

File Parser Installation Guide for WhereScape 3D

This is a guide for installing Source Enablement Packs for WhereScape 3D 8.6.5.1 or higher

Prerequisites

- Python 3.8
 - Download python installer from <https://www.python.org/downloads/>
 - Select "Add Python 3.8 to PATH" from installation Window
- PIP Manager
 - From Command Prompt (Run As Administrator) run below command

PIP Manager Install

```
python -m pip install --upgrade pip
```

- Python Packages
 - From Command Prompt (Run As Administrator) run below command -
Install Python Package

```
pip install pandas fastavro openpyxl xlswriter xlrd pyarrow fastparquet pyorc avro avro_python3
jsonpath_ng openpyxl Pillow pyarrow xmldict lxml
pip install --upgrade pandas

Amazon S3
pip install boto3

Azure DataLake Storage Gen2
python -m pip install azure-storage==0.36.0
python -m pip install azure-storage-file-datalake

Google Cloud
python -m pip install --upgrade gcloud
python -m pip install google_api_python_client google_auth_oauthlib protobuf google-cloud-core
google-cloud-datastore google-cloud-storage
```

NOTE: Above mentioned python packages can be installed by running install_WsIPython_Modules.bat(refer to section *Enablement Pack Setup Scripts*.)

Source Enablement Pack Installation

Installing a source enablement pack into your WhereScape 3D repository requires adding a set of .uiconfig files and scripts.

Importing UI Configurations

1. Go to Tools -> Settings -> Enabled functionality and make sure 'Advanced mode' has been selected.
2. Open the UI configuration manager by going to Tools -> Manage UI Configuration.
3. Click the blue import button and navigate to the 'UI Configurations' folder within the source enablement pack download file. Import all configuration files related to the source systems relevant to your project.

Importing browse scripts (in WhereScape 3D 8.6.5.1 - 8.6.6.0)

1. Open the template manager by going to Tools -> Manage templates.
2. Create a new template of type browse script called 'Browse_File_Parser' and insert the content of the 'Browse_File_Parser.py' file contained within the enablement pack download folder.
3. Select 'Python' in the Script Language drop down menu for the new browse script.
4. Click on the gear symbol next to Script Language drop down and enter the following:
 - a. Name: python
 - b. Description: python
 - c. Extension: py
 - d. Command: Python "\$SCRIPT_NAME\$"
5. Copy and paste all other files from the 'Scripts' folder to the following location: C:/ProgramData/WhereScape/Modules/WsIPython

Importing browse scripts (in WhereScape 3D 8.6.7.0)

1. Open the template manager by going to Tools -> Manage templates.
2. Create new templates of type browse script for each of the scripts within the 'Scripts' folder of the enablement pack download folder.
3. Select 'Python' in the Script Language drop down menu for each of the scripts that have been added.
4. Click on the gear symbol next to Script Language drop down and enter the following:
 - a. Name: python

- b. Description: python
- c. Extension: py
- d. Command: Python "\$SCRIPT_NAMES"

File Parser Connection Setup

Post install checks:

1. File Parser Browse Script - In 3D check the Menu: 'ToolsManage templates for the appropriate browse scripts.
2. UI Configurations - In 3D check the Menu: 'ToolsManage UI Configurations for the appropriate UI Configurations.

**Note: UI Configurations generally come in sets of 2 or 3 for a particular source type, a minimum set will have both a Connection UI Config and a Load UI Config, optionally a Column UI config may also be included.*

Amazon S3 Connection Setup

The screenshot shows a dialog box titled "Add a new connection" with a close button (X) in the top right corner. Below the title bar, there is a prompt: "Provide a name for the new connection and select its type".

