

MAR Geocoder 3.1 User Guide

(Master Address Repository)



Office of the Chief Technology Officer



August 22, 2013

Table of Contents

- Geocoding Defined 3
- MAR Geocoder Overview..... 3
- MAR Geocoder Availability 4
- Installing the MAR Geocoder 4
- Launching the MAR Geocoder 4
- Interface Description 5
- Functionality Matrix..... 6
- Data Preparation for Input..... 7
 - Limitations to Consider 8
- Advanced Logic Selection..... 9
- Batch Geocode 9
- Interactive Geocode..... 9
 - Map and Photography Display 11
 - Address Submission Sequence..... 12
- Results for Geocoding 13
- Generating a GIS Layer from Geocoding Results 13
- Additional Help and Resources 14

Geocoding Defined

Geocoding is the process of assigning a location in the form of geographic coordinates (often expressed as latitude and longitude) to geographic data. This is done by comparing the descriptive geographic data to known geographic locations such as addresses, blocks, intersections or place names. For example '2500 Q ST N.W.' is geocoded to the MAR address of '2500 Q STREET NW' and assigned the XY coordinates of 395476.16, 138058.36.

MAR Geocoder Overview

The MAR Geocoder 3.1 is designed to geocode a large number of records that are in an existing Microsoft Access 2003, 2007 and 2010 table or Excel spreadsheet 2003 and 2007. It is written to accept location records containing a single address, intersection, block, or a place name. The MAR Geocoder quickly matches geographic data to known locations in the District of Columbia's Master Address Repository (MAR). The MAR database includes addresses, blocks, intersections and place names located within the District of Columbia (Washington, DC); it will not geocode to locations outside of DC.

The MAR Geocoder is a standalone executable written as an implementation of the [MAR Web Service](#) (see p14 for full link) provided by OCTO GIS. The web service has several built in "scrubbing" tools that review records presented and parses it using look up tables for alternative names, abbreviations, misspellings, etc. The extra processing involved in "scrubbing" the data reduces the efficiency at which records can be geocoded. This reduced efficiency results in a **1.5 second average time to geocode each record**. The result is more accurate geocoding to the MAR standard.

MAR Geocoder Availability

The tool is available for download from the [Master Address Repository web page](#) (see p14 for full link). The tool is called the MAR Geocoder. Clicking on the download link will prompt you to save the installation executable to your local machine; this is the preferred method for installation.

The direct link is:

http://octo.dc.gov/DC/OCTO/DC%20GIS/Real%20Property/Publication%20Files/MAR_Installer_v31.zip

Installing the MAR Geocoder

You must have administrative privileges to your machine to successfully complete the installation.

After downloading the installer, double click to launch and follow the installation wizard instructions. Once installed, you will have a new icon on your desktop and folder in your *Windows Start* menu called *MAR Geocoder*. From this shortcut you can view the documentation, launch the software or uninstall the program from your system.

Launching the MAR Geocoder

The tool is accessible through the desktop icon or the *Windows Start* menu under the *MAR Geocoder* folder. Clicking on the shortcut will launch the MAR Geocoder dialog. From the dialog box you will be able to geocode your Microsoft Access and Excel Spreadsheet datasets utilizing the Master Address Repository (MAR) web service. The MAR Geocoder requires an active internet connection.

Interface Description

Select Access or Excel Spreadsheet

Path to database or spreadsheet

Name of table to Geocode

Column containing addresses, block etc.

