

What You Need to Know About Geocoding for Ridematching*

*includes bonus material on troubleshooting difficult addresses

Ross L. Edgar

Principal Geographic Information System Analyst

December 19, 2023

1

1

What is geocoding?

A procedure to derive location-based data from text

- Geocoding is the process of computing coordinates that correspond to some location on the Earth.
- We use computer software for speed although the process can be performed with accurate paper maps and a calculator (are your spherical trigonometry skills up to par?).
- We start with address text input that looks like
`777 NORTH CAPITOL ST NE, WASHINGTON, DC`
- When processing goes normally, the output resembles –
`77.00850521472817, 38.8998518867104`

2

2

Why is geocoding important?

It's the starting point for Location-Based Services (LBS)

- A substantial portion of our work is providing LBS to commuters.
- All commuting facilities and options—carpools, transit, park and ride lots, other commuters—are recommended based on a requesting commuter's location on the Earth.
- The system computes routes from origin coordinates to destination coordinates with turn-by-turn directions for commuters seeking carpools.
- For these LBS to function as expected, the TDM system must store mostly accurate geocodes for registrants seeking our help.

3

3

Up to 10%

The proportion of commuters in the database at any given time who have signed up for LBS but whose geocodes are inaccurate—or simply missing

4

4

Spurious geocodes—statistics

Study of the TDM database performed in April 2021

- There were 2,615 records that either lacked or had unusable geocodes left over from when we depended on Google for geocoding.
- The **GR** and **U**nified **A**ddress **L**ocator service hosted by ccgis.commuterconnections.org (**GRU_AL**) computed geocodes for almost 83% of these records in “batch mode” without human intervention.
- The remaining 17% needed analysis and human intervention. After going through those 449 records, there were just 8 records that could not be matched to a street or place in the data the service uses to compute coordinates.
- This is a surprisingly good result.

5

5

Reasons for missing or incorrect geocodes

#1: Malformed input for street address

- A common problem is the street address field contains a Post Office box number. The address locator cannot compute geocodes for post office boxes because they aren't in any of the reference data (the collection of streets, single house point features, and POIs against which we're geocoding).
- At least one user typed something like “3357 Court Something” when (s)he meant to type “3357 Something Court”. Besides transposing entire words, users often transpose digits in the house number, discussed next.

6

6

Reasons for missing or incorrect geocodes

#1: Malformed input for street address (continued)

- Many users did not provide a house number, or they provided one that doesn't exist, at least on our current map. Part of the geocoding process is to search the address range fields in the reference data to find a range that contains the house number. If the input is incorrect or absent, it won't be found.
- Some users must have gotten tired. One user entered "Shenandoah Va" for the street name instead of "Shenandoah Valley Drive". Since the current TDM System software provides address suggestions now, we can expect fewer mistakes like this.

7

7

Reasons for missing or incorrect geocodes

#2: Incorrect, misspelled, or missing city

- Less common than street address input mistakes.
- The service does not need the zip code, but it *does* need the city (post office).
- In one record, a user entered the city as "Ft Meade". This input doesn't work because the city name is "Fort George G Meade".

8

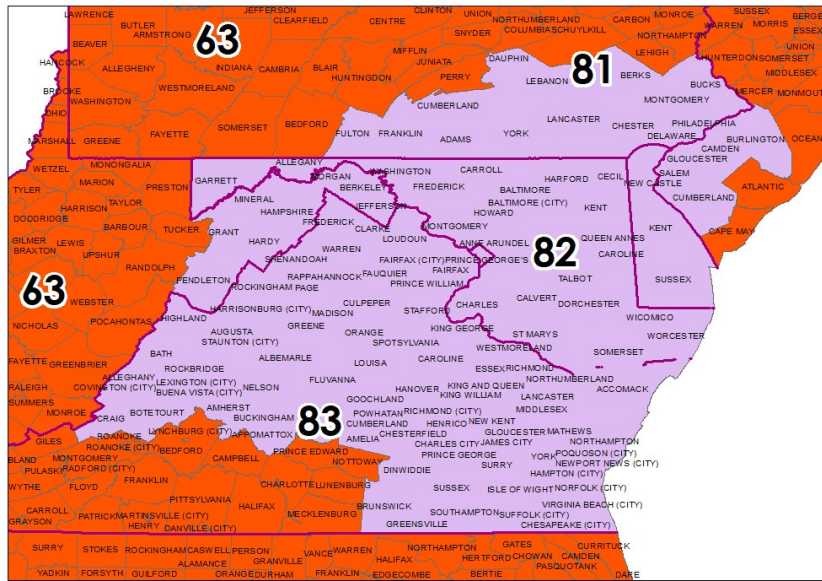
8

Reasons for missing or incorrect geocodes

#3: User does not commute in the Commuter Connections area

- The software cannot compute geocodes for places like Dallas, PA, Matthews, NC, Pulaski, VA, and Westfield, NJ because they are outside the area covered by the reference data.
- The coverage area changes from time to time. To build the current version of GRU_AL, some counties in far northern New Jersey were removed and several counties in southern Virginia from Richmond south to the state line were added.