The form contains the following fields and controls:

- Connection name:** Text input field containing "Amazon S3".
- Connection type:** Dropdown menu showing "Amazon S3".
- Connection Browse Script:** Dropdown menu showing "Browse_File_Parser".
- Load Table UI Configuration:** Dropdown menu showing "Load from Windows Data Parser".
- Secret Key:** Text input field with masked characters (asterisks) and a help icon (?).
- Access Key:** Text input field containing "AKIA533YNQGESIVPXJUD" and a help icon (?).
- S3 Bucket Name:** Text input field containing "iderabucket" and a help icon (?).
- S3 Region:** Text input field containing "us-east-1" and a help icon (?).

At the bottom right of the form area, there are two buttons: "Test Connection" and "Save Connection".

At the bottom of the dialog box, there are two buttons: "Next >" and "Cancel".

Azure Data Lake Storage Gen2 Connection Setup

Add a new connection X

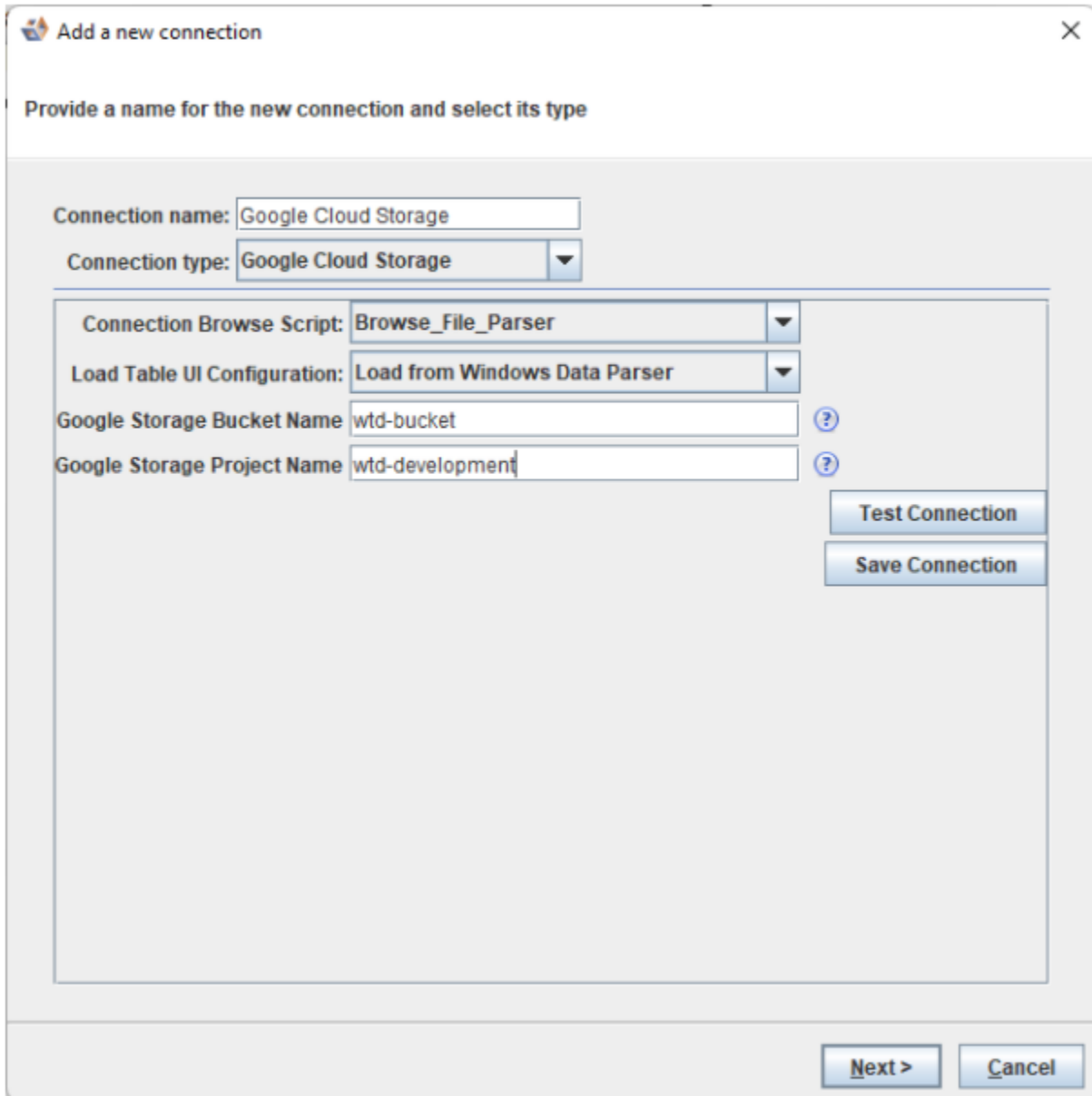
Provide a name for the new connection and select its type

