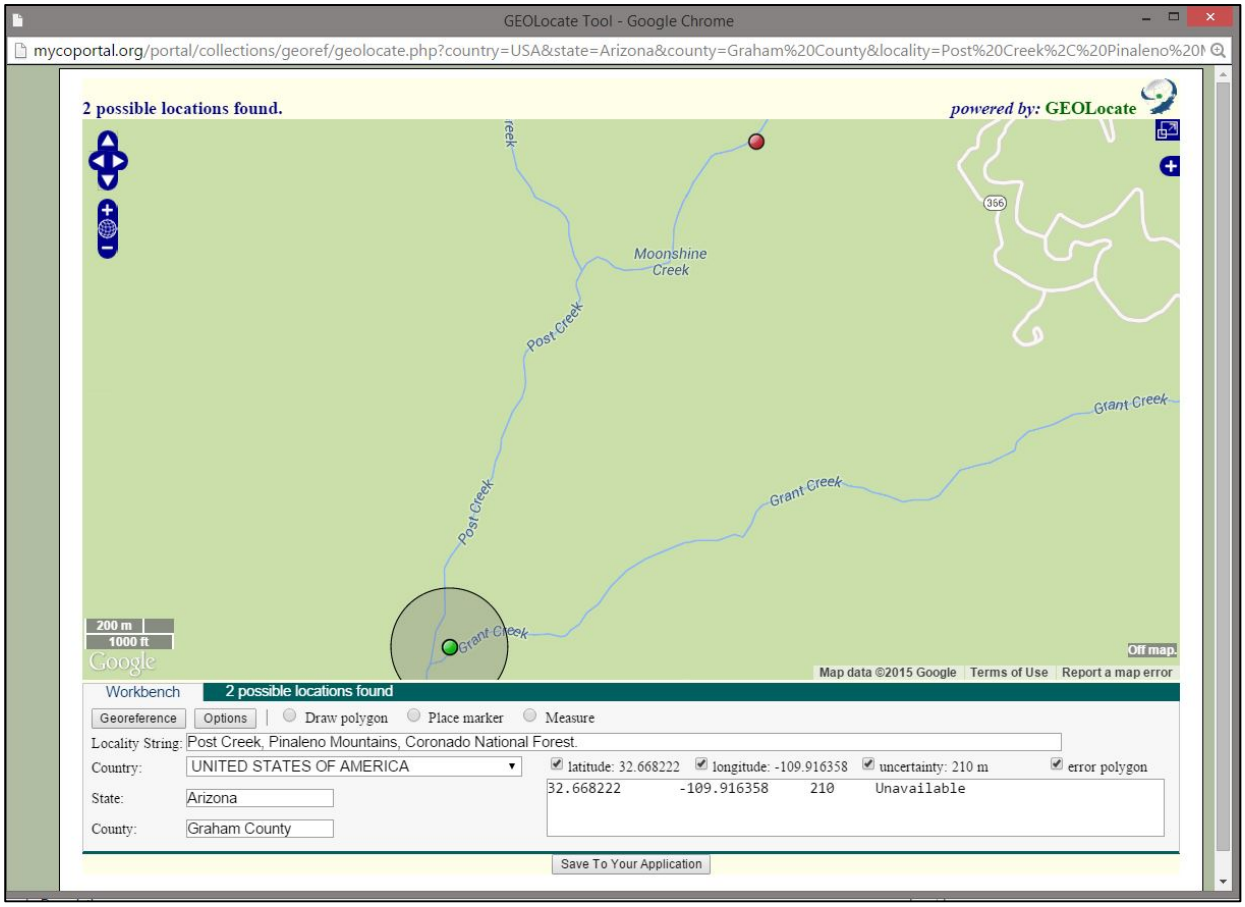
As you can see in the image above, there is no specific locality data for this specimen. So, I want to begin the polygon process by first establishing my lat/long coordinates and uncertainty radius. To do that, I click on the GEOLocate button located next to the *Tools* button.

> **NOTE:** In some cases, I'm required to draw a polygon around a locality that already has lat/long coordinates and an uncertainty radius. In those cases, I begin at step #2 of my procedure.

Once I click on the GEOLocate button, the following map appears. Notice that GEOLocate generated two possible georeferencing points for Post Creek.



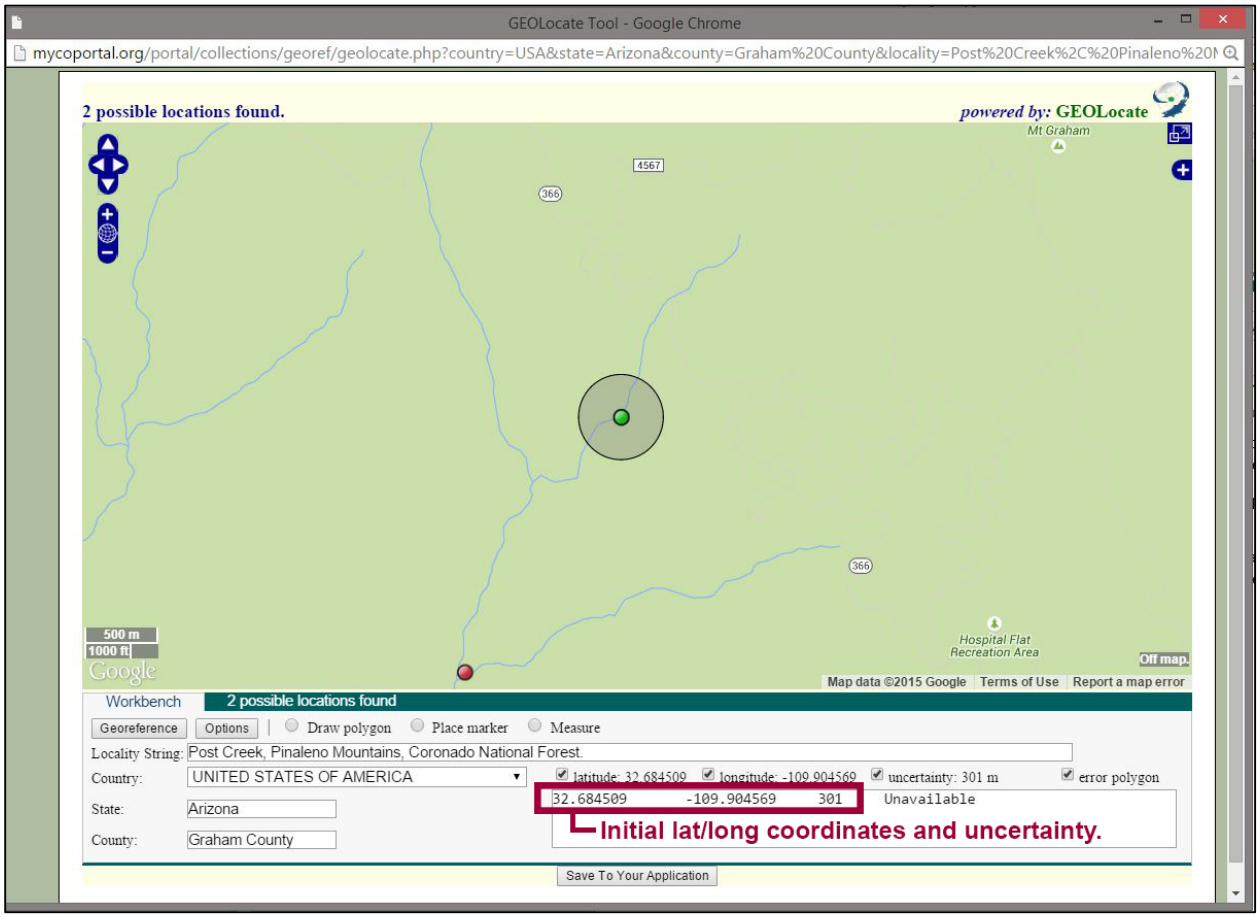
I need to verify if either of those georeferencing points are accurate before moving forward, so I zoom into the map to get a closer look.



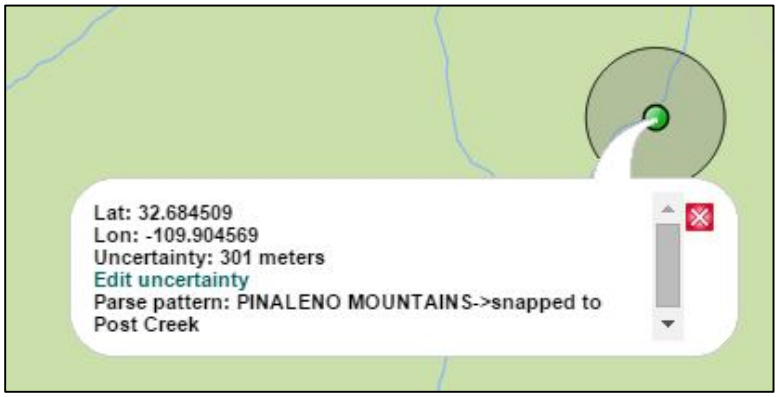
The first georeferencing point (marked in green) is located near the mouth of Post Creek at Grant Creek at the bottom of the map. The second georeferencing point (marked in red) is located along Post Creek northeast of Moonshine Creek at the top of the map.

Remember that I need to establish lat/long coordinates which represent the approximate midpoint of Post Creek. So, I'm going to begin with the red plotpoint near Moonshine Creek.

I scroll up the map so that the second plotpoint is at the center of the screen, then I zoom out a bit so that I can see the entirety of Post Creek. When I left click on the red plotpoint, it turns green, and the lat/long coordinates associated with that plotpoint are displayed at the bottom right corner of the Workbench (see below).

