Geocoding with Nominatim
How to Improve Geocoding Results

Geocoding: What is it?

Geocoding is the process of converting addresses (like "1600 Amphitheatre Parkway, Mountain View, CA") into geographic coordinates (like latitude 37.423021 and longitude -122.083739), which you can use to place markers on a map, or position the map.

Geocoding in the NLP Suite

The NLP Suite provides two free approaches to geocoding locations, via Nominatim and via Google. These options are available from GIS_main.

Other NLP Suite scripts (e.g., GIS_Google_Earth_main, SVO_main), also have geocoding capabilities. GIS_Google_Earth_main and SVO_main rely on Nominatim as the default geocoding option. If you wish to use Google for geocoding, please, use the GIS_main script.

Whichever option is chosen for geocoding (Nominatim, Google) the NLP Suite will carry out the geocoding automatically. But… other options are available. See below.

Nominatim

Nominatim is an open-source, freeware geocoding tool. Nominatim uses **OpenStreetMap** (OSM) data to find locations on Earth by name and address (geocoding).

How Nominatim Works

See https://nominatim.org/release-docs/latest/api/Search/ for the arguments used by the Nominatim script.

For every location passed to the Nominatim server for geocoding (for example, Athens), it will return up-to-fifty (default 10) best-matched OSM (OpenStreetMap) records with latitude and longitude values and address. The NLP Suite picks automatically, the top OSM record.

Thus, the sentence:

When I visited Athens, I thought it was a beautiful city. But Rome was really in my heart.

The two cities are mapped in Greece and Italy since the two top Nominatim OSM records are indeed:

Location	Latitude	Longitude	Address
Athens	37.98394	23.72831	Athens, Municipality of Athens, Regional Unit of Central Athens, Attica, 10431, Greece
Rome	41.89332	12.48293	Rome, Roma Capitale, Lazio, Italy

What if you want Athens and Rome as cities in the United States? There are dozens of Rome, for instance, in the United States https://whalebonemag.com/us-cities-named-rome/

Let's rewrite the sentence above:

When I visited Athens, in the United States, I thought it was a beautiful city. But Rome, in the United States, was really in my heart.

Now, Athens is mapped in Georgia and Rome in New York State.

| Location	Latitude	Longitude	Address
Athens, United States	33.95977	-83.3764	Athens, Athens-Clarke County, Georgia, United States
Rome, United States	43.21285	-75.4557	Rome, City of Rome, Oneida County, New York, 13440-5621, United States

Country bias widget

You obtain the same result with the same sentence without the country specification, the United States, but with the United States of America selected as the Country bias

When I visited Athens, I thought it was a beautiful city. But Rome was really in my heart.

Location	Latitude	Longitude	Address
Athens	33.95977	-83.3764	Athens, Athens-Clarke County, Georgia, United States
Rome	43.21285	-75.4557	Rome, City of Rome, Oneida County, New York, 13440-5621, United States
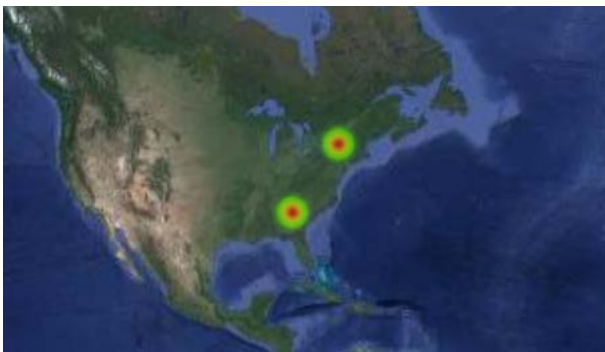
Area widget

What if the Rome you want to geocode and map is the one in Georgia? Let's go back to the sentence

When I visited Athens, I thought it was a beautiful city. But Rome was really in my heart.

In the Area widget of the GIS_main GUI, enter the coordinates for a rectangle encompassing Georgia, in the United States. Two sets of lat/long coordinates must be entered, the first for the top left-most point of an ideal rectangle and the second for the lower right-most point in the form (34.98527, -85.59790), (30.770444, -81.521974).

