**Geocoding Addresses in CSV Files**

**Smart Cville**

**Objective**

In this session, we look at ArcGIS Online’s capabilities for geocoding street addresses. Using the street addresses, ArcGIS Online adds longitude and latitude coordinates to the table so that we can now map our data. We will use the Virginia Geographic Information Network’s (VGIN) Virginia Composite Geocoding Service, which offers an excellent resource for mapping address data.

**Comma-Separated Variable (CSV) Files for Charlottesville Green Infrastructure (LEED Structures)**

The process starts with obtaining a comma separated value (CSV) file with street address data. CSV files, which contain data separated by commas, can be added to ArcGIS Online using the VGIN Composite Geocoding Service if they contain fields for street addresses, city, state, and optionally the zip code. (Different geocoders expect addresses in different formats, some use the complete address, city, state, and zip code in a single field.)

Here is an excerpt from City of Charlottesville’s Green Infrastructure (LEED Structures) CSV file that has been modified for use in this exercise …

OBJECTID,Entry,Owner,Address,City,State,Certification

1,Downtown Transit Station,City,615 East Water Street ,Charlottesville,VA,LEED Gold

2,ecoREMOD,City,608 Ridge Street,Charlottesville,VA,LEED Platinum

3,Smith Aquatic Center,City,1000 Cherry Ave,Charlottesville,VA,LEED Platinum

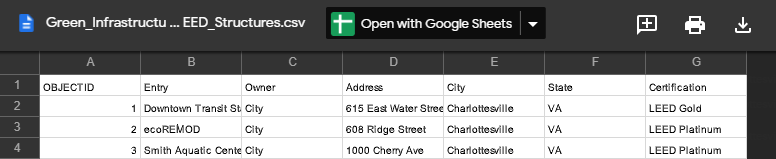
4,Charlottesville Area Transit,City,1545 Avon Street Extended,Charlottesville,VA,LEED Gold

5,Facilities Maintenance Building,City,315 4th Street NW ,Charlottesville,VA,LEED Gold

6,Fontaine Fire Station,City,2420 Fontaine Avenue,Charlottesville,VA,LEED Platinum

The first line is the header; it contains the field or variable names. The remaining lines are the data, with each field or variable separated by commas.

Download the file by entering the <https://drive.google.com/file/d/139c4h_fCYFKH4j8lzYTeZRomFIWW00kn/view?usp=sharing> in your browser.



Click the **Download Icon**, highlighted in red in image below, to save the spreadsheet to your system.

This process will vary depending on your browser. Remember where the file is located, as we will be using it later.

**VGIN Composite Geocoding Service Introduction**

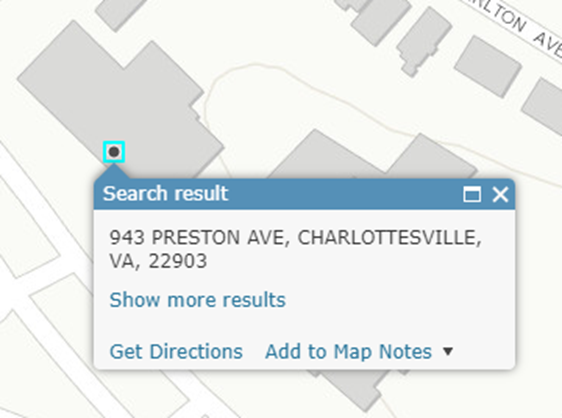
The VGIN Composite Geocoding Service is optimized for converting street addresses in Virginia to coordinate values that can be mapped. Composite geocoding services geocode data using multiple separate geocoders that can identify addresses using address points, interpolation of addresses along streets, zip codes, and jurisdiction names (from most to least precise). The geocoder will attempt to locate an address moving from the most to least precise source.

Here are examples of how a composite geocoder would work … they illustrate the different precisions that may be obtained depending on the information in the address and the type of geocoding available.

Address point (VGIN Geocoder)

CITY DRY CLEANERS, 943 PRESTON AVE, CHARLOTTESVILLE, ALBEMARLE, VA, 22901

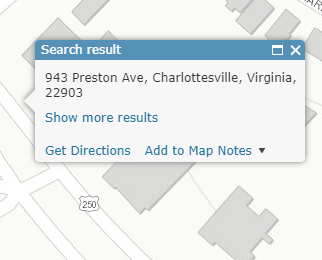
Address points are located on the building associated with the address. These are the most precise, but time-consuming and expensive to collect. Some, but not all, addresses in Virginia are mapped to address points.



Address interpolated along street (Esri Geocoder)

CITY DRY CLEANERS, 943 PRESTON AVE, CHARLOTTESVILLE, ALBEMARLE, VA, 22901

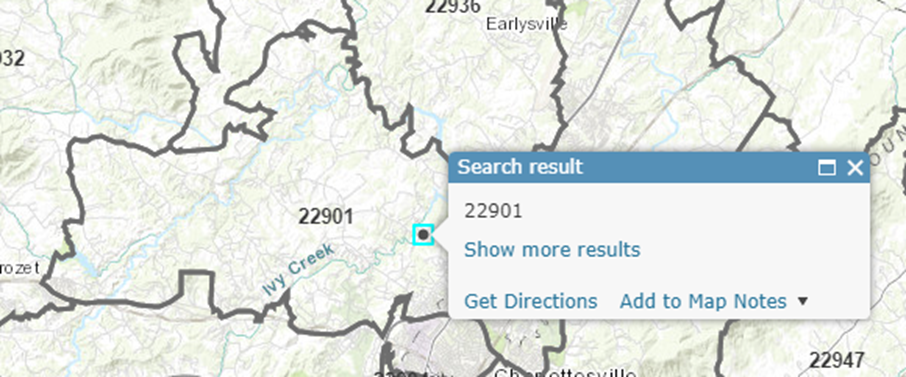
Addresses interpolated along a street use the numbered along each side of the street to estimate the location of the feature. The example below shows the interpolated result using Esri’s geocoder (which we will not use in this lesson). The Esri geocoder does not place a point and highlight it at the estimated location.