GRU_AL coverage area today



Reasons for missing or incorrect geocodes

#4: Outdated reference data

- Market for houses has cooled a bit but it's still busy.
- Home builders eager to meet demand have caused many new streets to appear on the Earth. Maps are always in danger of falling behind reality on the ground.
- Streets that were on our maps, but missing names or address ranges, have been given names and address ranges.
- Local governments readdress streets and occasionally change street names or alternate names.

11

11

Troubleshooting erroneous
address input data:

An example.

12

12

Troubleshooting #1 of 9

HOME ADDRESS

Share my home address with other registered commuters

Street Address: * Apt #:

City: *

State: * Zip Code: *

Home Jurisdiction:

Data entered by the commuter looks fine, except it isn't:

ADDRESS TYPE - NAME	ADDRESS	VIEW	START	DESTN.	DELETE
HOME -	1824 EDGEWOOD RD PARKVILLE MD 21234	Invalid Geocodes!	<input type="radio"/>		<input type="radio"/>
WORK - ANNE ARUNDEL COMMUNITY COLLEGE	101 COLLEGE PKWY ARNOLD MD 21012		<input checked="" type="radio"/>		<input type="radio"/>

13

13

Troubleshooting #2 of 9

Resources to help us find and fix the problem

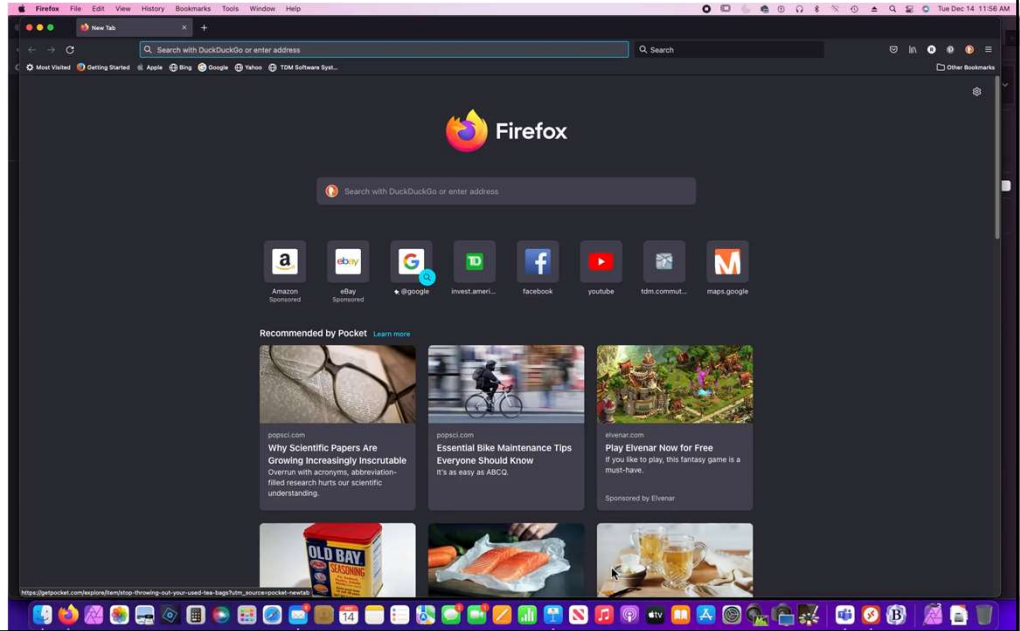
- USPS web site's Quick Tools -> Look up a Zip Code™
<https://www.usps.com>
- Commuter Connections ArcGIS Server
https://ccgis.commuterconnections.org/arcgis/rest/services/address_locators/GRU_AL/GeocodeServer/findAddressCandidates
- TDM System's Suggest As You Type functionality
- Google Maps
maps.google.com
- commutersupport@mwcog.org