Connection name:

Connection type:

Connection Browse Script:	<input type="text" value="Browse_file_Parser"/>	?
Load Table UI Configuration:	<input type="text" value="Load from Windows Data Parser"/>	?
Azure Data Lake Storage Gen2 Account	<input type="text" value="emeadatalake"/>	?
Azure Data Lake Storage Gen2 Account Access Key(Account Key)	<input type="text" value="*****"/>	?
Azure Data Lake Storage Gen2 Account SAS Token	<input type="text" value="*****"/>	?
Azure Data Lake Storage Gen2 File System	<input type="text" value="emeademo"/>	?

Google Cloud Connection Setup

A dialog box titled "Add a new connection" with a close button (X) in the top right corner. The main instruction is "Provide a name for the new connection and select its type".

Connection name:

Connection type: ▼

Connection Browse Script: ▼

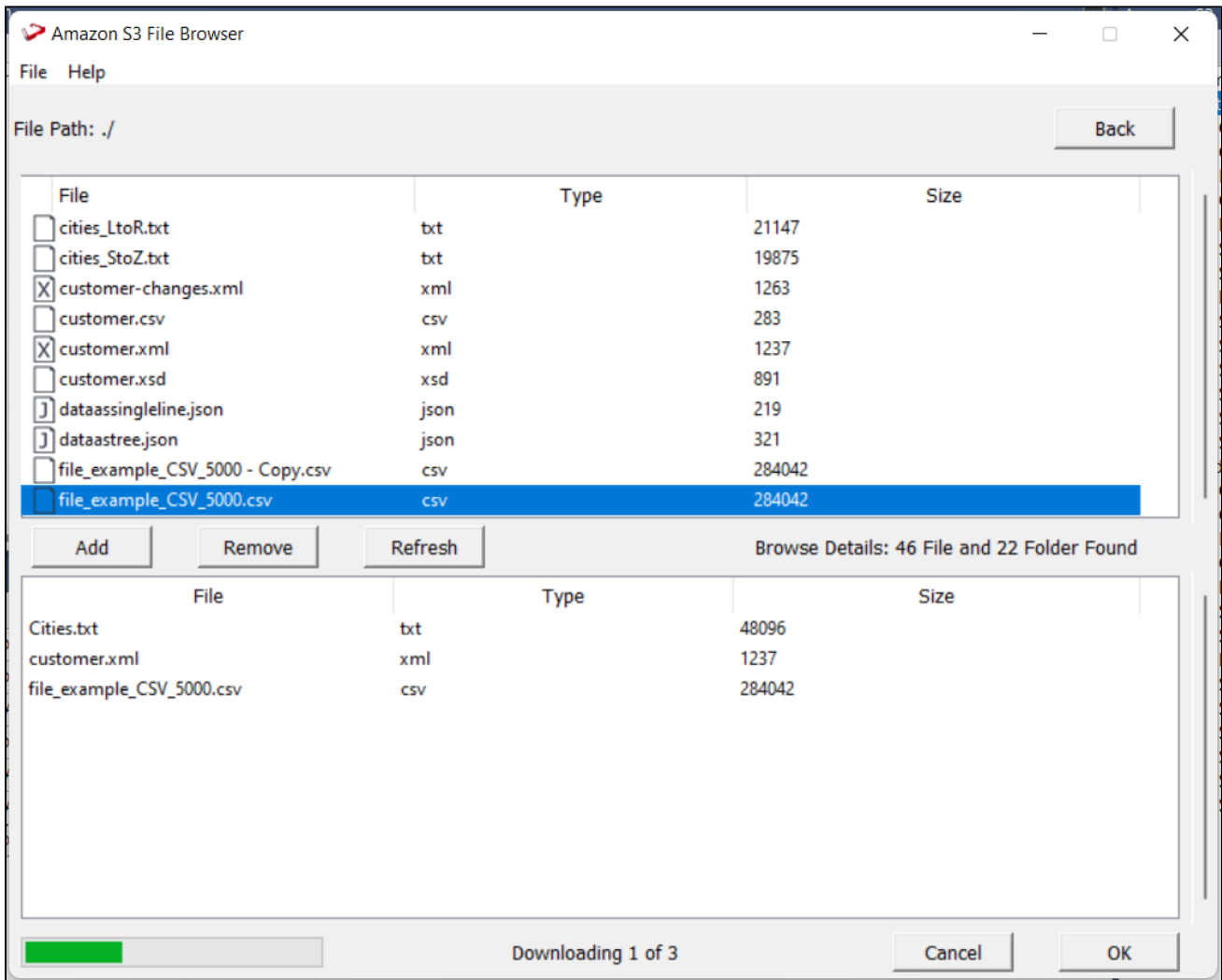
Load Table UI Configuration: ▼

Google Storage Bucket Name ?

Google Storage Project Name ?

NOTE: For Google cloud, install and configure Google Cloud SDK

Cloud Browser



1. Select and click Add to copy files to staging area.
2. Click Back to navigate to previous directory