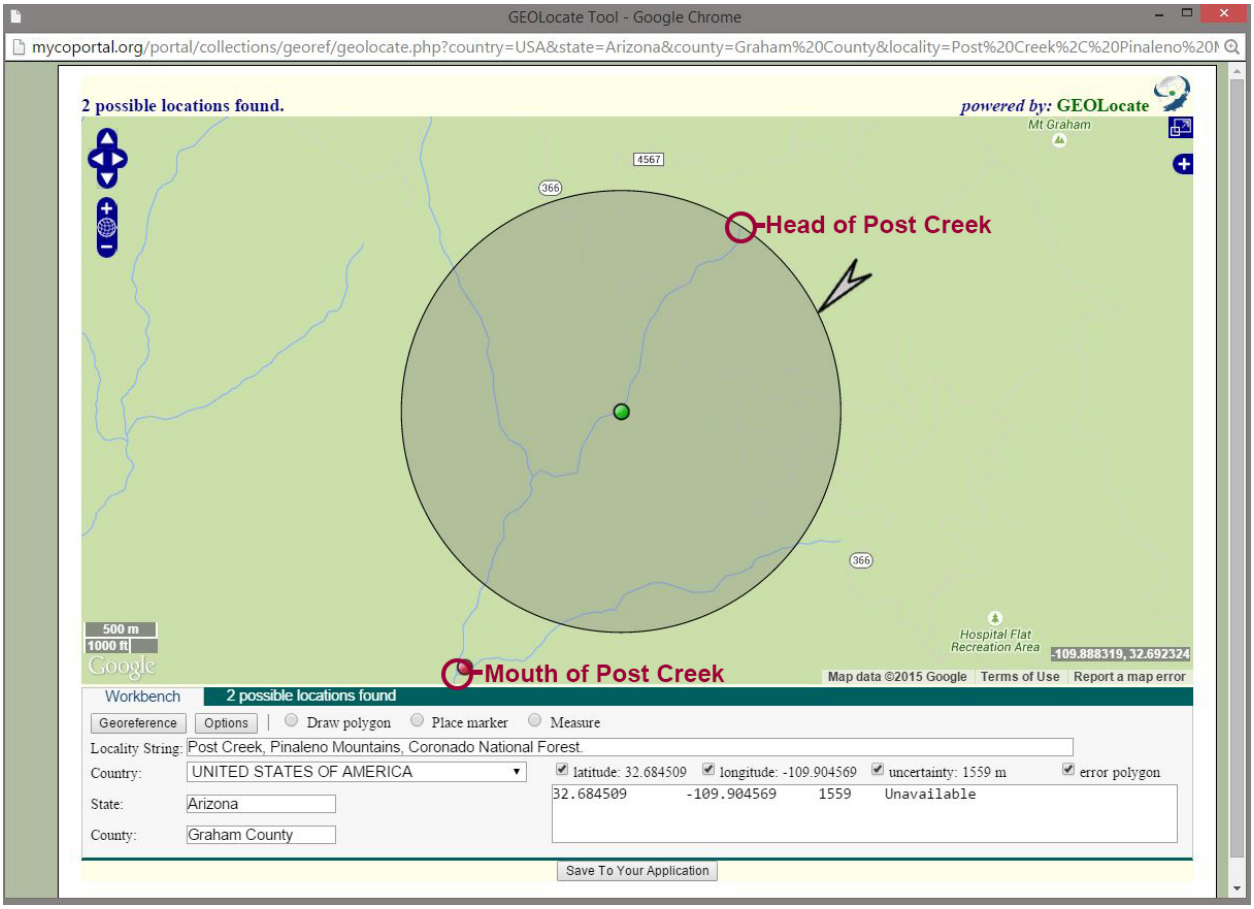


At first glance, it looks as if the green georeferencing marker is not quite at the midpoint of Post Creek. The only way I can accurately verify this is to expand the uncertainty radius. To do that, I left click on the green georeferencing marker to open up a display which contains the lat/long coordinates and the uncertainty in meters (see below).

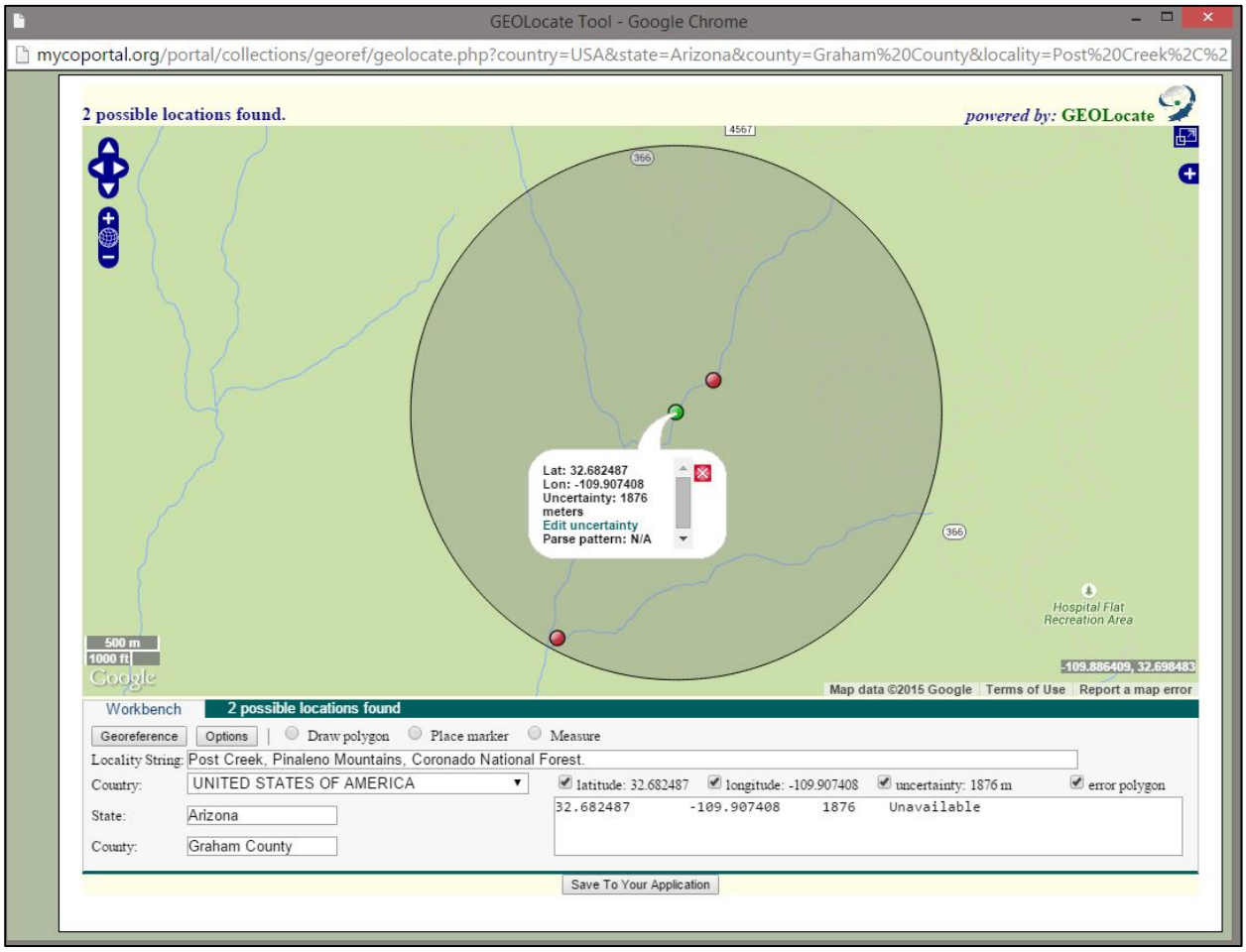


I click on *Edit uncertainty* option highlighted in blue, and a gray arrow appears at the upper right of the uncertainty circle. I expand the uncertainty radius by left clicking and holding on the gray arrow, dragging it away from the georeferencing point to expand the radius.

In the image below, you'll notice that although the northeastern edge of the radius reaches the head of Post Creek, the southwestern edge of the radius does not reach the mouth of Post Creek at Grand Creek. This means the green georeferencing marker is not at the midpoint of Post Creek. I need move the green plotpoint farther southwest along the creek.



Just like when I batch georeference specimens using the point-radius method, I sometimes have to move back and forth between moving the georeferencing marker and adjusting the uncertainty radius until I end up with the following result:



You can see that there's a third georeferencing point (marked green). This is the plotpoint I just established, now located at the approximate midpoint of Post Creek. I can verify this because the uncertainty radius borders the mouth of Post Creek (at Grant Creek) to the southwest and the head of Post Creek to the northeast.

I now have my final lat/long coordinates along with my uncertainty radius (measured in meters).

Lat: *32.682487*
Long: *-109.907408*
Uncertainty: *1876*

If I were georeferencing this specimen using the point-radius method, this would be the last step (before saving the data to the Portal). But of course I'm using the polygon method, so I need to draw my polygon *first* before saving any data.

> **NOTE:** Do not click on *Save To Your Application*. If I were to save the data to the Portal at this point, the map would disappear. There would be no way to access or *see* the map in GEOLocate with the lat/long coordinates and uncertainty radius I just established.

With that said, it's on to drawing a polygon!

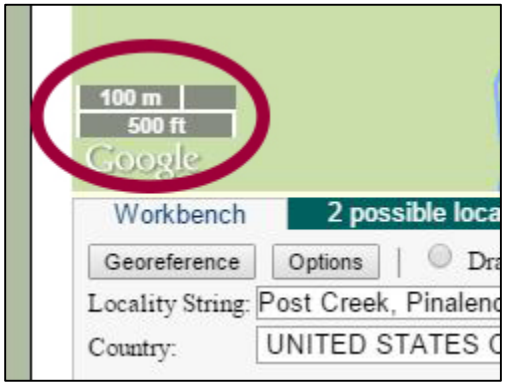
Step #2: Drawing the Polygon

Before moving forward, a quick note about the zoom feature on the map. Keep in mind that the closer I zoom into Post Creek, the more detailed the polygon will be as I draw it around the parameters of the creek. However, the closer I zoom in, the longer it will take to draw the polygon. I have to strike the right balance between specificity and efficiency.

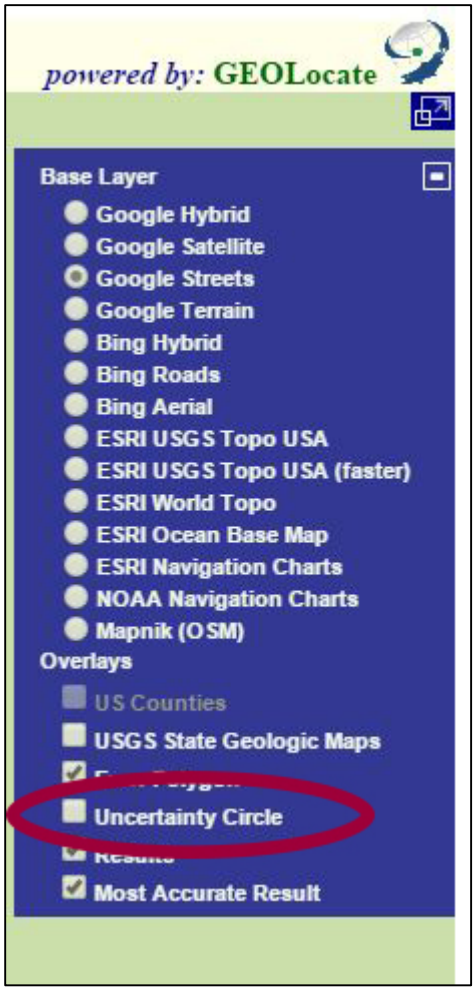
For example, when I'm drawing a polygon around a large canyon that spans several kilometers, I zoom out farther on the map than I would if I'm just drawing a polygon around a small creek which runs only a couple of hundred meters.