Optional Zip code

Standard Output Fields

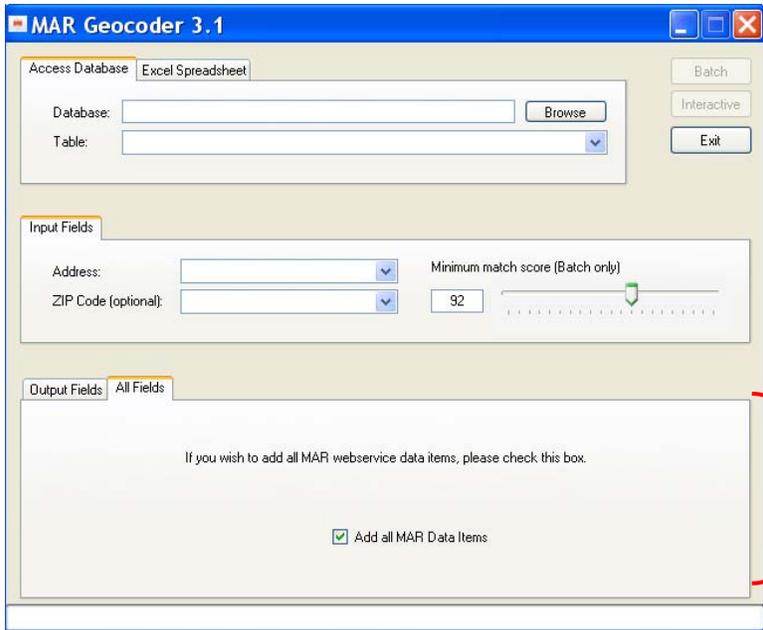
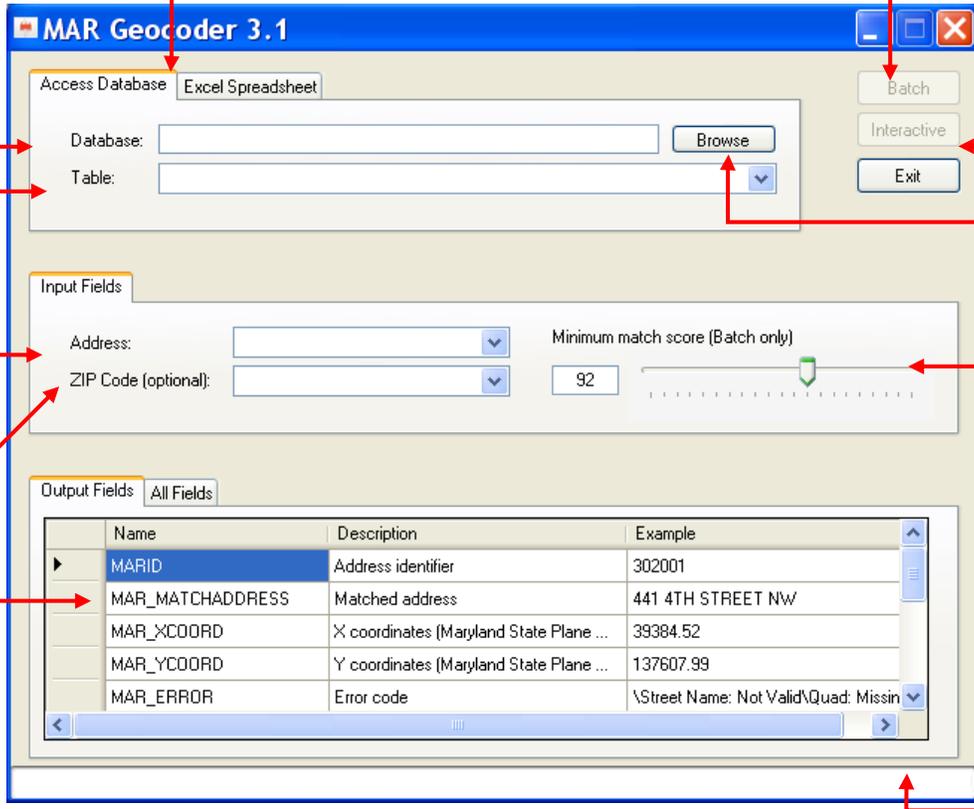
Run tool in batch match

Interactively match one-by-one

Browse for Data

Minimum accepted match score

Progress Bar



Add additional MAR fields to the Geocoding output.

