

Address Reference Guide

This reference accompanies the *Address Parsing Tool*. It describes the address components participating in full street addresses, the elements the *Address Parsing Tool* outputs, how to check the parsing results, and troubleshoot address parsing errors and flags.

Use this guide if you are utilizing the *Address Parsing Tool* to parse street addresses and want to learn about how to optimize the results of your address parse.

The *Address Parsing Tool* was developed as a part of the Wisconsin Statewide Parcel Initiative as a pathway for data contributors to process unparsed addresses into address sub-components meeting the statewide layer's attribute schema. The schema's address elements are based on the FGDC-endorsed *United States Thoroughfare, Landmark, and Postal Address Data Standard*¹ and include common elements with definitions. This tool is designed to interpret local addresses and output address components meeting the FGDC and Parcel Initiative standards to the highest degree possible, although variations in data at the local level may impede accuracy.

Why is This Guide Necessary?

In order to achieve the most optimal parsing results from this tool, it is important for the user to understand the nature of the address elements within the Parcel Initiative's schema as well as the various outputs provided as a result of this tool.

This is a reference guide specifically for those using the *Address Parsing Tool*, developed to meet the needs of the V2 schema in Appendix F of the V2 Submission Documentation.

Address Element Examples

1.1 Address standards and formatting

The attribute schema for the Wisconsin Parcel Initiative can be found in Appendix D, Schema Requirements for V2. The address elements in the table below are the targeted outputs of the Address Parsing Tool.

Address Element Outputs	
Schema Name (and Alias)	Corresponding Tool Output
ADDNUMPREFIX (Address Number Prefix)	NEW_ADDNUMPREFIX
ADDNUM (Address Number)	NEW_ADDNUM
ADDNUMSUFFIX (Address Number Suffix)	NEW_ADDNUMSUFF
PREFIX (Prefix)	NEW_PREFIX
STREETNAME (Street Name)	NEW_STREETNAME
STREETTYPE (Street Type)	NEW_STREETTYPE
SUFFIX (Suffix)	NEW_SUFFIX
UNITTYPE (Unit Type)	NEW_UNITTYPE
UNITID (Unit ID)	NEW_UNITID

Note that this tool aims to parse street addresses only and does NOT parse the following to output fields:

- Landmark Names
- Place Names (City/Town/Village Name)
- Zip-Codes

¹ https://www.fgdc.gov/standards/projects/FGDC-standards-projects/street-address/index_html

If Landmark Names, Place Names, or Zip-Codes exist within the address, they will be written to the extraneous data field while street address will be parsed to their respective elements.

Some address examples, and their expected *Address Parsing Tool* outputs are depicted in Figure 6. The address and ordering of sub-components to outputs depicted meet the specifications of the Parcel Initiative’s Schema.

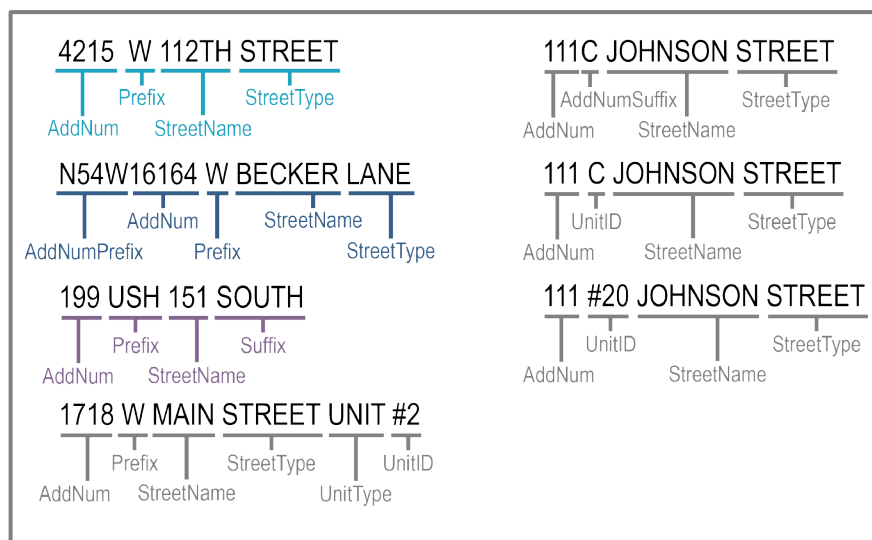


Figure 1. Sub-components of address outputs

Correcting Flags and Errors

2.1 Special characters

The address parsing tool is very sensitive to the inclusion or location of non-alphanumeric characters in the site address field. For this reason, character cleanup is vital to achieving optimal results. In many cases, you will need to manually remove characters flagged by the tool.

Characters flagged by the tool should be removed, with some exceptions where a character is acceptable.

Special Characters		
	Acceptable	Unacceptable – Requires Change
Hyphen (-) The hyphen character is most often used in a site address for indicating a range or compound words.	Hyphens are acceptable in compound street names, concurrently running streets, or unit ID ranges. <ul style="list-style-type: none"> 100 W MAIN ST UNIT A-E 100 TRI-COUNTY RD 100 BUS 18-151 	Hyphens are not allowed to indicate an address number range. A range of address numbers indicates multiple site addresses. There should only be one site address per parcel. To solve this problem, either remove the address number range or duplicate the parcel assigning each parcel a different number in the range. <ul style="list-style-type: none"> 100-104 W MAIN ST should be 100 W MAIN ST
Forward Slash (/) The forward slash is used frequently in a site address to denote different street names.	Forward slashes are acceptable when denoting concurrently running streets. <ul style="list-style-type: none"> 100 STH 25/58 	Forward slashes are unacceptable when indicating a street name alias. To solve this problem, remove one of the street names and the forward slash. <ul style="list-style-type: none"> 100 STH 26/MAIN ST should be 100 STH 26 or 100 MAIN ST
Pound sign (#) The pound sign is used to indicate a unit. For example, instead of “Unit 1” another way to write it would be “#1.”	Pound symbols are acceptable when used to indicate a unit. <ul style="list-style-type: none"> 100 W MAIN ST #1 100 #1 W MAIN ST 	All other situations. To solve this problem, remove the character and other associated text. Then add number if known. <ul style="list-style-type: none"> NO# W MAIN ST should be 100 W MAIN ST or W MAIN ST UNKNOWN # W MAIN ST should be 100 W MAIN ST or W MAIN ST