As of the publication of this blog, there is no established protocol regarding the distance in zoom for drawing polygons. Be sure to check with your project manager regarding how closely you need to be zoomed into the map. Keep in mind that how closely you are zoomed in on the map may change based on the type of locality you are georeferencing.

With that said, I'm going to start my polygon for Post Creek at a zoom of 100 meters/500 feet (as displayed on the legend at the bottom left corner of the map).

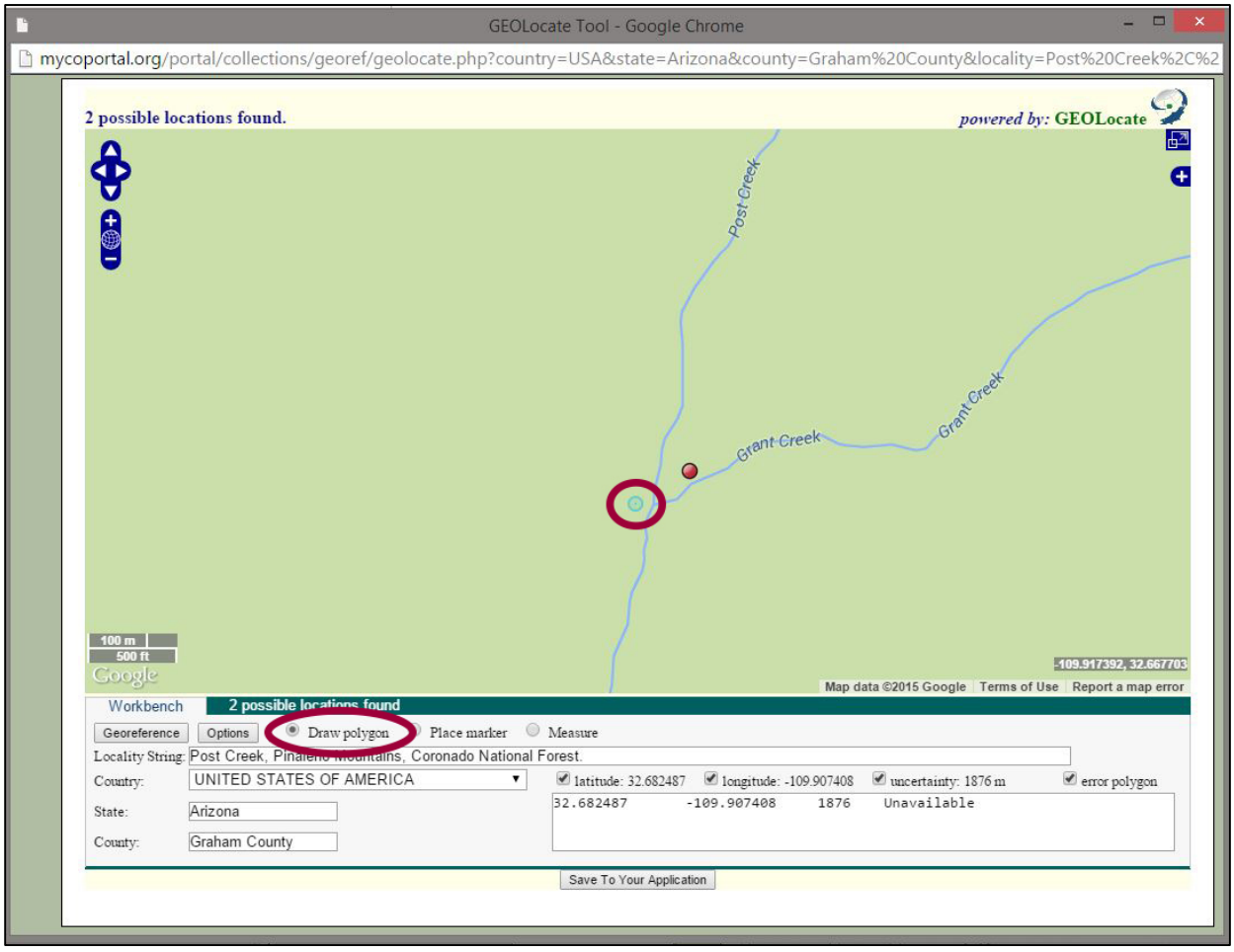


In addition, I want to remove the displayed uncertainty radius so I can see Post Creek more clearly. I do this by clicking on the plus sign (+) located at the upper right corner of the map. After the menu appears, I click on the *Uncertainty Circle* option under **Overlays**.



Now that I'm at a comfortable zoom distance, and the uncertainty radius has been removed, I'm ready to being. I find that the best place to start drawing a polygon for a creek is at the creek's mouth.

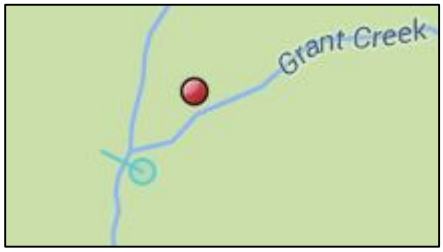
So, I move the map to where Post Creek intersects with Grant Creek. I click on the *Draw polygon* option located on the Workbench next to the *Place marker* option. I then left click near the mouth of Post Creek. A blue circle appears, indicating the starting point for the polygon.



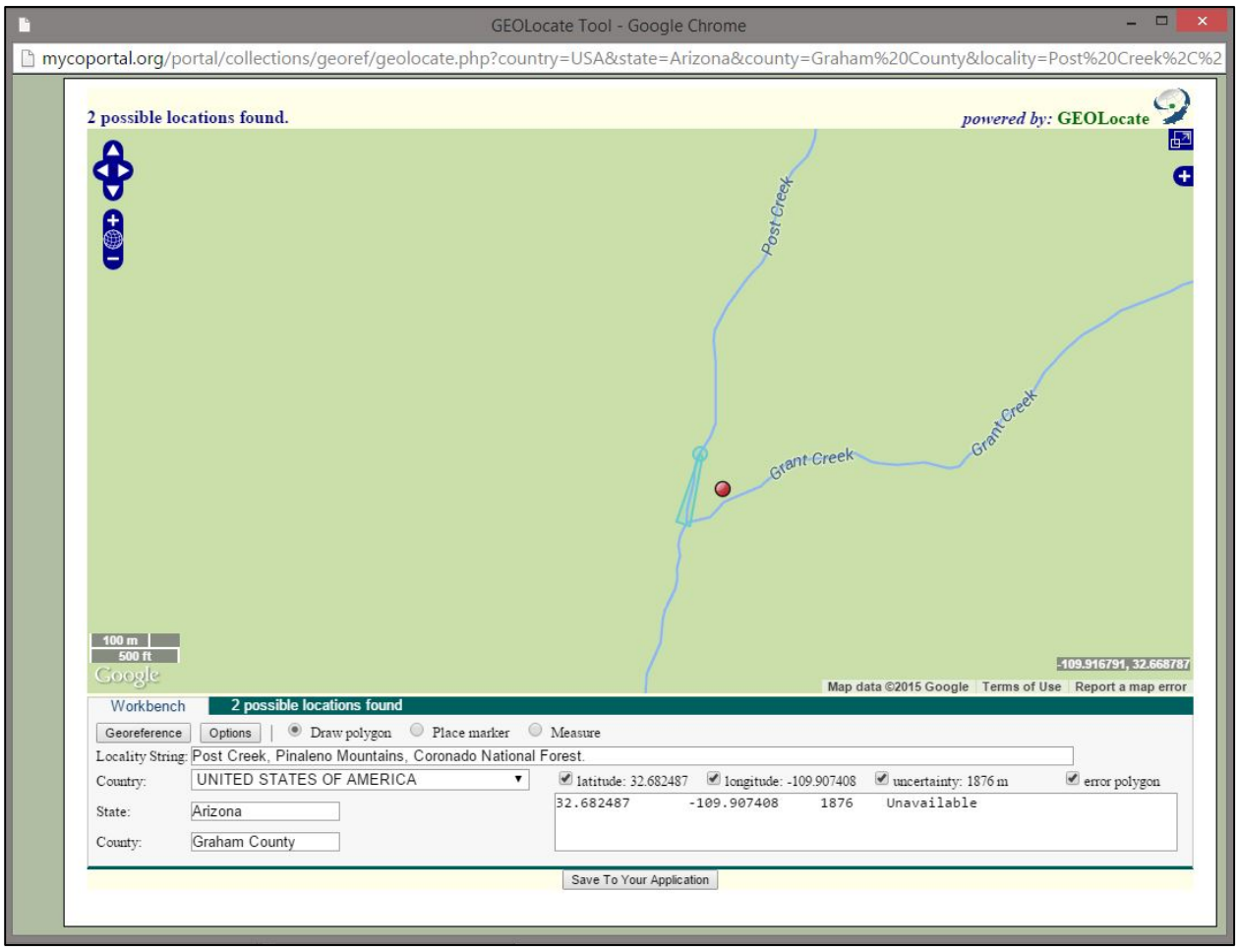
NOTE: If you make a mistake in placing the initial starting point of the polygon, no worries. Simply click on the *Draw polygon* option again and the blue circle disappears. Then click on the same option again to start drawing your polygon.

Now that I have my blue circle (indicating the starting point), I can begin drawing the polygon. When I move my cursor along the map, a blue line follows the cursor with one end originating at the starting point, and the other end attached to the cursor.