Functionality Matrix

Description	Purpose
Select Access or Excel Spreadsheet	Users may choose to geocode a Microsoft Access database or Excel Spreadsheet.
Path to Database or Spreadsheet	This is the path to your input database or spreadsheet; it will need to be in Microsoft Access or Excel Spreadsheet format.
Browse for Data	Navigate to your database or spreadsheet; it will populate the, “path to database or spreadsheet.”
Name of table to Geocode	Select from the drop down menu to choose the table within the database or spreadsheet that contains your locations to geocode.
Column containing addresses	The column within your table that has the locations string (addresses, blocks, intersections, or place names).
Optional Zip code	Select the column containing zip code data (optional). The column type must be “string.”
Minimum accepted match score	The match score is the relative accuracy of the point found based on the information given. Recommended to maintain a score of 92% or greater.
Run tool in Batch Mode	To Geocode all records together; click to start the process.
Interactively match one-by-one	Once Batch Mode is complete, the interactive tool allows you to see the potential locations for each unmatched location. Choose the correct location manually.
Standard Output Fields	The following listed fields will be automatically added to your new geocoded table. See descriptions in the tool for each new field.
Progress Bar	Approximates completion for Batch Match.
Add additional MAR fields to the Geocoding output	Checking this option will append additional field from the MAR such as PSAs, Census Tracts Numbers and Roadway Segment ID Numbers.

Data Preparation for Input

The tool requires data to be in either a Microsoft Access Table or Excel Spreadsheet as the input. The addresses for geocoding will also need to be in a **single column**. Name the column accordingly and avoid using keywords, [SQL reserve words](#), (see p14 for full link) spaces or special characters. The information in the “address” or “location” column must meet the [DC Addressing Standards](#) (see p 14 for full link)

Single Address:

Address				
Number	Suffix	Name	Type	Quadrant
615		3rd	St	NE

Intersection:

Address_Intersection					
Name	Type	Quad	&	Name	Type Quad
F	St	NE	&	3rd	St NE

Block of:

Address_Block_Of				
Hundred Blk	Block of	Name	Type	Quad
400	Block of	4th	St	NE

Block between:

Address_Block_Between					
Name	Type	Quad	between	Name	Type Quad
F	St	NE	between	3rd	St NE
			&	4th St NE	

Place Name

Alias Name
Alias Name
Washington Monument

The cleaner the locations are formatted, the more efficient the geocoding results will be. The tool is designed to handle imperfections; however, there are some limitations and considerations highlighted below.

Limitations to Consider

The tool may **not** handle and skip the following format issues:

Hyphens: Globally remove all hyphens from address data.

123-124 Sample Street NW

1234 – Sample Street NW

Spaces in ordinal number: Globally replace the space before the exact phrases ‘th’, ‘rd’, ‘nd.’

123 4th Street NW

123 2nd Street NW

Quoted Street “LETTER:” Globally remove all quotation marks.

123 “H” Street NW

The tool **will** handle pound “#” signs:

#123 Sample Street NW

The MAR Geocoder requires an internet connection and this increases the time it takes to geocode an address to about 1.5 seconds per active address. By checking the option of ‘If you wish to add all MAR Web service data items, please check this box,’ the geocoding will be significantly slower. Large files over 150,000 records often fail to geocode due to their large size.

Although the tool was tested among common datasets successfully, other limitations and errors may exist. If users identify consistent errors and limitations, please contact dcgis@dc.gov so they can be resolved and documented.

Advanced Logic Selection

Some addresses or intersections may have two 100% exact matches and multiple correct answers that are lower than 100%. This is due to the nature of the street fabric in DC; an example would be the intersection of Florida Avenue NW and U Street NW. The two streets legally intersect twice in the same quadrant, which one is correct? Without knowing anything additional, it is impossible to know the correct location.