14

14

Troubleshooting #3 of 9

First, try the USPS web site

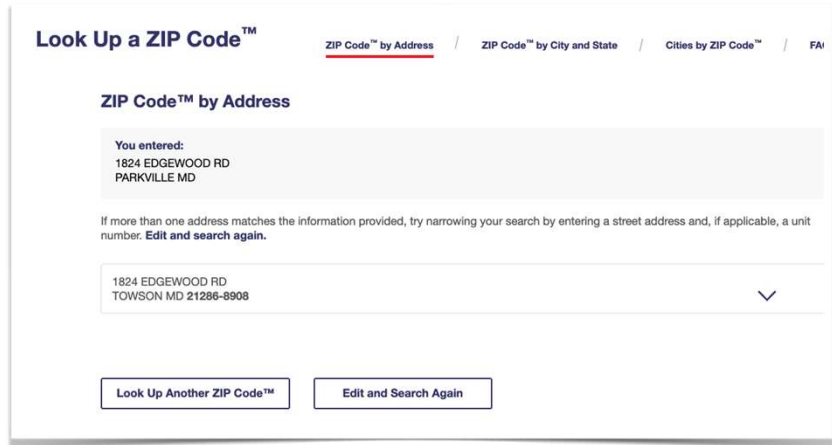


15

Troubleshooting #4 of 9

Sometimes the problem just dissolves

- This one turned out to be easy because the USPS has told us the address is not in Parkville — it's in Towson.
- Furthermore, the zip code is not 21234 as it was entered by the commuter, it's really 21286.



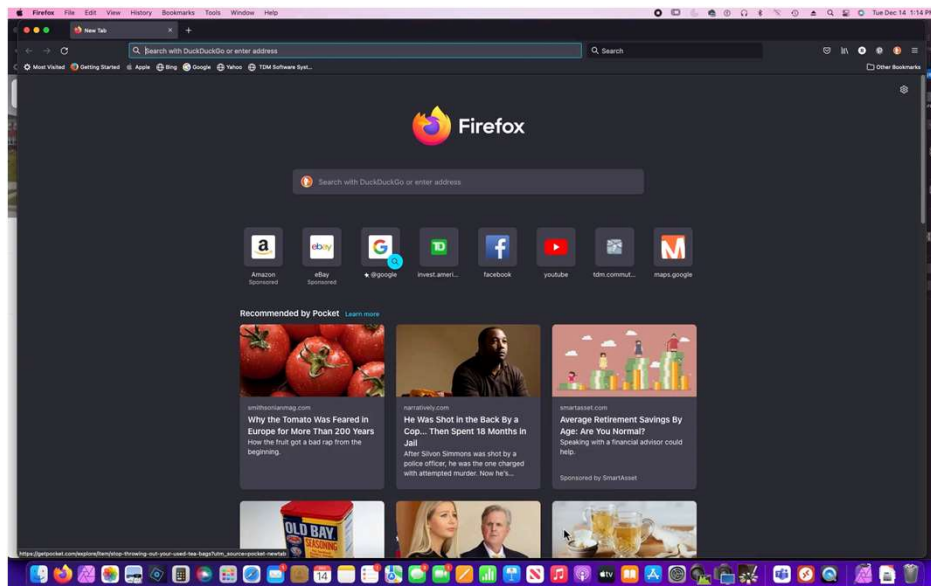
16

16

Troubleshooting #5 of 9

Optional step — Confirm result using the ArcGIS Server

- This is the same server the TDM System uses to obtain geocodes.
- You can then copy/paste the coordinates into the TDM web application.



17

Remember:

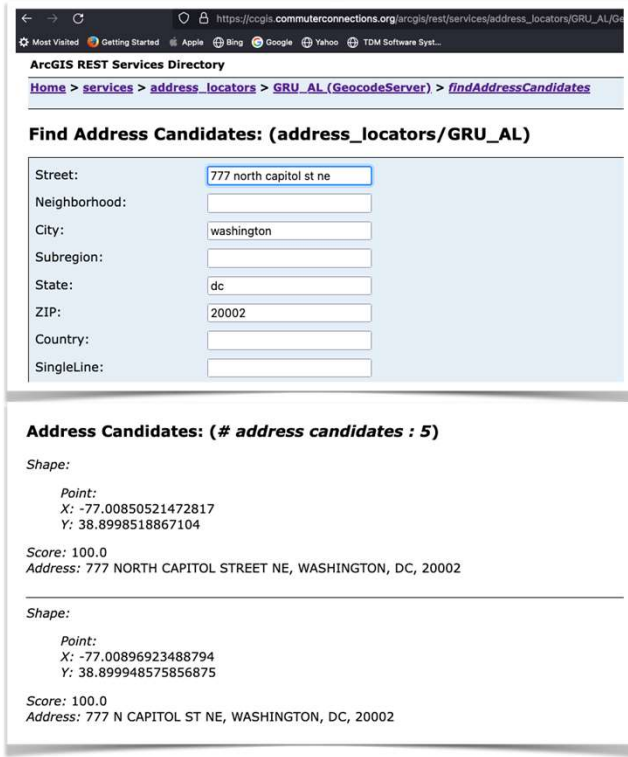
When using the Commuter Connections ArcGIS Server, it's easiest to type the address into the SingleLine widget on one line:

1824 EDGEWOOD RD , TOWSON , MD

18

18

If you're more comfortable using the individual address component widgets, you can do that, too, although it requires more mouse clicks.



19

19

Troubleshooting #6 of 9

Tips for entering addresses to get candidates and geocodes from the Commuter Connections ArcGIS Server

- If you have street, city, state, and zip *it's best not to enter the zip code*. The geocoding service works well without it. If you use it, then it *must* be correct. This is because the software assigns the highest importance to the zip code and it ignores the city and state.
1824 EDGEWOOD RD, TOWSON, MD, 21234 will not geocode (wrong zip)
1824 EDGEWOOD RD, TOWSON, MD works.
- If you're certain the zip code you're entering is correct, you can skip the city and state fields and enter the street and zip. For example,
1824 EDGEWOOD RD, 21286
works fine because 1824 EDGEWOOD RD *is* in zip 21286.

20

20

Troubleshooting #7 of 9

Suppose things don't go so well next time?

- If the USPS web site doesn't know of the address, or it warns you the address exists but they don't deliver mail to that address, the next step is to contact the commuter.
- Perhaps the commuter has moved into a recently constructed home that wasn't addressed at the time **GRU_AL** was built. When this happens, you can work with the commuter to use the manual geocoder to choose a point on the map. You should also inform commutersupport@mwcog.org so we can put the house on the map.
- Factoid: At one time, every cell tower registered with the FCC was assigned a street address. The USPS will not deliver mail to those addresses.

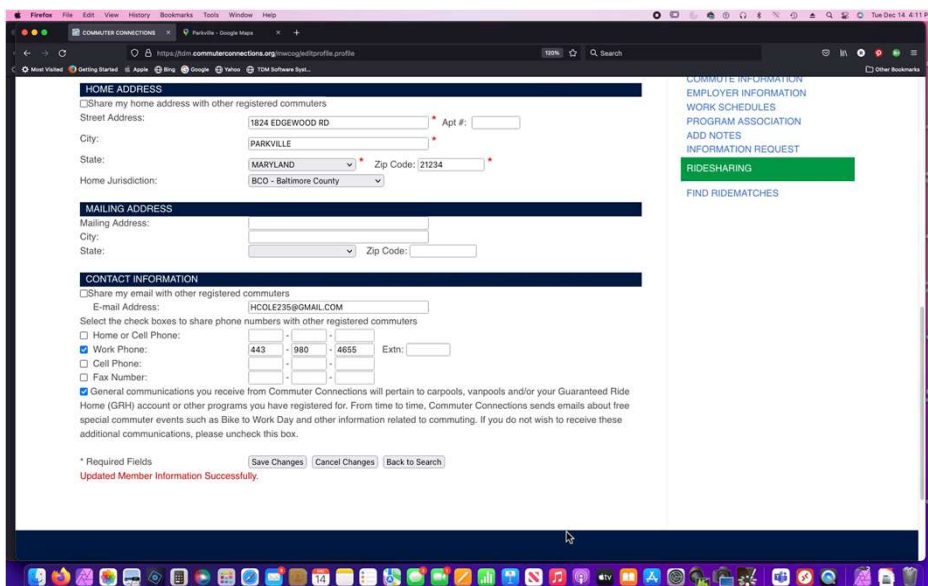
21

21

Troubleshooting #8 of 9

Make changes to the commuter's data

- Armed with valid information, you merely need to make an update in the **MEMBER INFORMATION** page of the TDM web site.



22

Troubleshooting #9 of 9

Things look much better now

- With the address changes made and saved, the commuter's account can receive ride matches ...
- ... and mail.