3. Click OK to download files for parsing.

Windows Parser Connection Setup

Add a new connection ×

Provide a name for the new connection and select its type

Connection name:

Connection type:

Connection Browse Script:

Load Table UI Configuration:

Browse Parser

Choose parser as per file type

Options to Parse File: Cities.txt

Field Delimiter		Parser	DELIMITED
Record Delimiter	\n	Skip Lines	0
Field Enclosure Delimiter	"	Header Record	<input type="checkbox"/>
Row Limit For Profiling	100	Header Lines	0
SheetName	None	Coefficient	

Save Options For All Files

OK

```

city_code|city_name|state_code|country_code
278|Assen||1
279|Auckland|28|107
280|Christchurch|27|107
281|Den Bosch||1
282|Den Haag||1
283|Dunedin|31|107
284|Glasgow||109
285|Groningen|32|1
286|Haarlem||1
287|Hamilton|2|107
288|Leeuwarden|22|1

```

If the files are of same type and parsing options are same, check highlighted box to save same options.

Options to Parse File: Cities.txt

Field Delimiter		Parser	DELIMITED
Record Delimiter	\n	Skip Lines	0
Field Enclosure Delimiter	"	Header Record	<input checked="" type="checkbox"/>
Row Limit For Profiling	100	Header Lines	0
SheetName	None	Coefficient	