Batch Geocode

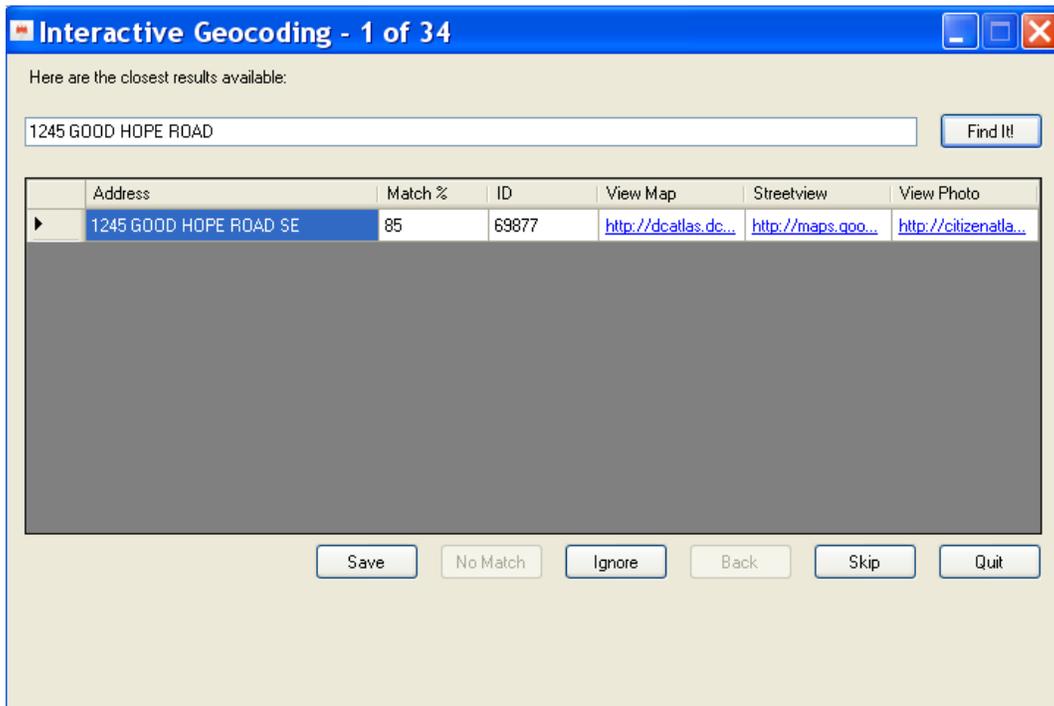
In almost every case, the Batch Geocode mode should be performed before starting the Interactive Geocode. After setting up all the input criteria (see page 5), click Batch. The Batch Geocode will begin; records will automatically be geocoded to the MAR. Use the progress bar to monitor progress. Once complete, a summary of the results of the Batch Geocode will be shown.



Interactive Geocode

Once the batch process is complete, examine the results by sorting the records (in Access or Excel) that have not been geocoded (those that do not have an X or Y coordinate). Use the Interactive Geocode function (in the MAR Geocoder) to continue geocoding records individually.

The Interactive Geocode button will allow users to manually look at each unmatched record with all potential locations. Users can then determine the correct match. Depending on the total unmatched records to go through, the process can take some time. If the number of unmatched records seems high, go back into the table or spreadsheet and investigate.



The interactive match function will list the original location descriptor from the table along with potential matches below. Users have several options:

- Select the correct match location and click **Save**.

This will match the selected location to the record and move users to the next.

- Edit the original record and click **Find It!** to get a new list of candidates. In this example, you could add the quadrant to find a better match.