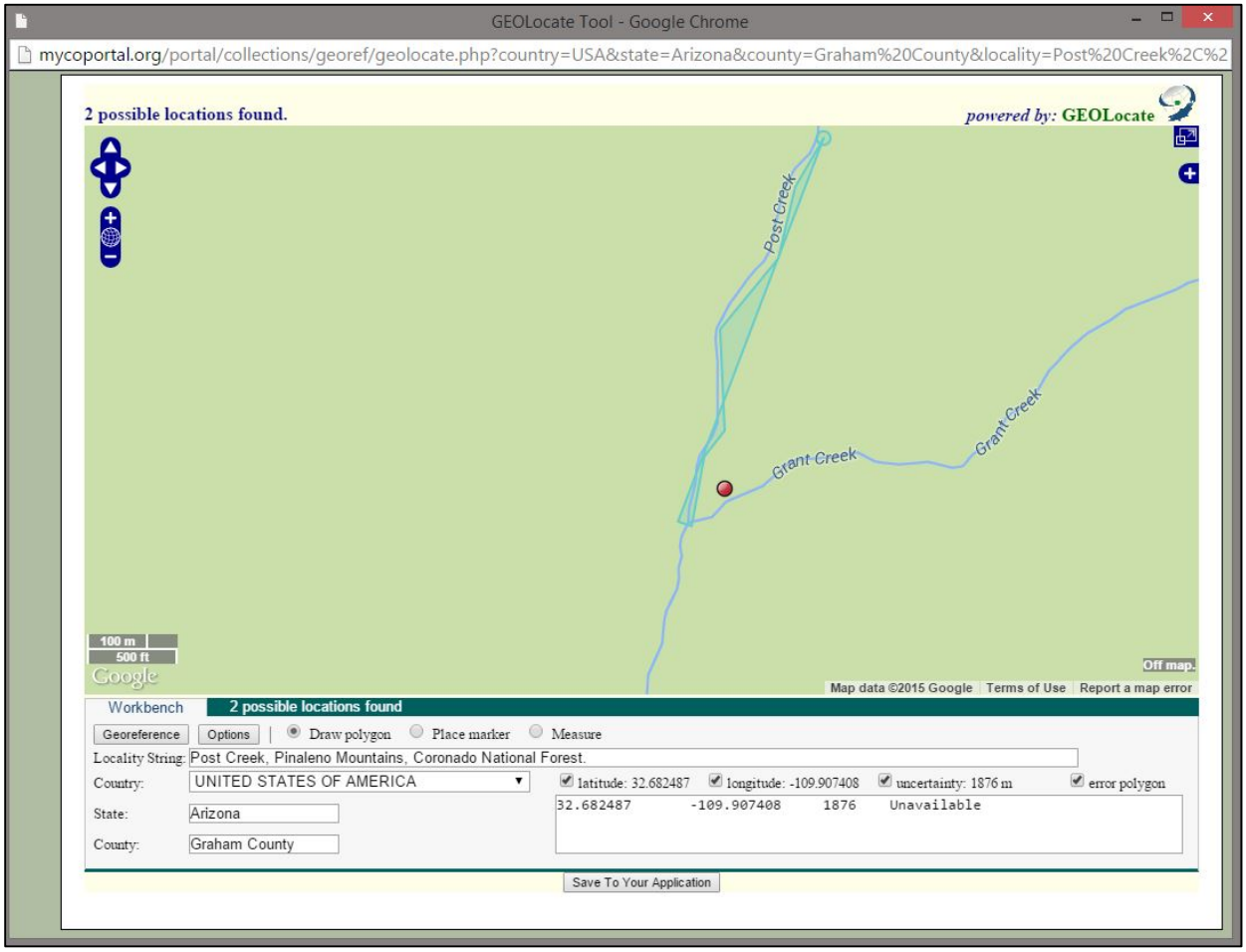
I use that blue line to draw the lower section of the polygon which will run perpendicular to Post Creek just south of the creek's mouth.



I then left click, and a second starting point of the polygon is generated where I can draw a new angle (or bend) in the polygon. I continue this process, now drawing the polygon parallel to Post Creek, moving north along the west bank (see below).



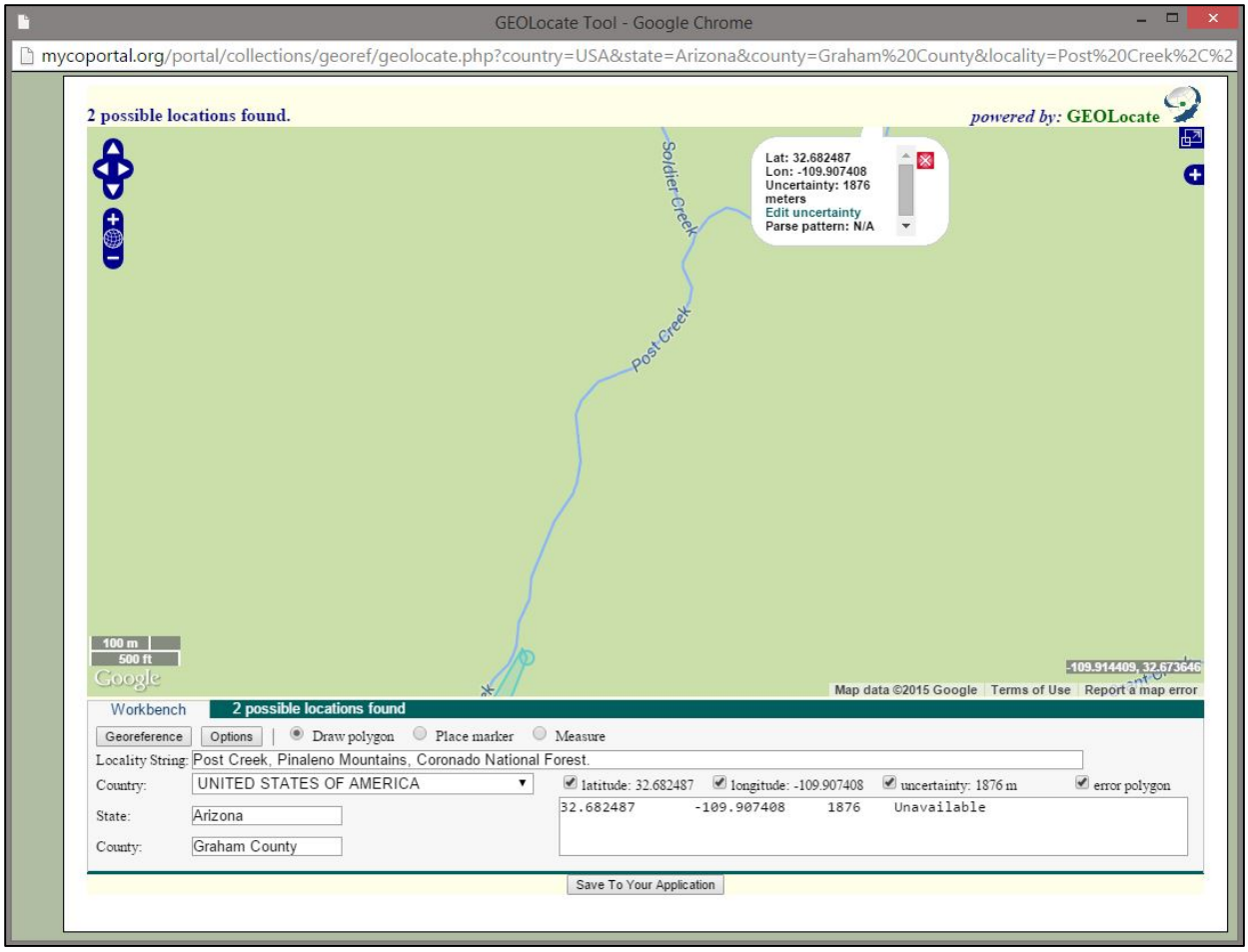
I continue to left click as I move northeast along the east side of Post Creek, being careful to ensure that each bend or angle of the polygon corresponds with the contours of the meandering creek.



Notice the polygon I've drawn at this point looks a bit odd, overlapping not only the creek but the polygon itself. This will correct itself once I start drawing the polygon along on the opposite side of Post Creek.

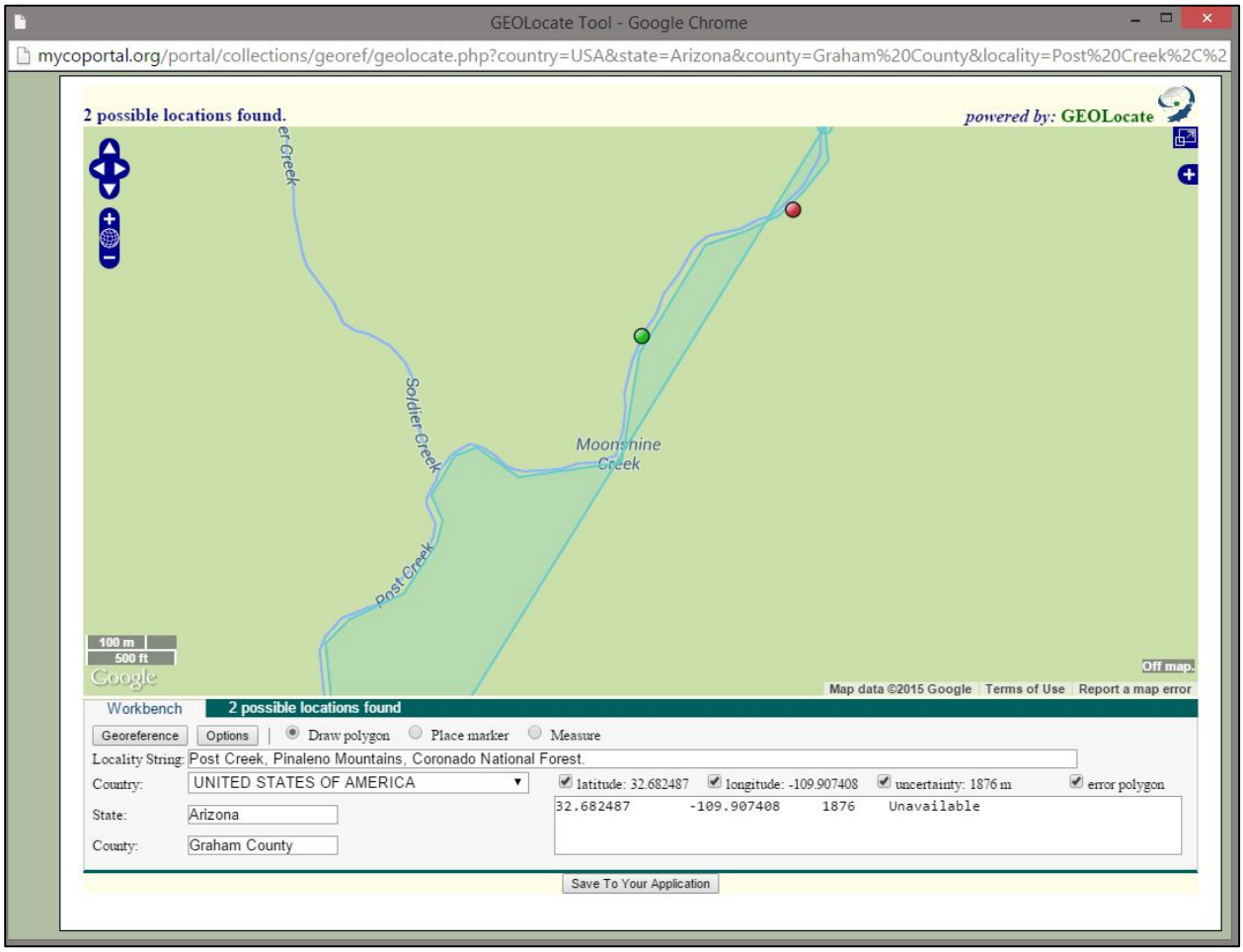
You can also see in the example above that I've reached the top of the map with my polygon, even though Post Creek continues north off of the displayed map.