ADDRESS TYPE - NAME	ADDRESS	VIEW	START	DESTN.	DELETE
HOME -	1824 EDGEWOOD RD TOWSON MD 21286	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
WORK - ANNE ARUNDEL COMMUNITY COLLEGE	101 COLLEGE PKWY ARNOLD MD 21012	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
PARK&RIDE - Cromwell Bridge	I-695/Exit-29A Cromwell Bridge Rd E Baltimore MARYLAND	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PARK&RIDE - Providence Road	I-695/Exit-28 Providence Rd Baltimore MARYLAND	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PARK&RIDE - Carney	MD-147 (Harford Rd) at Jomat Ave Baltimore MARYLAND	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PARK&RIDE - Gardenville P&R	US-1 south of Hamilton Ave Baltimore City MARYLAND	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
PARK&RIDE - Light Rail-Lutherville Station	Ridgley Rd / Timonium Mall Baltimore MARYLAND	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

23

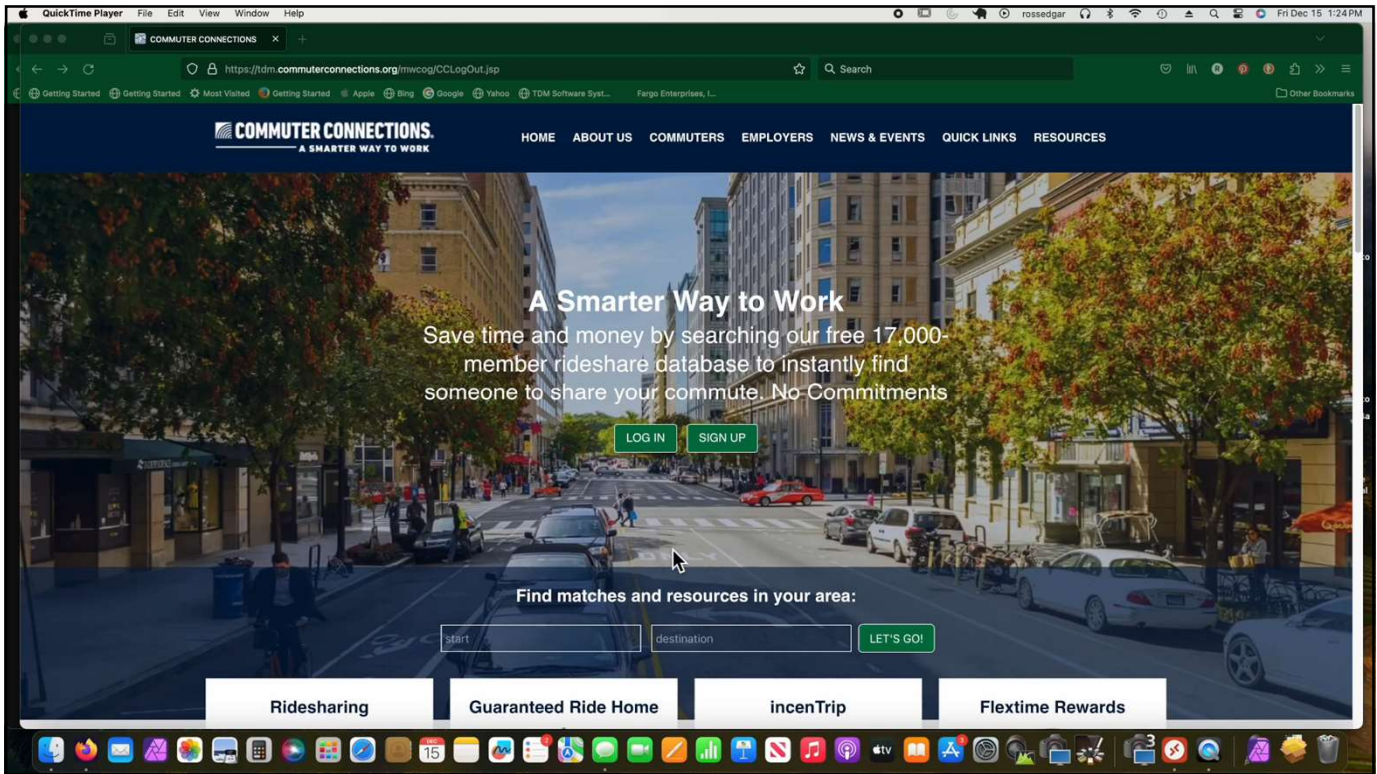
23

Working with challenging addresses:

A second example that shows how to deal with new streets.

24

24



25

Finally ...

How to get help when you get stuck on one of these

- Send an email to commutersupport@mwco.org
- Put the words "Geocoding Problem" in the subject line so it lands on the right desk quickly.
- Be sure to include the commuter ID and brief description of what you have done so far to try to resolve the problem.

26

26

Questions, suggestions, would you like to learn more?

27