Zip code location (VGIN Geocoder)

CITY DRY CLEANERS, CHARLOTTESVILLE, ALBEMARLE, VA, 22901

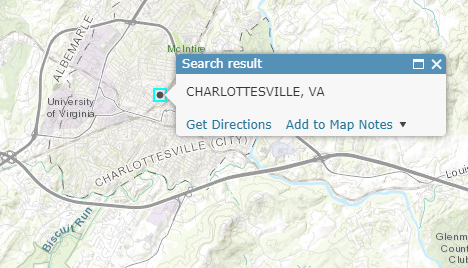
If the zip code is the only information or the most precise information in the address, the address will be mapped to a location in the zip code.



City or County location (VGIN Geocoder)

CITY DRY CLEANERS, CHARLOTTESVILLE, VA

If the city and state (or county and state) is the only valid information available, the address will be mapped to a location in the city (or county).



The Virginia Composite Geocoding Service web page is located at

<http://www.arcgis.com/home/item.html?id=64216809ef54479e95535fa822cbd6ee>

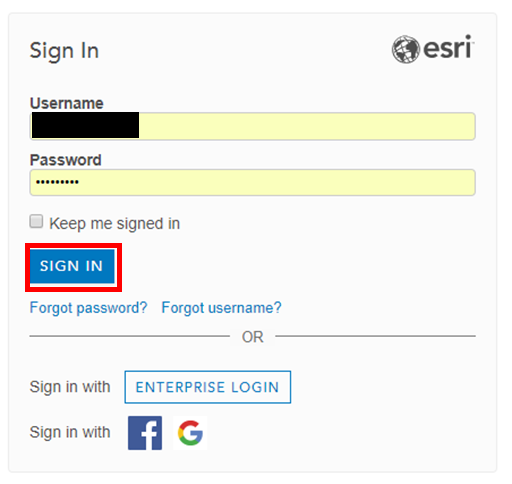
Details about the service are found at

<https://ftp.vgingis.com/Download/Documentation/VGIN_Geocoding_Service_Overview_and_Access_Instructions.pdf>

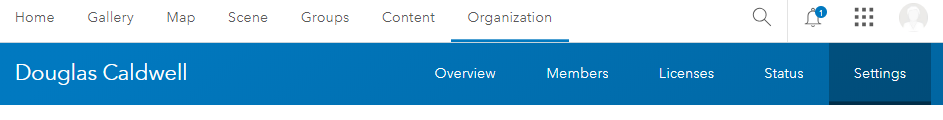
**Add the VGIN Composite Geocoding Service to Your Organization**

We need to add the VGIN Composite Geocoding Service to our organizational account in order to use it. Open a new browser tab and sign in to your ArcGIS Online account, open a new browser tab by typing “<https://www.arcgis.com/home/signin.html>” in the URL window to sign in. Enter your *Username* and *Password* and then click the **SIGN IN** button. **Sign in to your ArcGIS Online Account**

To sign in to your ArcGIS Online account, open a new browser tab and type <https://www.arcgis.com/home/signin.html> in the URL window to sign in. Enter your *Username* and *Password* and click the **SIGN IN** button.

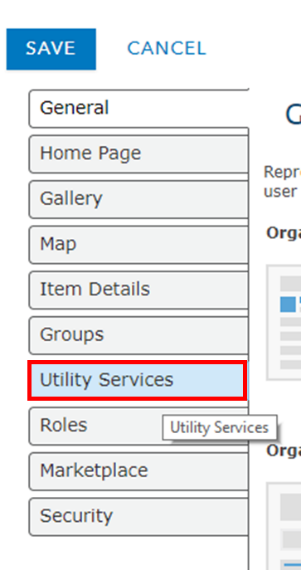


This will take you to your ArcGIS Online **Organization** section. Click on **Settings** to set your geocoded. Note: This screen and all others may vary depending on your screen resolution.