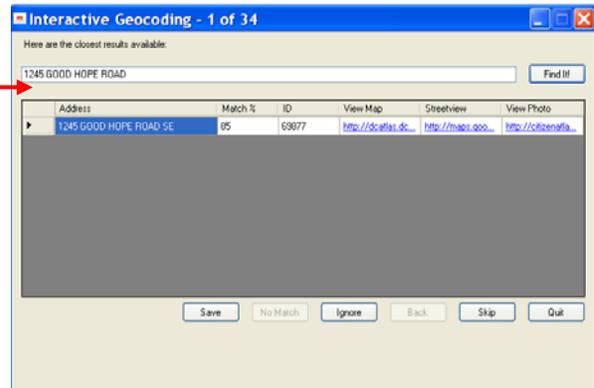
- Review the record later by clicking **Skip**.

- Click to **Ignore** the record.

- Move to the previous record by clicking **Back**.

- Click **Quit** to save, exit and continue at a later time.

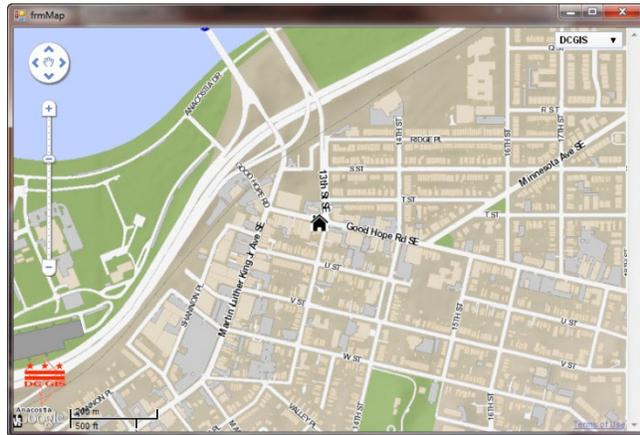
- If the original location is valid, however not verified, select **No Match** to launch the MAR Address Submission Sequence form (see page 12).



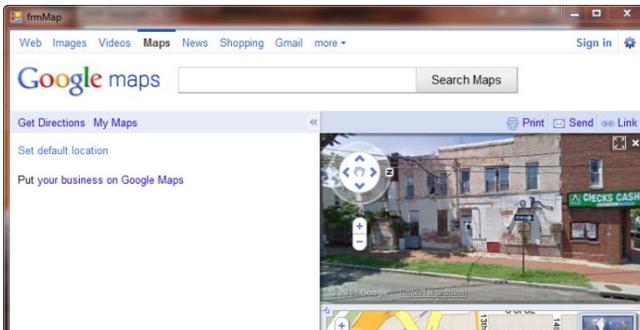
Map and Photography Display

The interactive match tool provides users with the ability to view maps and photography for any potential match candidates. Scroll over to the right of the table and click to view a map from DC Guide, photography from Google Street View or a frontage photo from the Office of Tax and Revenue. Street frontage photos are only available for property addresses and not available for blocks, intersections or place names. The displays use map web services and have basic navigation tools available including zoom in, zoom out and pan.