Now, both Athens and Rome will be mapped in the state of Georgia, U.S. regardless of Country bias selected and/or ticking the Restrict checkbox.

Location	Latitude	Longitude	Address
Athens	33.95977	-83.3764	Athens, Athens-Clarke County, Georgia, United States
Rome	34.25704	-85.1647	Rome, Floyd County, Georgia, United States

Restrict checkbox

The Country bias and Area widgets tell Nominatim to use the selected Country bias and Area rectangle as *preferred* locations.

If, however, you tick the Restrict checkbox

Only OSM records falling in the selected rectangle will be returned.

Editing the LOCATIONS csv file

Let's run the following sentence through the NLP Suite GIS pipeline.

During my trip to Europe, last year, I visited several countries, from England, to France, Belgium, Italy, and Greece. I was quite a tour de force, from London to Paris, Brussels, Rome and Florence, and Athens.

The algorithms will return a csv file containing a set of locations extracted from the input text by the Stanford CoreNLP NER (Named Entity Recognition) annotator (LOCATIONS, caps, in the filename).

Location	NER Value
Europe	LOCATION
England	COUNTRY
France	COUNTRY
Belgium	COUNTRY
Italy	COUNTRY
Greece	COUNTRY
London	CITY
Paris	CITY
Brussels	CITY
Rome	CITY
Florence	CITY

It will also return the csv file of geocoded values…

Location	Latitude	Longitude	Address
Athens	37.98394	23.7283052	Athens, Municipality of Athens, Regional Unit of Central Athens, Attica, 10431, Greece
Belgium	50.64028	4.6667145	Belgium
Brussels	50.84656	4.351697	Brussels, City of Brussels, Brussels-Capital, 1000, Belgium
England	52.53102	-1.2649062	England, United Kingdom
Europe	51	10	Europe
Florence	43.76987	11.2555757	Florence, Metropolitan City of Florence, Tuscany, Italy
France	46.60335	1.8883335	France
Greece	38.99537	21.9877132	Greece
Italy	42.63843	12.674297	Italy
London	51.50732	-0.1276474	London, Greater London, England, SW1A 2DX, United Kingdom
Paris	48.85889	2.320041022	Paris, Ile-de-France, Metropolitan France, France
Rome	41.89332	12.4829321	Rome, Roma Capitale, Lazio, Italy

And it will map the geocoded values.

Besides manipulation the Country bias and Area widgets, you can also edit the LOCATIONS file directly to alter the geocoded values. If you do not have too many locations to edit, this approach will give you the exact results you want.

Location	NER Value
Europe	LOCATION
England	COUNTRY
France	COUNTRY
Belgium	COUNTRY
Italy	COUNTRY
Greece	COUNTRY
London	CITY
Paris, United States	CITY
Brussels	CITY
Rome, New York, United States	CITY
Florence	CITY
Athens, Georgia	CITY

Location	Latitude	Longitude	Address
Athens, Georgia	33.9597677	-83.3764	Athens, Athens-Clarke County, Georgia, United States
Belgium	50.6402809	4.666715	Belgium
Brussels	50.8465573	4.351697	Brussels, City of Brussels, Brussels-Capital, 1000, Belgium
England	52.5310214	-1.26491	England, United Kingdom
Europe	51	10	Europe
Florence	43.7698712	11.25558	Florence, Metropolitan City of Florence, Tuscany, Italy
France	46.603354	1.888334	France
Greece	38.9953683	21.98771	Greece
Italy	42.6384261	12.6743	Italy
London	51.5073219	-0.12765	London, Greater London, England, SW1A 2DX, United Kingdom
Paris, United States	33.6617962	-95.5555	Paris, Lamar County, Texas, 75460, United States
Rome, New York, United States	43.2128473	-75.4557	Rome, City of Rome, Oneida County, New York, 13440-5621, United States