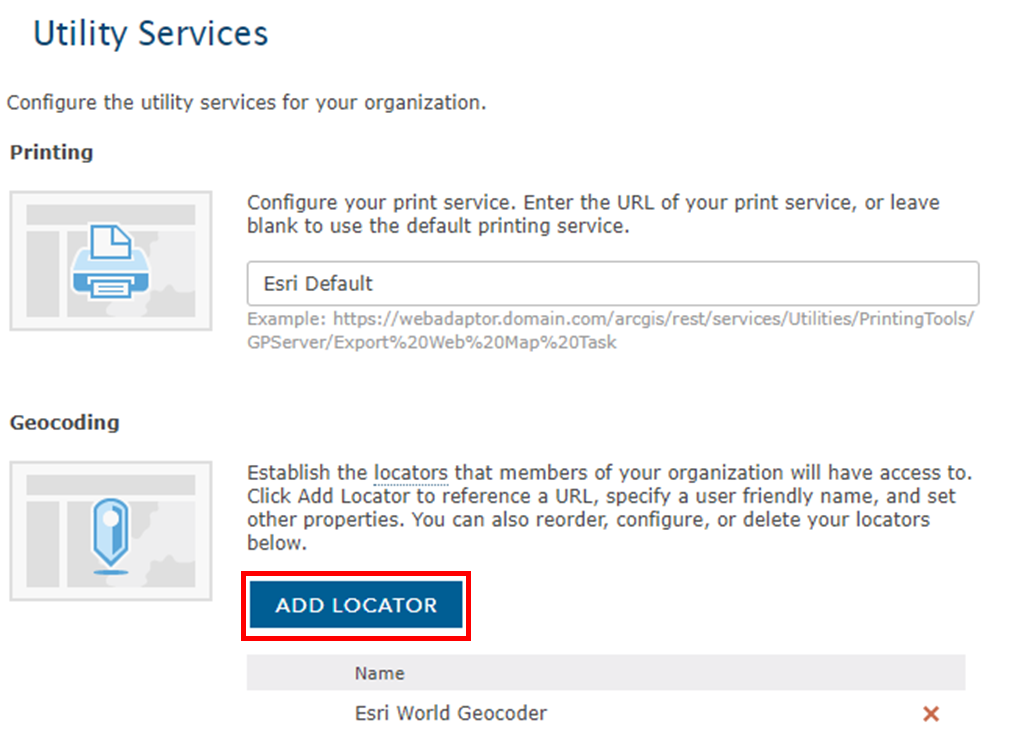


**Explore the Search Interface**

Click on the **Utility Services** menu option.

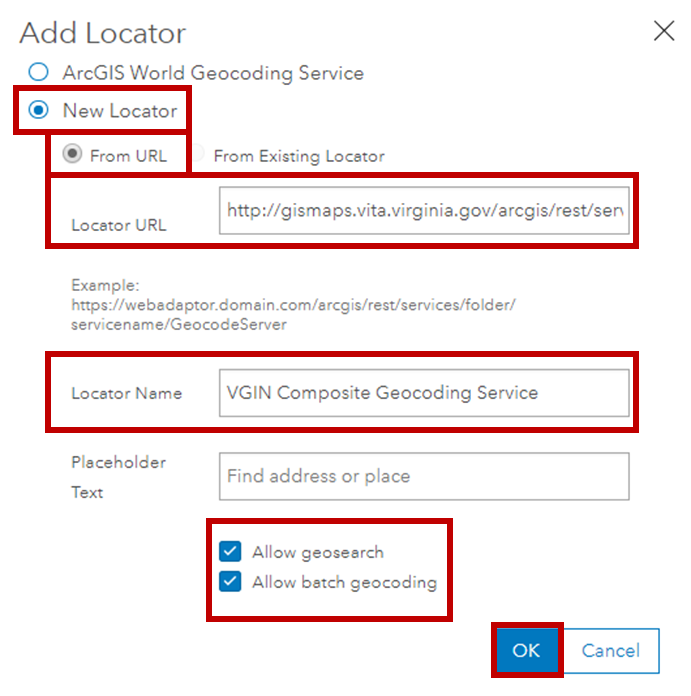


Scroll down to the *Geocoding* section and click on the **ADD LOCATOR** button.

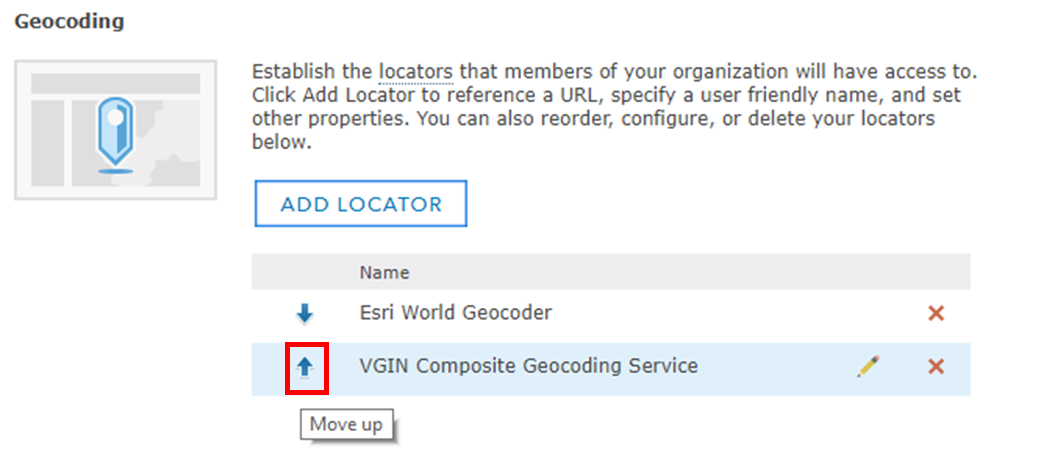


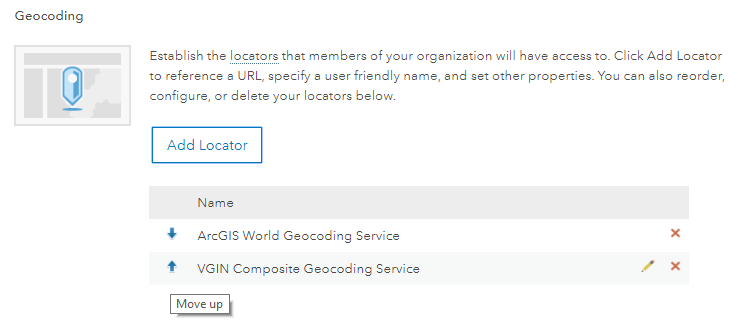
Check the **New Locator** button to expand the form.