Save Options For All Files

OK

```

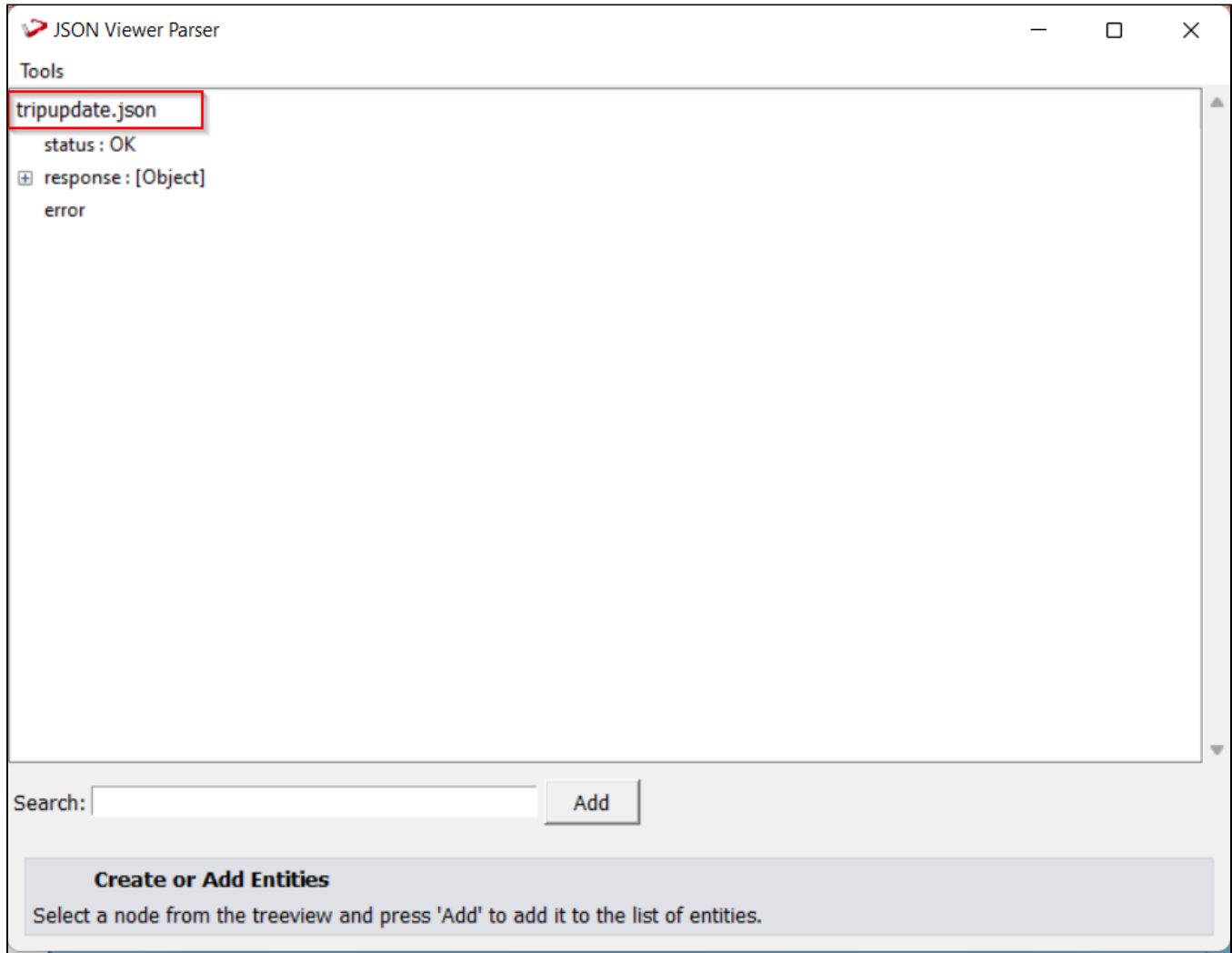
city_code|city_name|state_code|country_code
278|Assen||1
279|Auckland|28|107
280|Christchurch|27|107
281|Den Bosch||1
282|Den Haag||1
283|Dunedin|31|107
284|Glasgow||109
285|Groningen|32|1
286|Haarlem||1
287|Hamilton|2|107
288|Leeuwarden|22|1