View Map



Street View



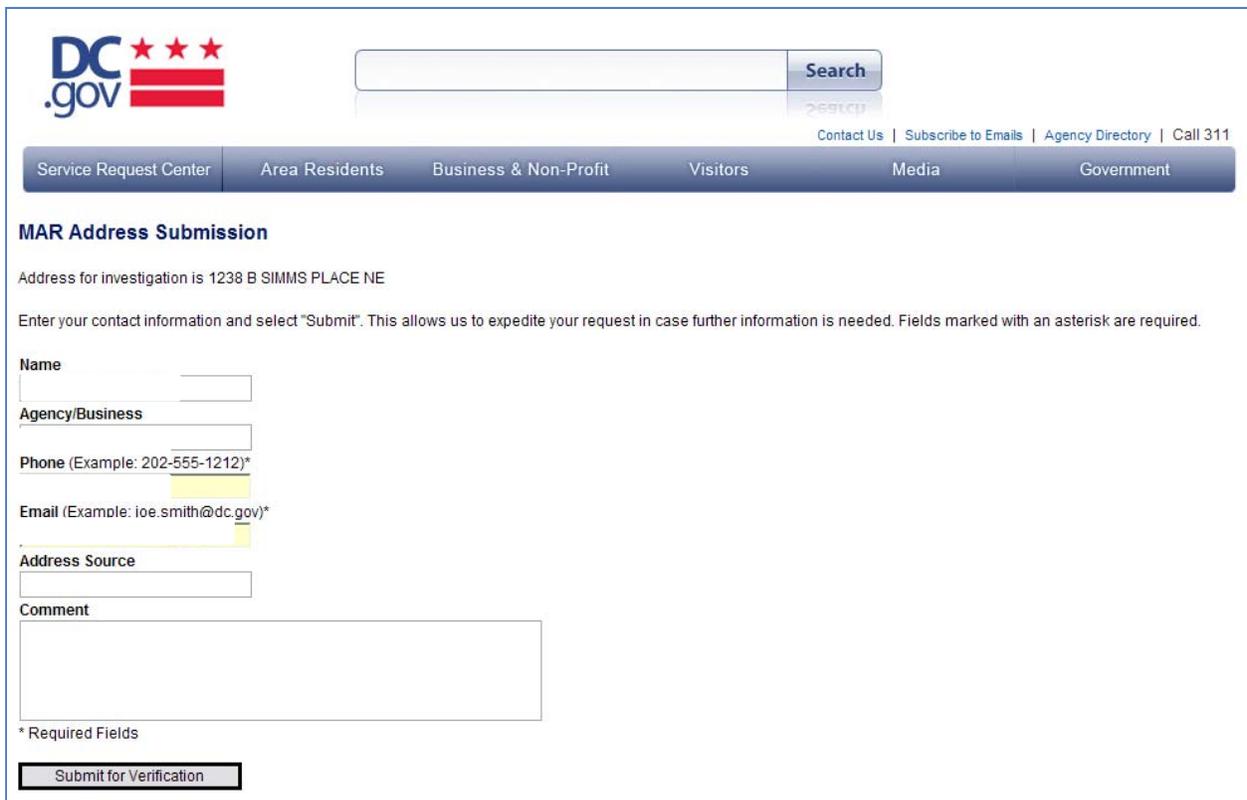
View Photo



Address Submission Sequence

There will be occasions where users find address records not currently in the Master Address Repository, and possibly should be. The *No Match* button will enable itself if an address record in the table is valid with no available match, but not yet verified. A valid address follows the District's Address Standard containing all the elements of an address and is within a valid address range (see p7; Data Preparation for Input). Select *No Match* to launch the *Address Submission Sequence* form. This is an automated way for a potential new address to be added to the MAR. Users will be presented with a dialog box to enter contact information. The District's Master Address Repository team will review the submission and typically reply within two business days.

The potential address is pre-populated on the form. Depending on a user's internet browser settings, contact information will also be pre-populated. Otherwise users must enter the required information.



The screenshot shows the DC.gov website interface. At the top left is the DC.gov logo with three stars. To its right is a search bar with a "Search" button. Below the search bar is a navigation menu with links: "Service Request Center", "Area Residents", "Business & Non-Profit", "Visitors", "Media", and "Government". On the right side of the navigation menu are links for "Contact Us", "Subscribe to Emails", "Agency Directory", and "Call 311".

The main content area is titled "MAR Address Submission". Below the title, it states: "Address for investigation is 1238 B SIMMS PLACE NE". A note reads: "Enter your contact information and select 'Submit'. This allows us to expedite your request in case further information is needed. Fields marked with an asterisk are required." The form contains the following fields:

- Name:
- Agency/Business:
- Phone (Example: 202-555-1212)*:
- Email (Example: joe.smith@dc.gov)*:
- Address Source:
- Comment:

At the bottom left of the form, there is a note: "* Required Fields". At the bottom center, there is a "Submit for Verification" button.