In the expanded *Add Locator* form, make sure that the **From URL** is checked. Type “<http://gismaps.vita.virginia.gov/arcgis/rest/services/Geocoding/VGIN_Composite_Locator/GeocodeServer>” in the *Locator URL* text box. Type “VGIN Composite Geocoding Service” in the *Locator Name* text box. Leave the text in Placeholder Text as is and make sure that the **Allow geosearch** and **Allow batch geocoding** checkboxes are checked. Click the **OK** button.

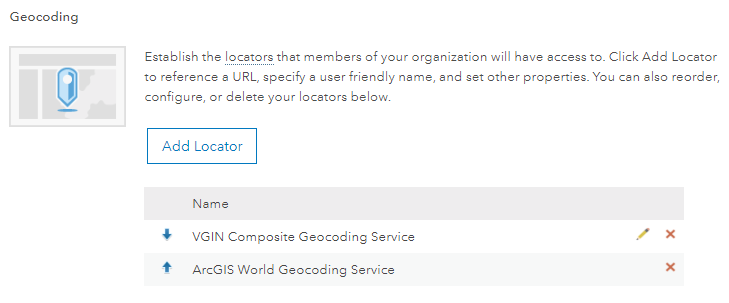


Then move the VGIN Composite Geocoding Service up to the primary service by clicking on the **Up Arrow** icon.

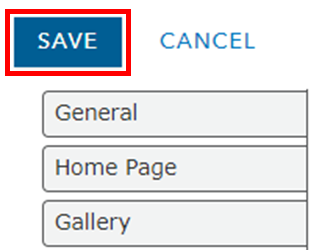




Your display should look like the image below, with the VGIN Composite Geocoding Service on top.



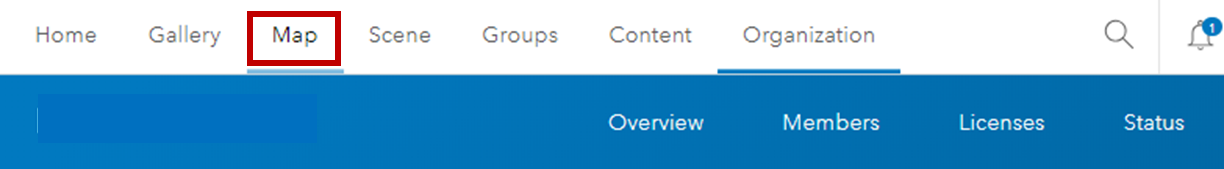
Click on the **SAVE** button in the upper left corner of the Utility Services web page.



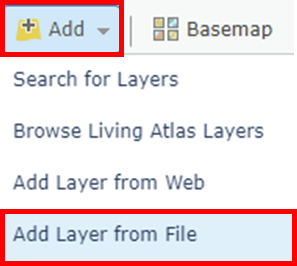
***Important Note: Using the VGIN Composite Geocoding Service will not use any credits , while using the ArcGIS World Geocoding Service will consume many of your ArcGIS service credits. Geocoding is particularly aggressive at consuming your credits, so it is important that you use the VGIN Composite Geocoding Service and not Esri’s service (unless you can afford the costs). If you follow the instructions above, you will only be using the VGIN Composite Geocoding Service and should not have a problem with credit usage. You can purchase additional credits if you wish to use the ArcGIS Online geocoder. A detailed discussion on Service Credits and the services that use them, see*** [***http://www.esri.com/software/arcgis/arcgisonline/credits***](http://www.esri.com/software/arcgis/arcgisonline/credits)

**Geocoding your Comma Separated Variable (CSV) File**

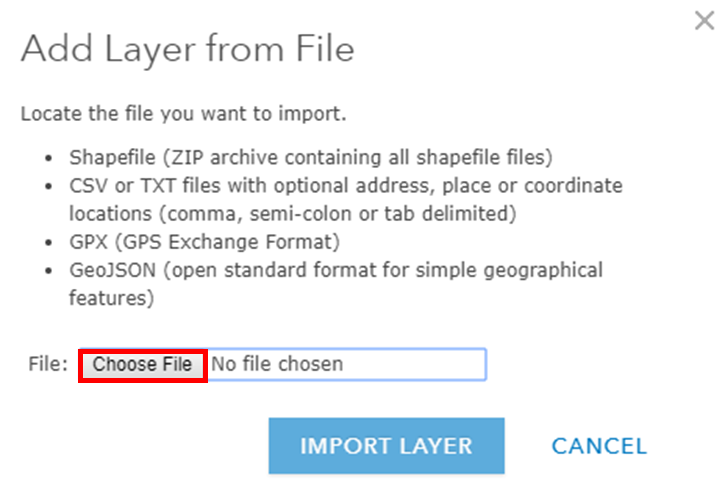
Return to your ArcGIS Online tab in the browser and click on **Map**.