```

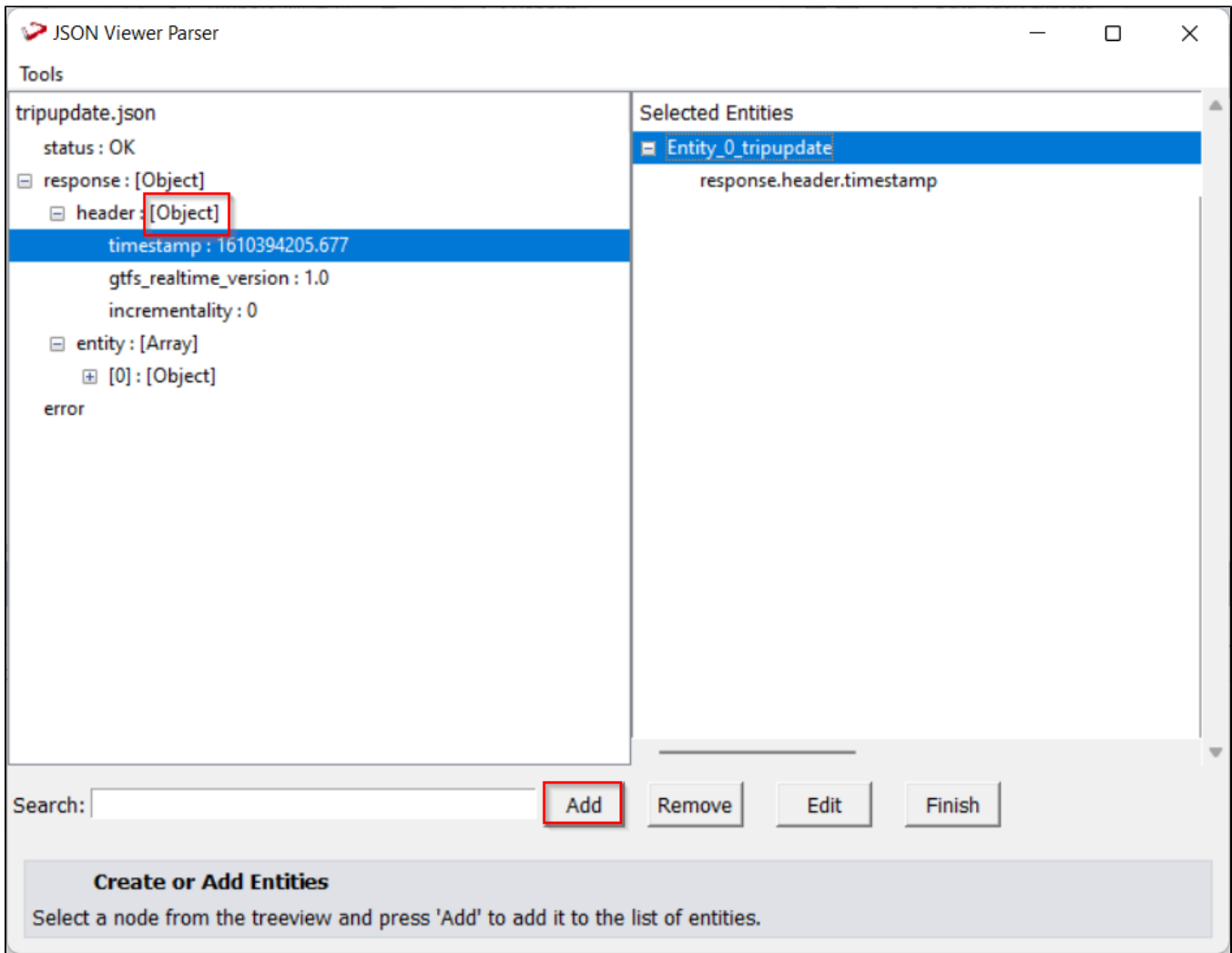
Parser for JSON and XML Files

The JSON parser GUI's main pane, The file name is highlighted, and the JSON tree structure is shown below it.

Hovering the cursor over any widget or element in the GUI will display information about that widget or element in the bottom help box.