To scroll north on the map (so I can see the continuation of Post Creek), I simply left click *and hold*, pulling the map down so I can see more of Post Creek. I then continue to draw the polygon by left clicking at each angle of the polygon (see below).

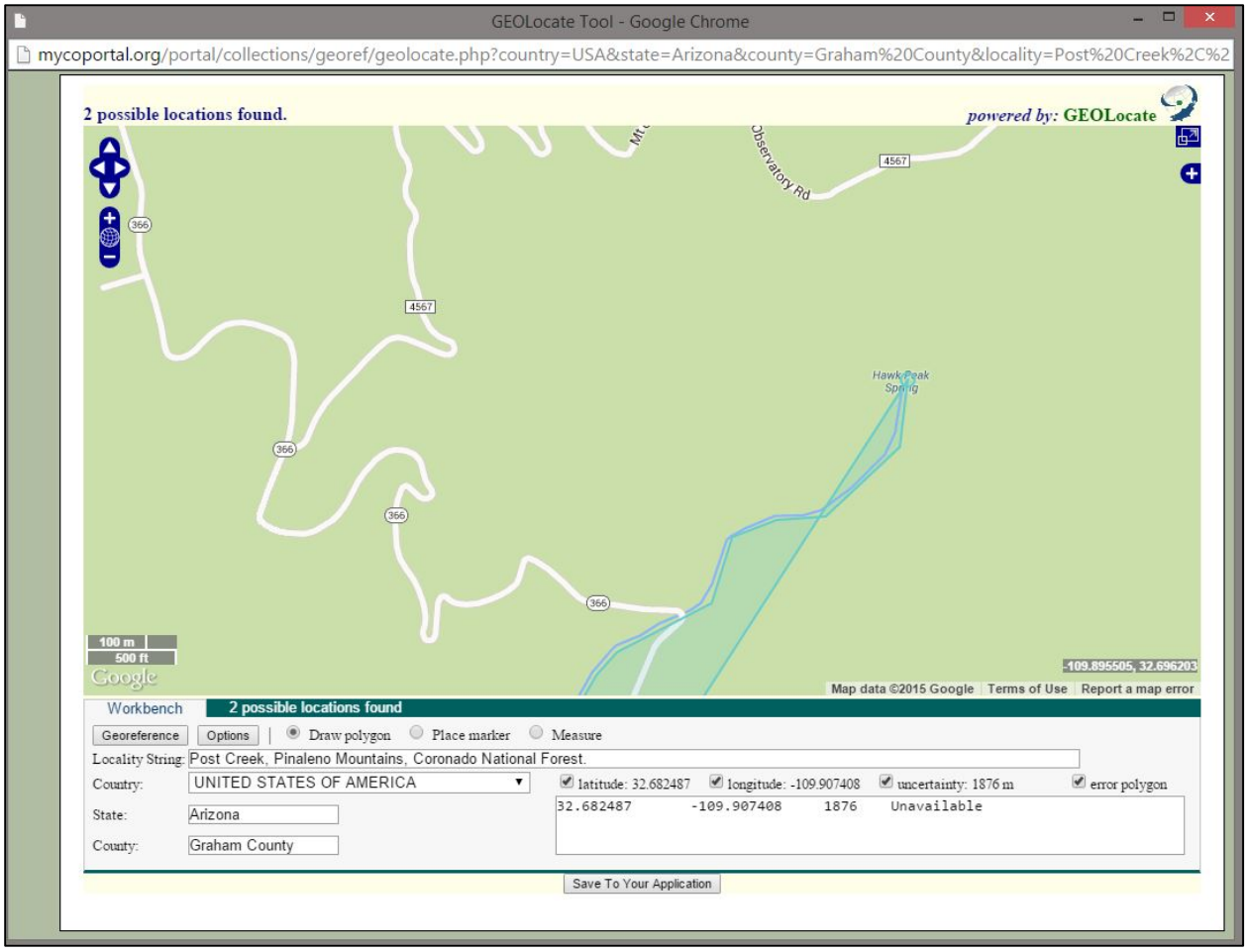


Remember that if you make a mistake in drawing the polygon, you can always click on the *Draw polygon* option and start again.

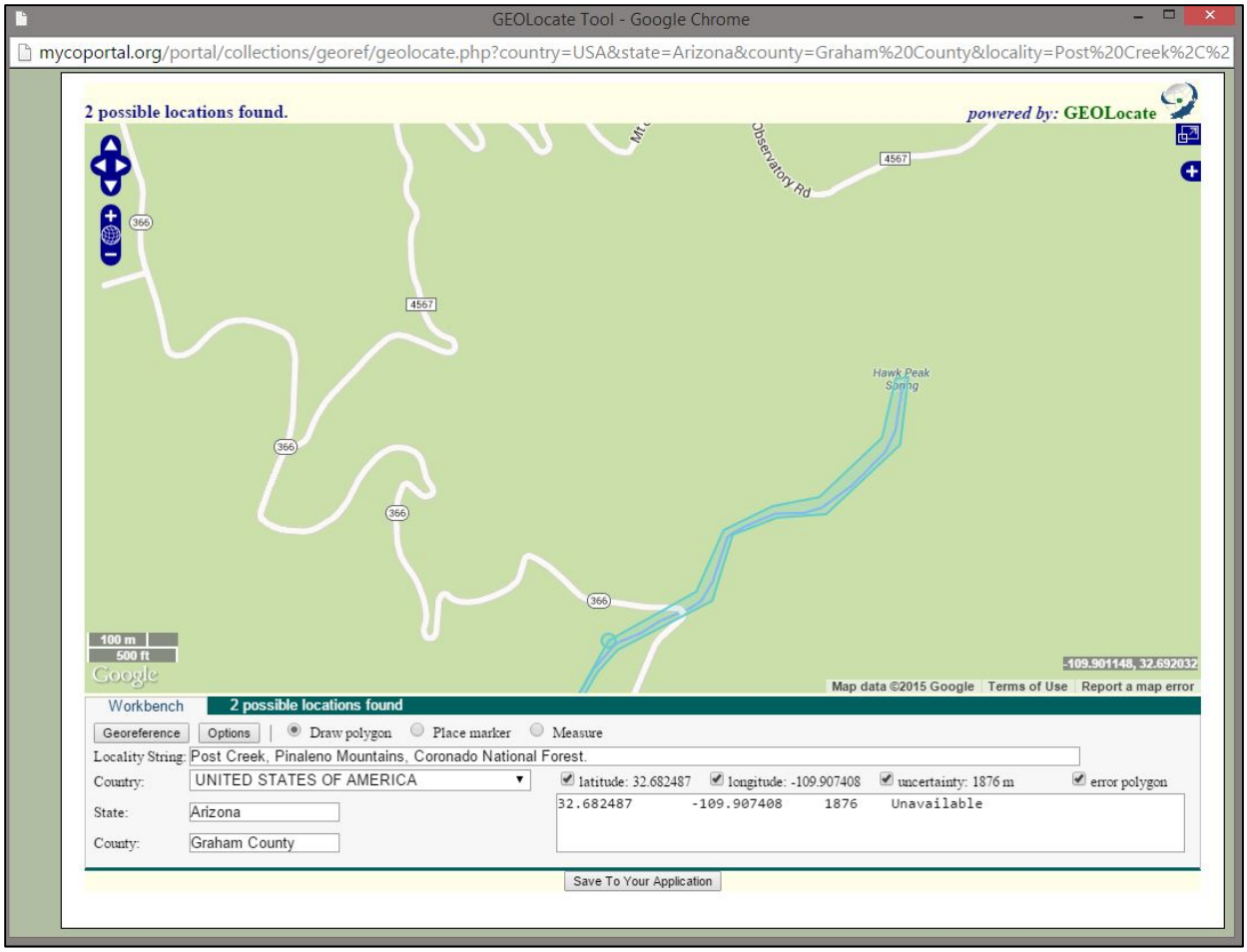
Below you can see once again that I've reached the top of the map, so I left click *and hold* to scroll north, then continue to draw my polygon along Post Creek. I also removed the white bubble (which contained the lat/long coordinates) from the green georeferencing point so that the creek wasn't masked by the information bubble.