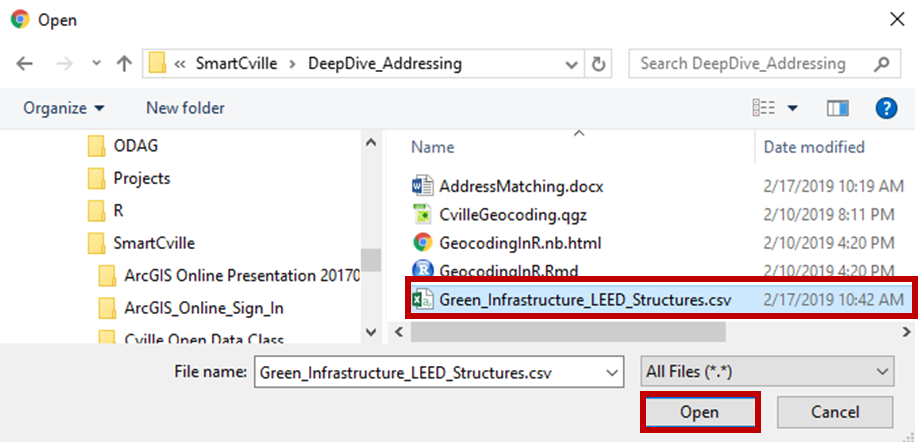
Click on **Add** and **Add Layer from File**.



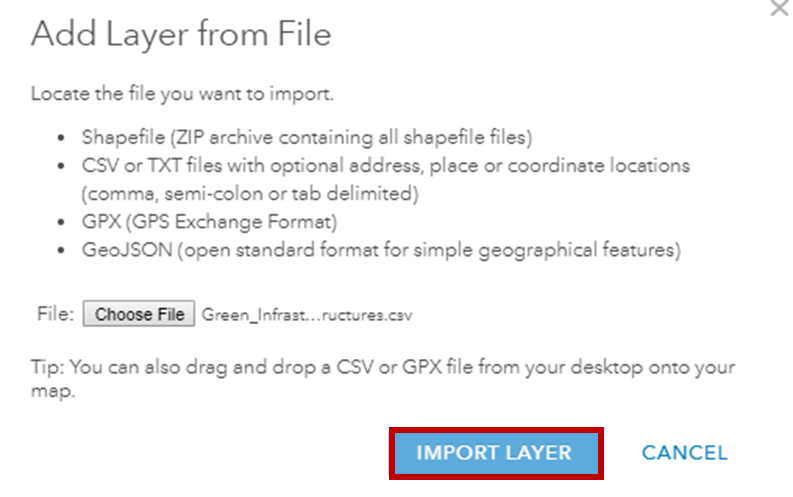
Click on **Choose File.**



Go to the directory where you download the Green\_Infrastructure\_LEED\_Structures.csv file and select the file. Click on the **Open** dropdown to save the file.



Click on **IMPORT LAYER** on the *Add Layer from File* form to begin the process of adding the layer.

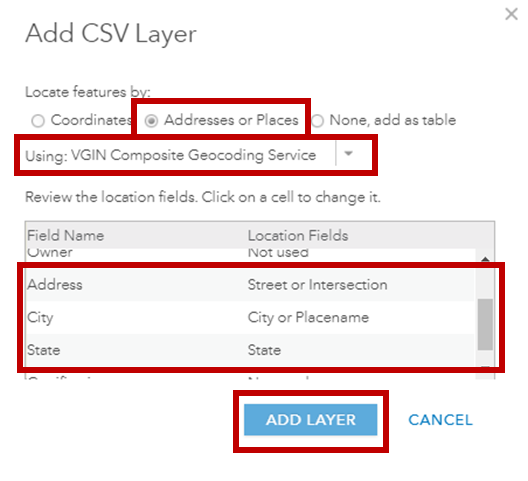


Using the *Add CSV Layer* form, we can match the fields in our CSV file to the fields used in the VGIN Composite Geocoding Locator. Make sure that the **radio button** for *Addresses or Places* is checked and that we are **Using** the VGIN Geolocator. Select the VGIN Geolocator from the dropdown. Note: it is very important to use the VGIN Geolocator if we don’t want to use any credits.