2.2 Extraneous Data

The **Extraneous_Data_Flag** field is generated when you run the parser over the site address field. It contains various data that the parser was unable to assign to a common address element. **In most cases this extraneous information should be removed from the full address prior to re-running the parse tool.**

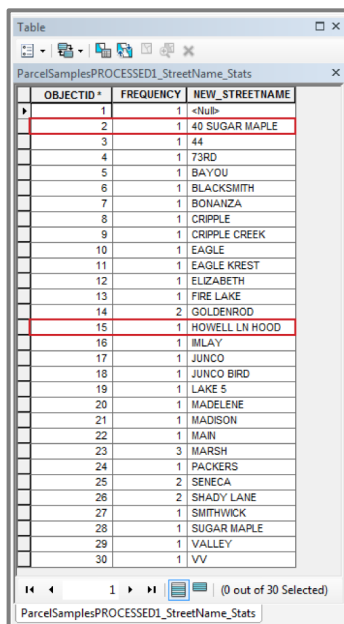
A few examples of data that could be found in this field include the latter portion of an address range or some type of property descriptor, or an alternate street name attached to the full address.

Please note that while in most cases this information is not an acceptable address element, it is possible that the parser incorrectly placed an acceptable element into this field. It is recommended you take time after your final parse to manually move elements that may have inadvertently ended up in the **Extraneous_Data_Flag** field to their appropriate field.

Leverage the Tool's Summary Tables

3.1 Address Fields Summary output tables

Once you have successfully run through as many iterations as necessary to remove **Parse_Error_Flags**, corrected issues associated with **Incomplete_Data_Flags** and **Character_Flags**, it is recommended that you select the tool check box that will create **Address Fields Summary** for your Prefix, StreetName, StreetType, and Suffix fields. These summary tables will provide an additional method for double checking the parse success and identify any potential errors that may not have been identified by the various flags listed above. Below are some examples of the summary tables from the (SampleData.gdb » ParcelSamples) feature class provided with the tool.



OBJECTID	FREQUENCY	NEW_STREETNAME
1	1	<Null>
2	1	40 SUGAR MAPLE
3	1	44
4	1	73RD
5	1	BAYOU
6	1	BLACKSMITH
7	1	BONANZA
8	1	CRIPPLE
9	1	CRIPPLE CREEK
10	1	EAGLE
11	1	EAGLE KREST
12	1	ELIZABETH
13	1	FIRE LAKE
14	2	GOLDENROD
15	1	HOWELL LN HOOD
16	1	IMLAY
17	1	JUNCO
18	1	JUNCO BIRD
19	1	LAKE S
20	1	MADELENE
21	1	MADISON
22	1	MAIN
23	3	MARSH
24	1	PACKERS
25	2	SENECA
26	2	SHADY LANE
27	1	SMITHWICK
28	1	SUGAR MAPLE
29	1	VALLEY
30	1	VV

Figure 2. Highlighting of incorrect street names

Figure 7 shows a summary of the **StreetNames** from the **ParcelSamples** feature class. The two highlighted records "40 SUGAR MAPLE" and "HOWELL LN HOOD" are incorrect street names. Once you have read through the summary table and identified incorrect names, you can make corrections in various ways:

Option 1

Sort the attributed field alphabetically, open an editing session, navigate to the problematic street name and make the necessary correction. Be sure to make the correction to the full address field that you are running the parse over too, especially if you plan to re-run the parse.

Option 2

Another option is to use the **Select by Attributes** option within the attribute table, open an editing session and construct a definition query that will select the incorrect records and allow you to make the necessary corrections. Be sure to make the correction to the full address field that you are running the parse over too, especially if you plan to re-run the parse.

OBJECTID	FREQUENCY	NEW_STREETTYPE
1	5	<Null>
2	3	AVE
3	1	AVENE
4	1	BIRD
5	3	BLVD
6	1	CIR
7	1	CREEK
8	2	DR
9	1	KREST
10	2	LN
11	2	PL
12	1	PS
13	1	R
14	4	RD
15	6	ST
16	1	WAY

Figure 3. Highlighting of incorrect street types

Figure 8 shows a summary of the **StreetTypes** from the **ParcelSamples** feature class. The highlighted records uncommon/incorrect street types. Once you have identifier incorrect street types, you can make corrections in various ways.

Note that often when you see uncommon street types this is a result of a multi-word street name that is missing its street type. In many cases, the second word of the street name will be placed in the street type attribute field.

Option1

Sort the attributed field alphabetically, open an editing session, navigate to the problematic street name and make the necessary correction. Be sure to make the correction to the full address field that you are running the parse over too, especially if you plan to re-run the parse.

Option 2

Another option is to use the **Select by Attributes** option within the attribute table, open an editing session and construct a definition query that will select the incorrect records and allow you to make the necessary corrections. Be sure to make the correction to the full address field that you are running the parse over too, especially if you plan to re-run the parse.