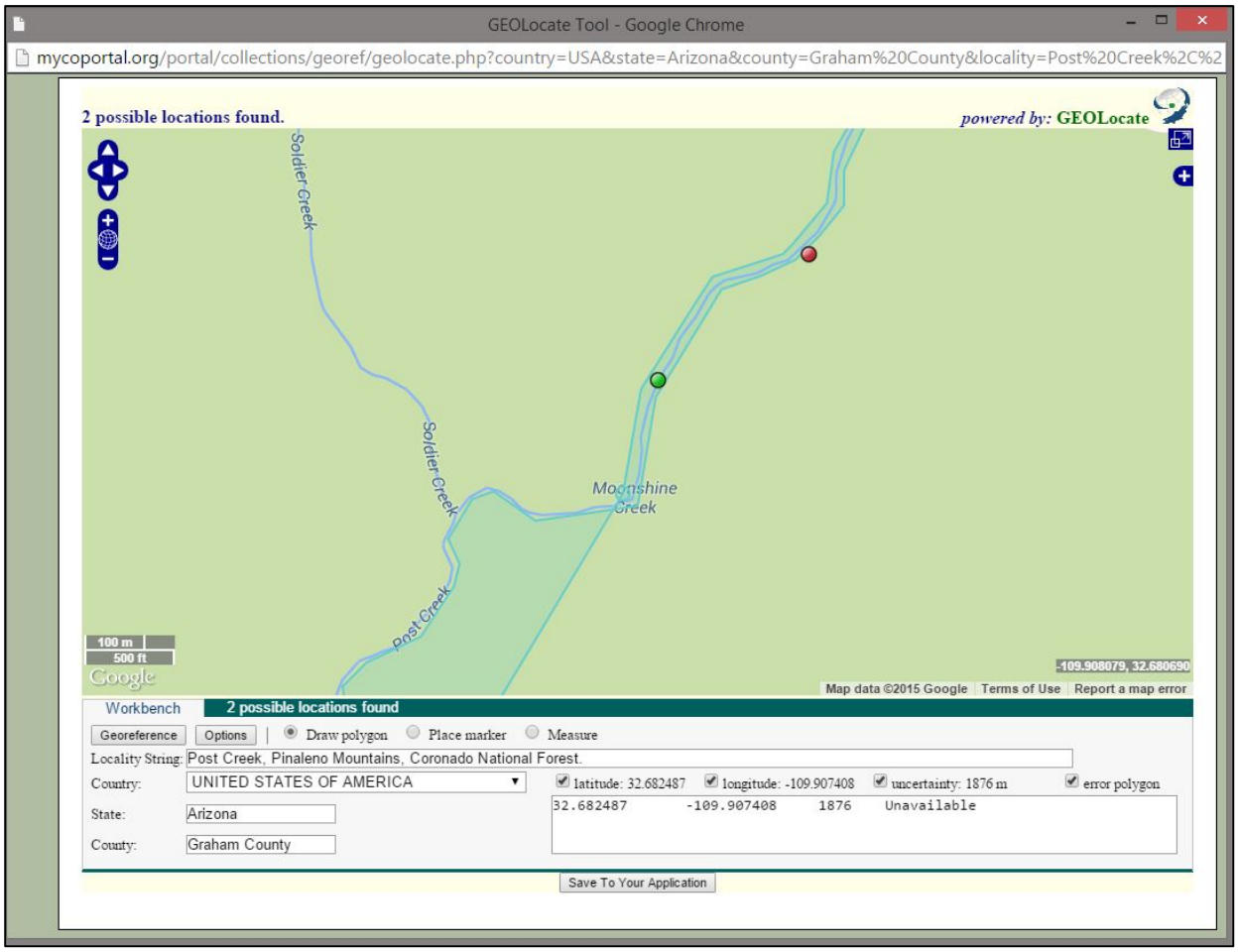
I continue left clicking along the east bank of Post Creek, moving the map accordingly until I reach the creek head near Hawk Peak Spring.



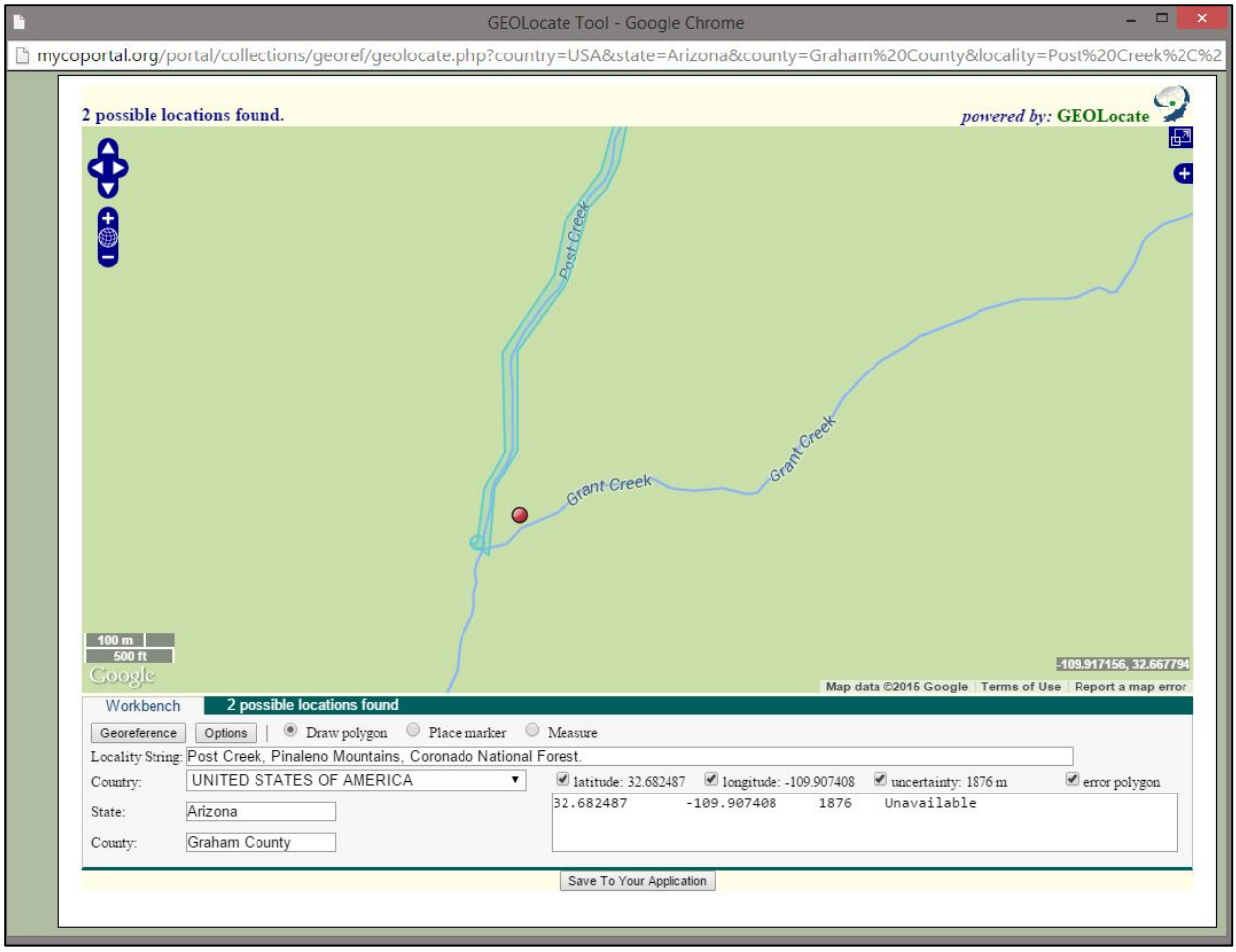
Now that I've reached the head of Post Creek, I begin placing each plotpoint of the polygon along the west side of the creek, working my way back down to the creek's mouth.



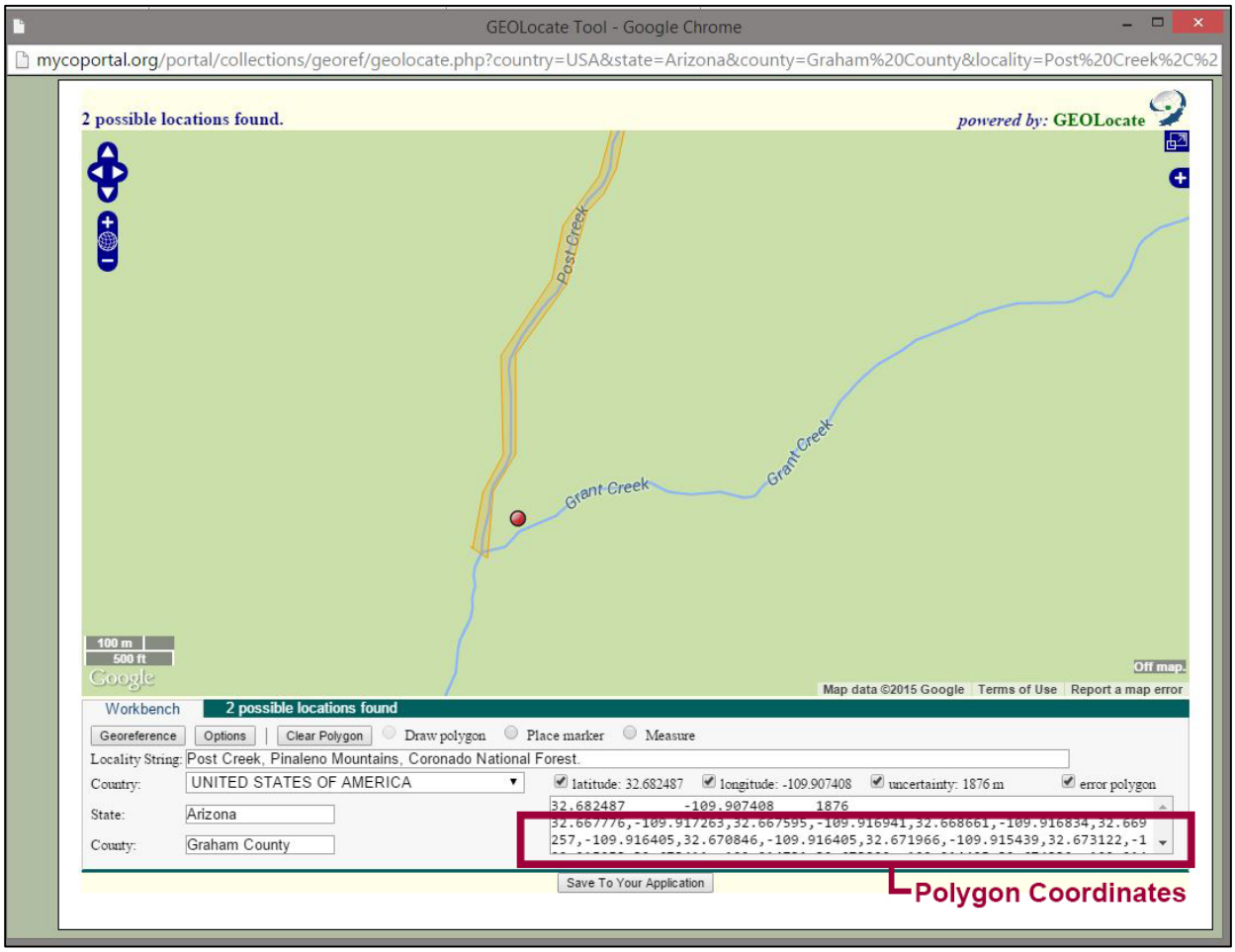
As I work my way back down the west side of Post Creek, notice the north side of the polygon is no longer overlapping on itself, neatly framed around Post Creek as I move southwest along the creek's bank.