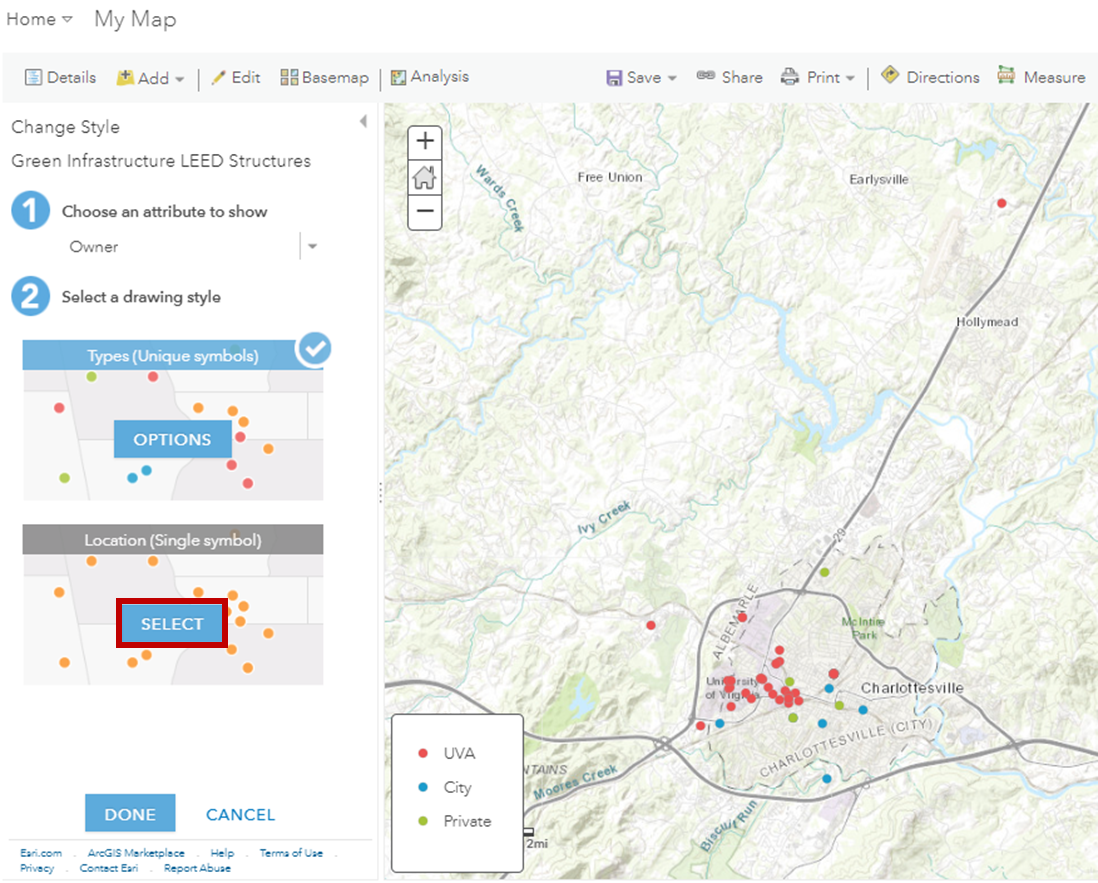
You can review the location fields. The names in the *Field Name* column are the names in our CSV file. The names in the *Location Fields* column are the names that the geocoder is using. Some of these may be matched automatically.

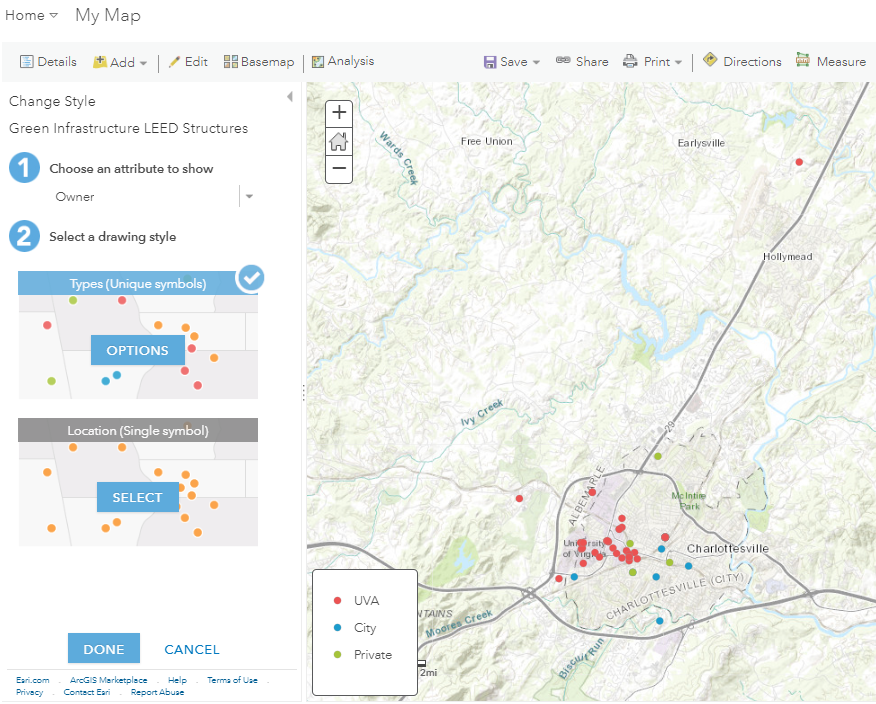
Click on the *Location Fields* to select the appropriate match to the *Field Name*. Map the Address field name to the dropdown **Street or Intersection** Location Field; City to the dropdown **City or Placename**, and *State* to the dropdown **State** Location Field. This file does not have any Zip Codes. If Zip Codes were available, they would improve the geocoding. If we used different field names, we would manually match our Field Names to the Location Fields.

Click the **ADD LAYER** button to add the layer to our map.