Results for Geocoding

The tool will append new fields within the geocoded table or spreadsheet. These fields are automatically populated by the tool based on the set minimum match score. Set the minimum match score to 100 to ensure that the addresses matched are the absolute correct ones. However, by default the tool is set to a minimum match score of 92 in order to return the maximum correct matches. Below are the standard output fields that will be appended to the table or spreadsheet:

MAR_MATCHADDRESS	MAR_XCOORD	MAR_YCOORD	MAR_LATITUDE	MAR_LONGITUDE
Matched Address in MAR	X Coordinates (Maryland State Plane Meters)	Y Coordinates (Maryland State Plane Meters)	Latitude (decimal degrees)	Longitude (decimal degrees)
MAR_WARD	MAR_CENSUS_TRACT	MAR_ZIPCODE	MARID	MAR_ERROR
Ward	Census Tract	Zipcode	MAR Unique Identifier	The error message.
MAR_SCORE	MAR_SOURCEOPERATION	MAR_IGNORE		
Confidence Level of Geocoding	Table Source for Joining	Determines if record should be interactively geocoded again		

Generating a GIS Layer from Geocoding Results

After successfully geocoding the address table or spreadsheet using the MAR Batch Geocoder, users will likely need to add this data to a map. The following instructions are based off of the [ESRI ArcGIS Desktop product suite using the ArcView](#) (see p 14 for full link) application.

1. Launch ArcMap
Note: For District of Columbia employees and contractors, OCTO hosts and maintains this application in a Citrix environment. To learn more about accessing these services please visit [DC Intranet – CitrixGIS Desktop Applications](#) (see p14 for full link) .
2. Add the table or spreadsheet to ArcMap. It will appear in the Layers table of contents.
3. **Right-click** on the file and choose to Display X,Y Data.

The next window will have the X and Y fields automatically chosen. If not, select the correct fields for X and Y (MAR_XCOORD and MAR_YCOORD).

4. For *Spatial Reference of Input Coordinates* click the **Edit** button.
5. Click Select and navigate to → *Projected Coordinate Systems* → *State Plane* → *NAD 1983* and choose the following: **NAD 1983 State Plane Maryland FIPS 1900**

This will produce a new layer event in the ArcMap table of contents. Each visible point represents a geocoded (mapped) address. This event layer can be saved or exported to a new map layer by right-clicking and accessing the export options.

Additional Help and Resources

- ArcGIS Resource Center: <http://resources.esri.com/arcgisdesktop>
- ESRI Virtual Campus: <http://training.esri.com>
- ESRI Online Support Center: <http://support.esri.com>
- MAR Sample Client: <http://dcatlas.dcgis.dc.gov/mar/>
- DC Atlas Plus: <http://atlasplus.dcgis.dc.gov>
- MAR Website:
<http://octo.dc.gov/DC/OCTO/Maps+and+Apps/Online+Mapping/All+Online+Maps/Master+Address+Repository>
- MAR Data Dictionary:
<http://octo.dc.gov/DC/OCTO/Agency+Support/IT+Standards/Geography+Standards/MAR+Data+Dictionary>
- DC Address Standards
<http://octo.dc.gov/DC/OCTO/Agency+Support/IT+Standards/Geography+Standards/MAR+Address+Standards>
- DC Government Full Time Employee Training:
<http://octo.in.dc.gov/node/164692>
- Learn How to Use GIS Web Services
<http://octo.dc.gov/DC/OCTO/Agency+Support/Development,+Training+and+Support/How+To+Guides+for+GIS+Web+Services>
- SQL Reserved Sites
[http://msdn.microsoft.com/en-us/library/aa238507\(SQL.80\).aspx](http://msdn.microsoft.com/en-us/library/aa238507(SQL.80).aspx)
- DC Intranet Citrix Desktop Applications
<http://octo.in.dc.gov/node/164688>

Support for the MAR Geocoder 3.1 tool is available to District agencies and their contractors. Limited support is available for the general public. Please email dcgis@dc.gov for further information.

Updated by David Jackson - October 2012, August 2013