I'm almost finished, drawing the polygon back to my original starting point at Post Creek and Grant Creek.



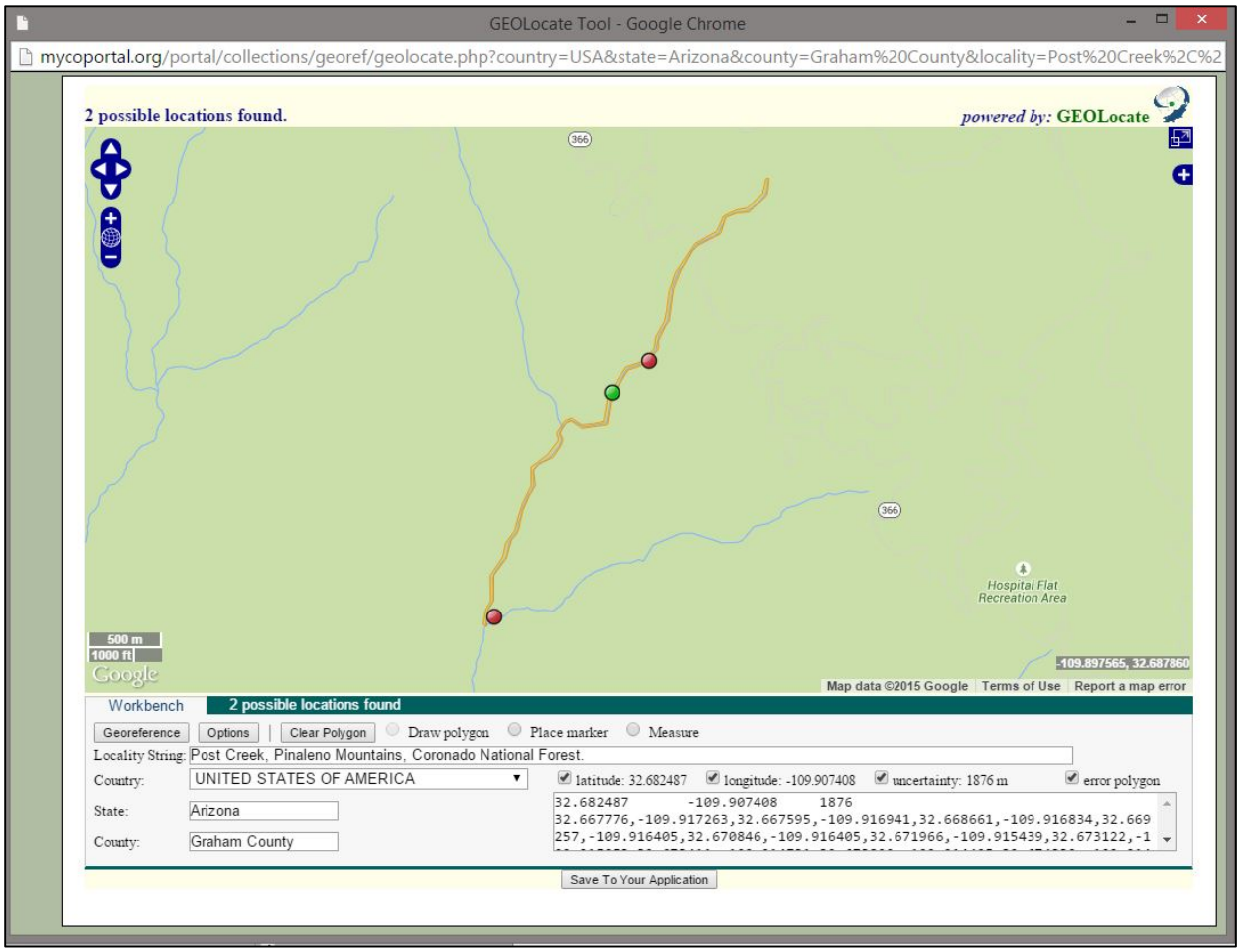
Now that I've reached the starting point, I double click on the original plotpoint of the polygon. The polygon will turn from blue to yellow, and all the pairs of polygon coordinates I just established will appear in the field located in the bottom right of the Workbench (see below).



Arrows have appeared in the field that contain the coordinate pairs, indicating there are numerous coordinates I can now view by scrolling down through that field (using those arrows).

These coordinate pairs represent every angle (or bend) of the polygon I just drew. In other words, every time I left clicked while drawing the polygon, GEOLocate transcribed that plotpoint as a lat/long coordinate pair, displaying the data in the Workbench.

But before I save any of my data, I want to zoom back out on the map so I can examine how accurately I drew the polygon around Post Creek.



As you can see, the polygon is tightly enclosed around the banks of Post Creek, following it in its entirety from the mouth at Grant Creek to the head of the creek.

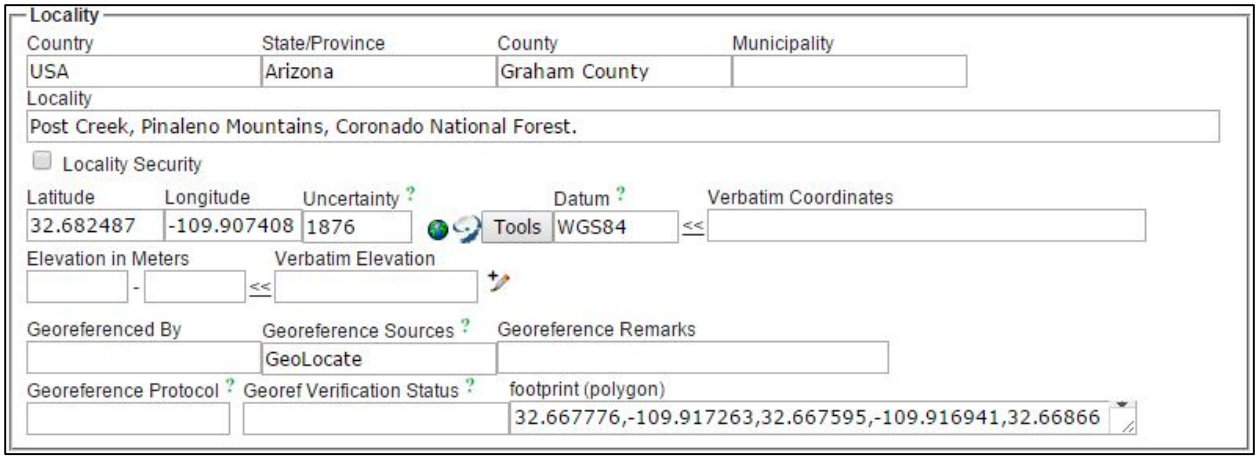
> **NOTE:** Be aware that if you draw the polygon outside of the green georeferencing plotpoint, GEOLocate will generate a warning at the top of the screen that reads: "Warning: green marker is outside the uncertainty polygon."

Now that I have established my polygon (along with the lat/long coordinates and uncertainty radius), and I've verified its accuracy, I'm ready to save all the locality data to the Portal.

Step #3: Saving Polygon Coordinates

It's important to emphasize that once I save the locality data to the Portal, I *cannot* access the polygon that I just drew. If I want to make any adjustments, I'll have to redraw the entire polygon from the beginning.

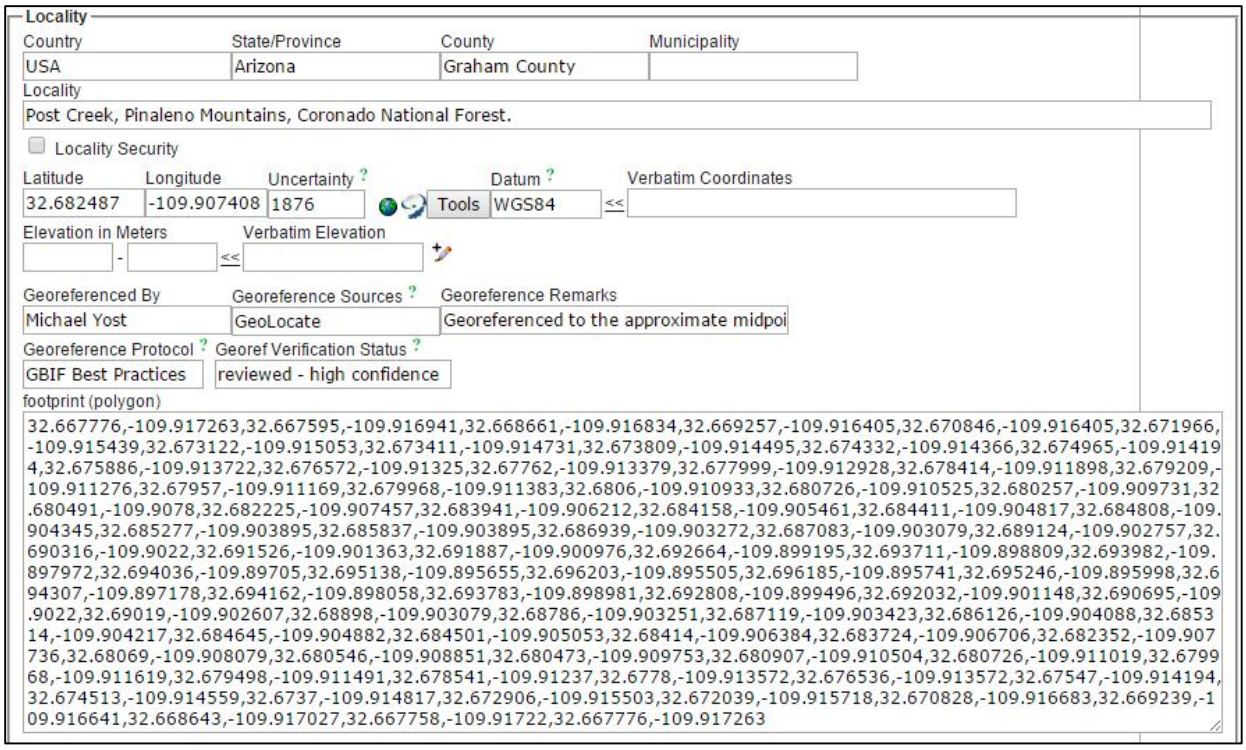
But after viewing the map in the previous step, I'm satisfied with the polygon I've established. So, I save the data by clicking on *Save To Your Application* at the bottom of the map. The map disappears, and all the locality data is transferred the Portal's editor page for the corresponding specimen (see below).