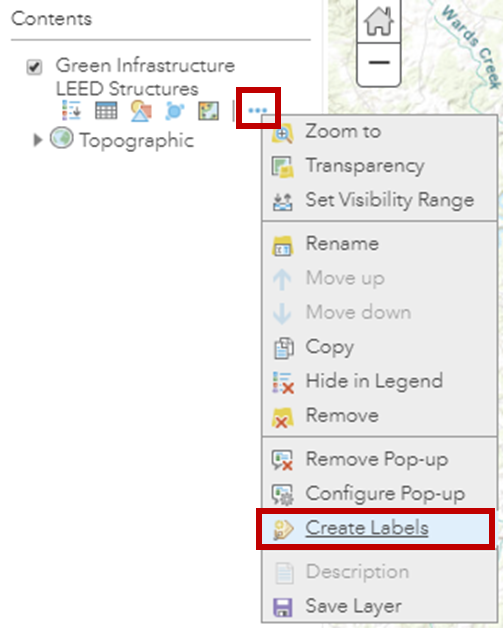


Your map will appear. Click on the **Select** button in the *2. the Location (Single symbol) option*. Then click the **DONE** button. This will display all the circles as black circles.





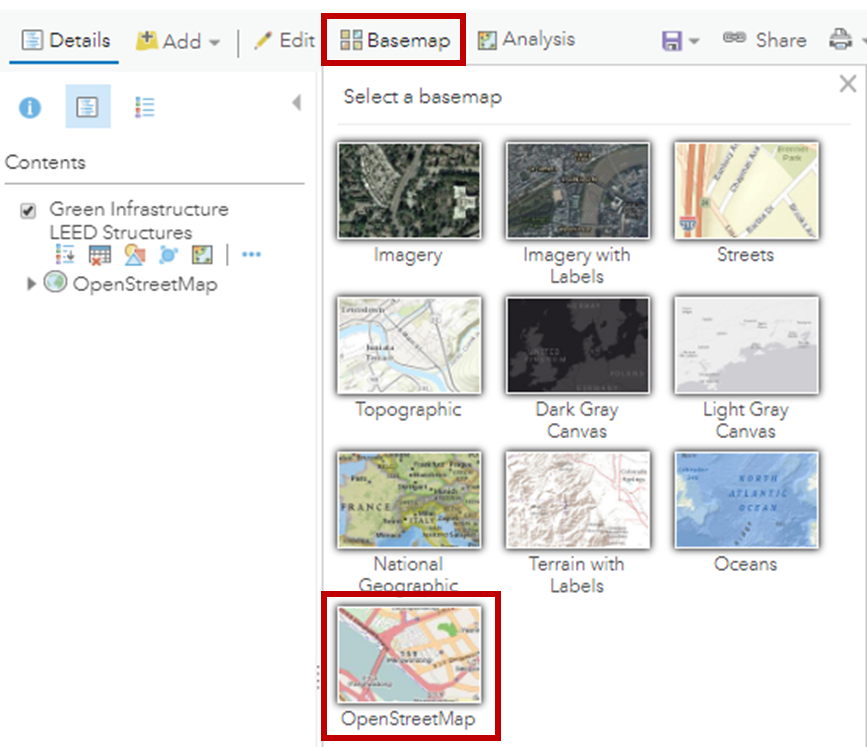
Let’s add labels to the points. Click on the **three horizontal dots** (More Options) next to *Green Infrastructure and LEED Structures* in the *Table of Contents* and then click **Create Labels**.



When the *Label Features* pane appears, click the **OK** button to use the default values.



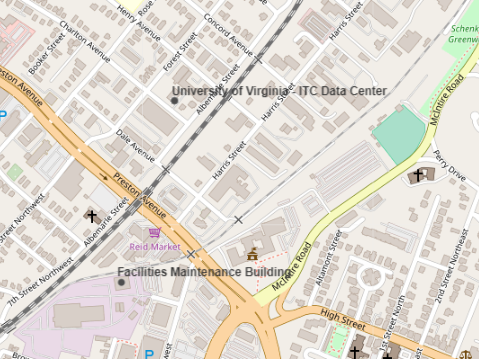
Let’s change the Basemap to the OpenStreetMap basemap. OpenStreetMap is a crowdsourced data base that has more detailed information on addresses and buildings. This will let us compare our results.



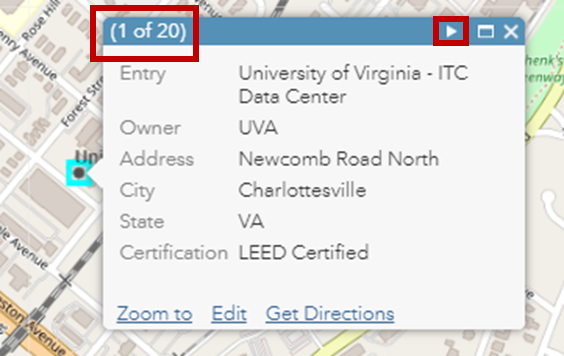
Zoom in the UVa Campus and look at some of the results. Click on the black dot symbols to see all the information for each Building. The results in this area seem pretty good.



Now move over to the area near Preston Avenue and Harris Street. You will see a feature labeled University of Virginia – ITC Data Center.



Click on the dot associated with the University of Virginia – ITC Data Center to open up the pop-up for this location.



Notice that 20 of the 59 records were geocoded to this location. This is the generic location for Charlottesville. Any address that cannot be matched to the *Address* field will be matched to the next highest geography, which is the city in this case. You can click on the right pointing arrow to loop through these records and see which could not be located to the address level. Many of these are street addresses that include the street name, but do not include a street number. Without a number, the geocoder is not able to locate the address along the street.