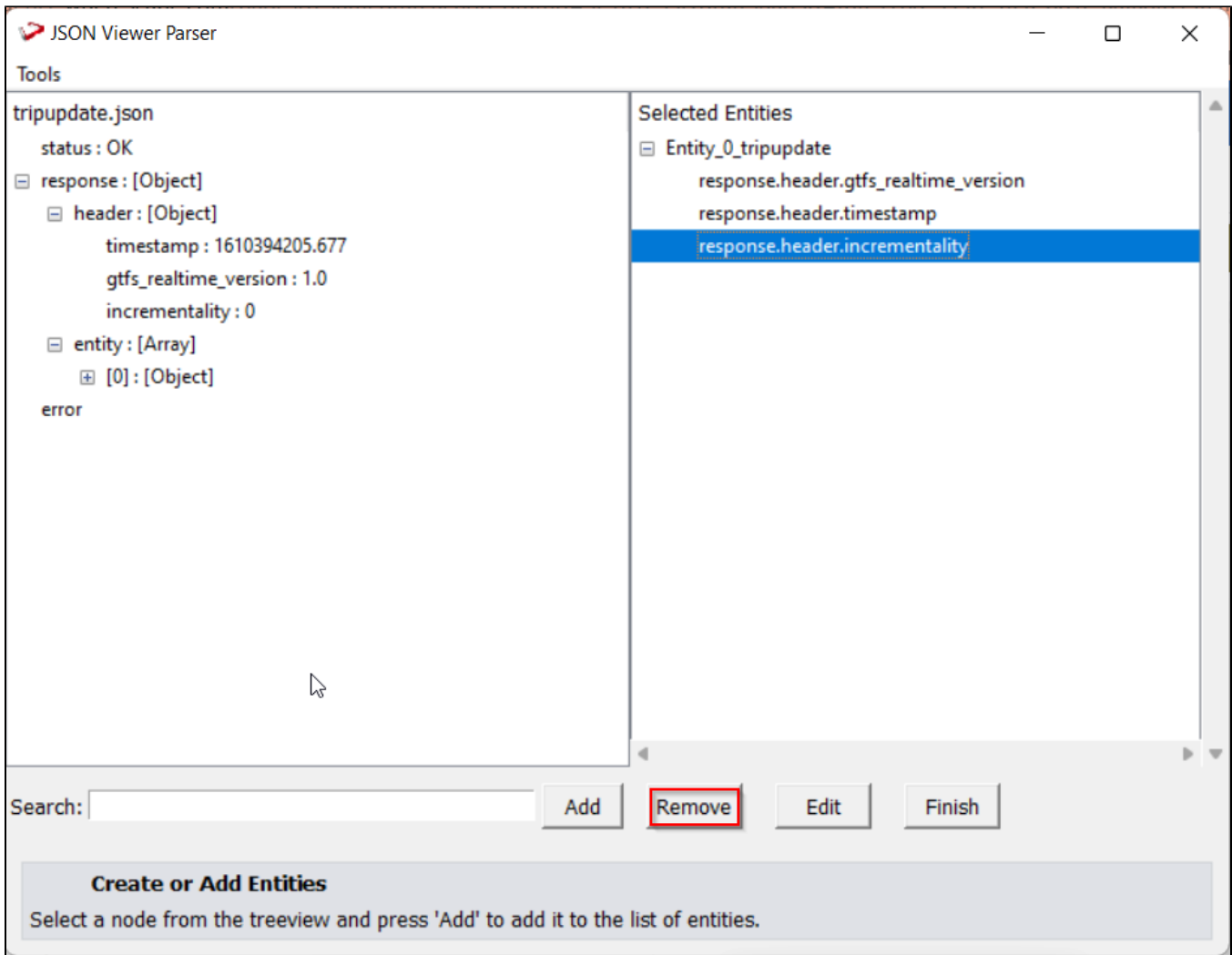


Select any node in the JSON tree and press the "Add" button at the bottom to create a new entity. On the right side of the window, a new pane will appear. The name of the new entity will be "Entity 0" by default. If the selected node is a leaf node (key value pair), this new entity will include only its key; if the selected node is an object or array, this new entity will include all of its children. The data type of the node is highlighted in the figure below.