NOTE: Keep in mind that the locality data associated with this location is *not yet saved* to the Portal, merely displayed on the editing page.

The corresponding lat/long coordinates and uncertainty I established are displayed. In addition, the coordinate pairs corresponding to the polygon I drew are visible in the *footprint (polygon)* field.

If I left click and hold on the right side of the *footprint (polygon)* field, I can expand that field to view all of the coordinate pairs relating to the polygon I just drew (see below).



Before entering the rest of the locality information, I always save the data to the Portal. There have been instances where I've lost my connection to the server, losing all the polygon data and forcing me to draw the polygon once again. (Not fun!)

I scroll down to the bottom of the editing page and click on *Save Edits*.

Voila! I now have the following locality information saved to the Portal:

Lat: *32.682487*
Long: *-109.907408*
Uncertainty: *1876*

Datum: *WGS84*
(This will always be the Datum when using Google Maps in GEOLocate.)

Georeference Sources: *GeoLocate*
(This will automatically be generated by the Portal when using GEOLocate.)

footprint (polygon): *32.667776,-109.917263,32.667595,-109.916941,32.668661,-109.916834,32.669257,-109.916405,32.670846,-109.916405,32.671966,-109.915439,32.673122,-109.915053,32.673411,-109.914731,32.673809,-109.914495,32.674332,-109.914366,32.674965,-109.914194,32.675886,-109.913722,32.676572,-109.91325,32.67762,-109.913379,32.677999,-109.912928,32.678414,-109.911898,32.679209,-109.911276,32.67957,-109.911169,32.679968,-109.911383,32.6806,-109.910933,32.680726,-109.910525,32.680257,-109.909731,32.680491,-109.9078,32.682225,-109.907457,32.683941,-109.906212,32.684158,-109.905461,32.684411,-109.904817,32.684808,-109.904345,32.685277,-109.903895,32.685837,-109.903895,32.686939,-109.903272,32.687083,-109.903079,32.689124,-109.902757,32.690316,-109.9022,32.691526,-109.901363,32.691887,-109.900976,32.692664,-109.899195,32.693711,-109.898809,32.693982,-109.897972,32.694036,-109.89705,32.695138,-109.895655,32.696203,-109.895505,32.696185,-109.895741,32.695246,-109.895998,32.694307,-109.897178,32.694162,-109.898058,32.693783,-109.898981,32.692808,-109.899496,32.692032,-109.901148,32.690695,-109.9022,32.69019,-109.902607,32.68898,-109.903079,32.68786,-109.903251,32.687119,-109.903423,32.686126,-109.904088,32.685314,-109.904217,32.684645,-109.904882,32.684501,-109.905053,32.68414,-109.906384,32.683724,-109.906706,32.682352,-109.907736,32.68069,-109.908079,32.680546,-109.908851,32.680473,-109.909753,32.680907,-109.910504,32.680726,-109.911019,32.679968,-109.911619,32.679498,-109.911491,32.678541,-109.91237,32.6778,-109.913572,32.676536,-109.913572,32.67547,-109.914194,32.674513,-109.914559,32.6737,-109.914817,32.672906,-109.915503,32.672039,-109.915718,32.670828,-109.916683,32.669239,-109.916641,32.668643,-109.917027,32.667758,-109.91722,32.667776,-109.917263*

Notice that the uncertainty of 1876 meters has not changed, despite the fact that I've drawn a polygon around the creek to further truncate the field of uncertainty. This is because the uncertainty field only relates to the radius I drew in the first step.

As I mentioned earlier, there is currently no way to *see* that reduction in uncertainty in Symbiota with the embedded GEOLocate tool (even with the drawn polygon saved to the server). However, the polygon data is used by different programs to limit and visualize the uncertainty, increasing the locality utility for this specimen. These programs will be examined in more detail in an upcoming blog.

Now that the polygon data is saved, I move on to entering the additional locality information.

Step #4: Supplemental Locality Data

Let's walk through the rest of the locality information, entering the data as follows:

Georeferenced By: *Michael Yost*

(Enter your own name here, unless your name is Michael Yost, too. If so, I am pleased to meet you . . . or me. Pleased to meet me!)

Georeferencing Remarks: *Georeferenced to the approximate midpoint of Post Creek, with a polygon drawn around the entirety of Post Creek (at 100 meter zoom) from its mouth at Grant Creek to the southwest and the creek head to the northeast.*

(Just like when I was using the point-radius method, I want the explanation of how I georeferenced the specimen to be comprehensive yet concise.)

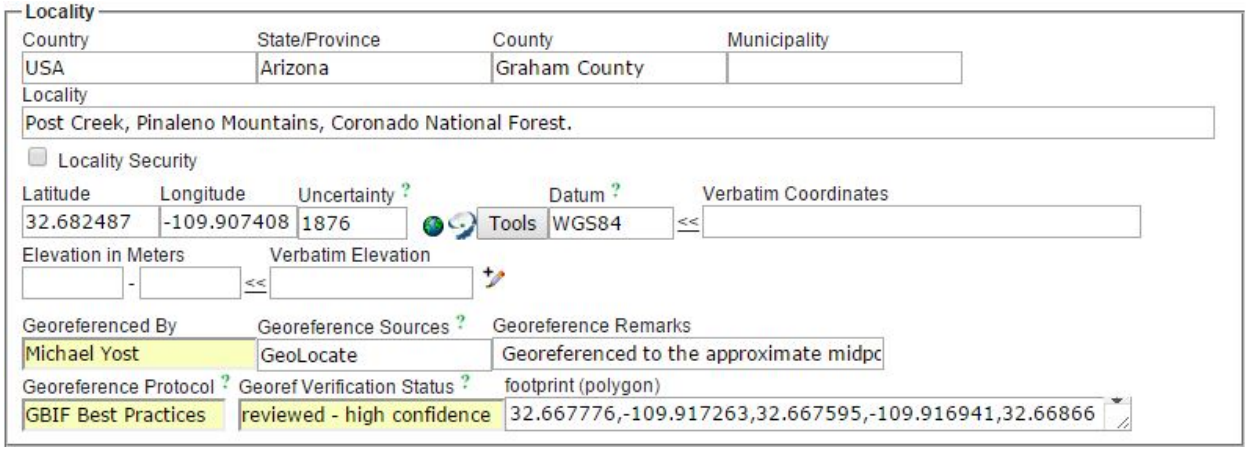
Georeferencing Protocol: *GBIF Best Practices*

(Always enter this into the protocol field.)

Georef Verification Status: *reviewed - high confidence*

(Always enter this into this field.)

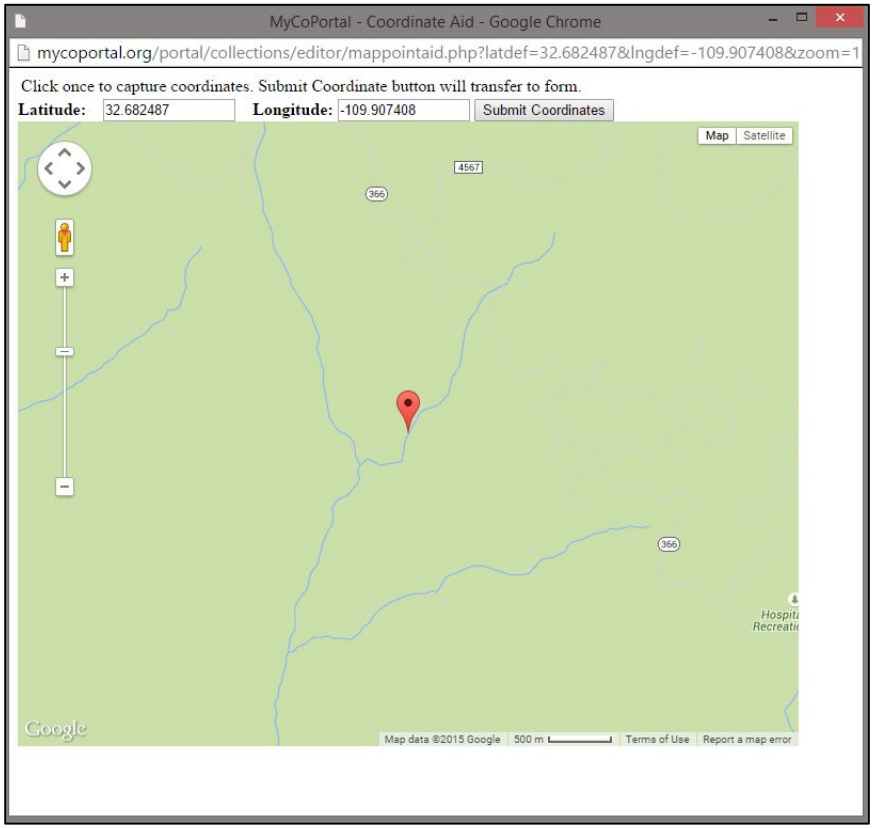
Now that I've finished entering all the locality information (see below), I'm ready to save the rest of the data to the Portal by clicking on the *Save Edits* button at the bottom of the editor page.



Step #5: Verification

Even though I can't visualize the polygon or uncertainty radius I just drew around Post Creek, I still want to verify my lat/long coordinates by clicking on the small earth located next to the *Uncertainty* field. The

Portal will open a map. If I zoom in, I can see that my lat/long coordinates are exactly where I set them in the first step of this process, the midpoint of Post Creek.



Eureka! I've successfully georeferenced a specimen using the polygon method!

Conclusion

As with my previous blog on batch georeferencing, this procedure is only a basic walkthrough of the polygon method. There are, of course, hundreds of potential localities, all of which can present their own set of challenges when drawing a polygon around a collection site.

I plan on drafting additional blogs to discuss some of those challenges. In the meantime, your own methods may vary—you may have even found a more efficient method!

Be sure to share your own experiences with the iDigBio community as we work together to establish proper polygon protocols.

Thank you for reading, and have a spectacular day!

Mike Yost
MaCC Project Assistant