Our geocoder matched all of our addresses, but the results were not what we expected. That said, by examining the records, it is easy to see how we can greatly improve our matching … find out the street addresses of the buildings and update the CSV file.

Geocoding is an imperfect process. Different geocoders will give different levels of accuracy and it is important to review your address lists and check to ensure that addresses follow addressing standards and include as much of the address information as possible.

**Extra Credit**

**1. Comparing Geocoders**

The objective of this extra credit exercise is to compare results among different geocoders. *Cville LEED Structures* is an ArcGIS Online map that contains the original Cville data, the address data geocoded in ArcGIS Online using the VGIN geocoder, and the address data geocoded in QGIS using the OpenStreetMap Nominatim geocoder. You can access the map at …

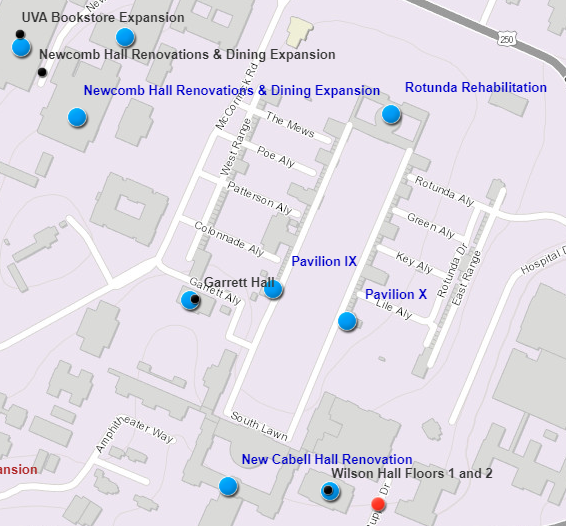
Short URL  
<https://arcg.is/1quamC>

Long URL (try this if the short URL does not work)  
<https://www.arcgis.com-home/webmap/viewer.html?webmap=f1b73966b6bb4e67acfadff568d97b82&extent=-78.5211,38.0282,-78.4894,38.0414>

The map should open up to the UVa campus. The symbols are as follows:

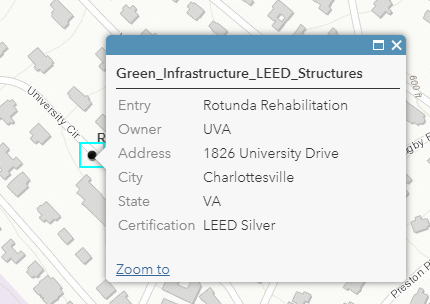
Small Black Circles Geocoded from Addresses using VGIN geocoder and ArcGIS Online  
Medium Red Circles Geocoded from Addresses using Nominatim Geocoder and QGIS  
Large Blue Circles Original Hand Coded Data from the City of Charlottesville

This is a good chance to explore the data. You will see all circles on top of or close to each other in cases where the geocoders match. In some areas, like the Lawn at UVa, only the city’s hand-matched symbols are shown.

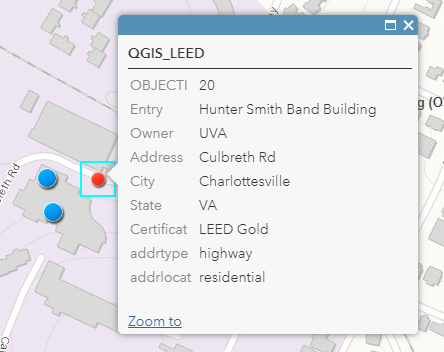


These locations could not be identified by the geocoders using the addresses, so you will have to search elsewhere on the map to find them.

Take a look at this example for the VGIN geocoding of 1826 University Drive for the Rotunda Rehabilitation. This seems pretty far from the actual location. Can you guess what the geocoder did?



There is another interesting example of the difference between geocoders. We saw earlier that the VGIN geocoder would locate features at the city level coordinate if it could not determine the numeric component of an address. In the figure below, the QGIS geocoder locates the feature in the middle of the street feature if no address number is present. This is closer to the actual location, but may not actually be near the location. You should check for this type of artifact when reviewing your data.



Explore the map and click on the circles to see the addresses. See if this will help you understand why some geocoded addresses are close to the City of Charlottesville’s hand coded addresses and others are not.

You should not rank the geocoder quality from the LEED structure example, given the large number of poorly formed addresses. With better formed addresses, the QGIS Nominatim geocoder can give very good results, even better than the VGIN geocoder, with building level matches.

Again, this shows the importance of having high quality address information if you are interested in obtaining high quality results. Also, you should check your data, correct any errors, and consider checking multiple geocoders.

**Additional Resources**

**Related Help**

ArcGIS Online Help for loading CSV, TXT, and GPX files

<https://doc.arcgis.com/en/arcgis-online/reference/csv-gpx.htm>

**Geocoding**

Geocoding: What is Geocoding? (University of Illinois)

<http://guides.library.illinois.edu/Geocoding/About>

Geocoding (Wikipedia)

<https://en.wikipedia.org/wiki/Geocoding>

Three Standard Geocoding Methods

<https://www.directionsmag.com/article/3406>

**Contact Information**

Please send any corrections or suggestions to

Doug Caldwell ([caldwellinva@verizon.net](mailto:caldwellinva@verizon.net))

February 17, 2019