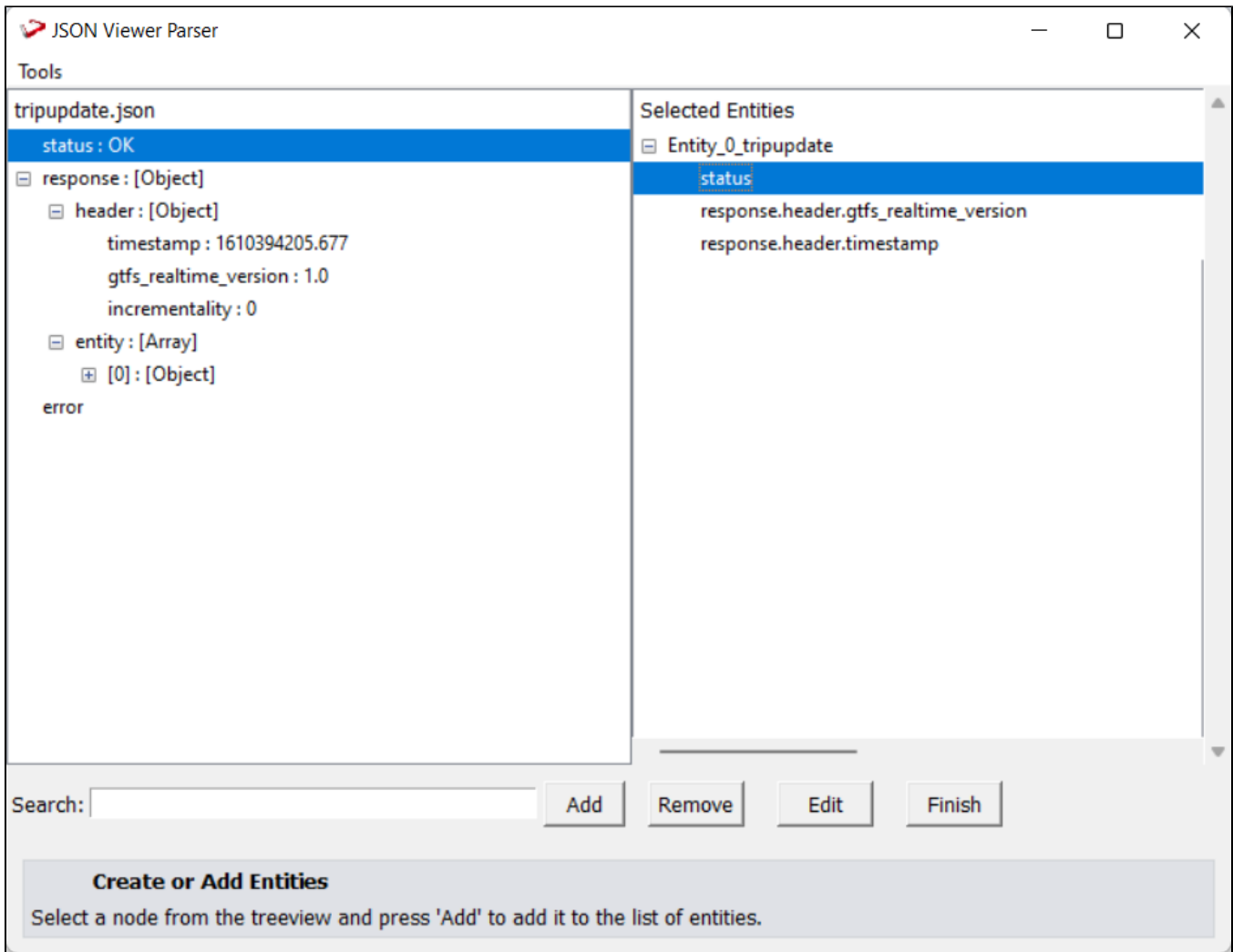


Select the entity and use the "Remove" button to remove any specific node. To remove the entire entity object, choose the primary node (for example, Entity_0) and press the "Remove" button in the same way.

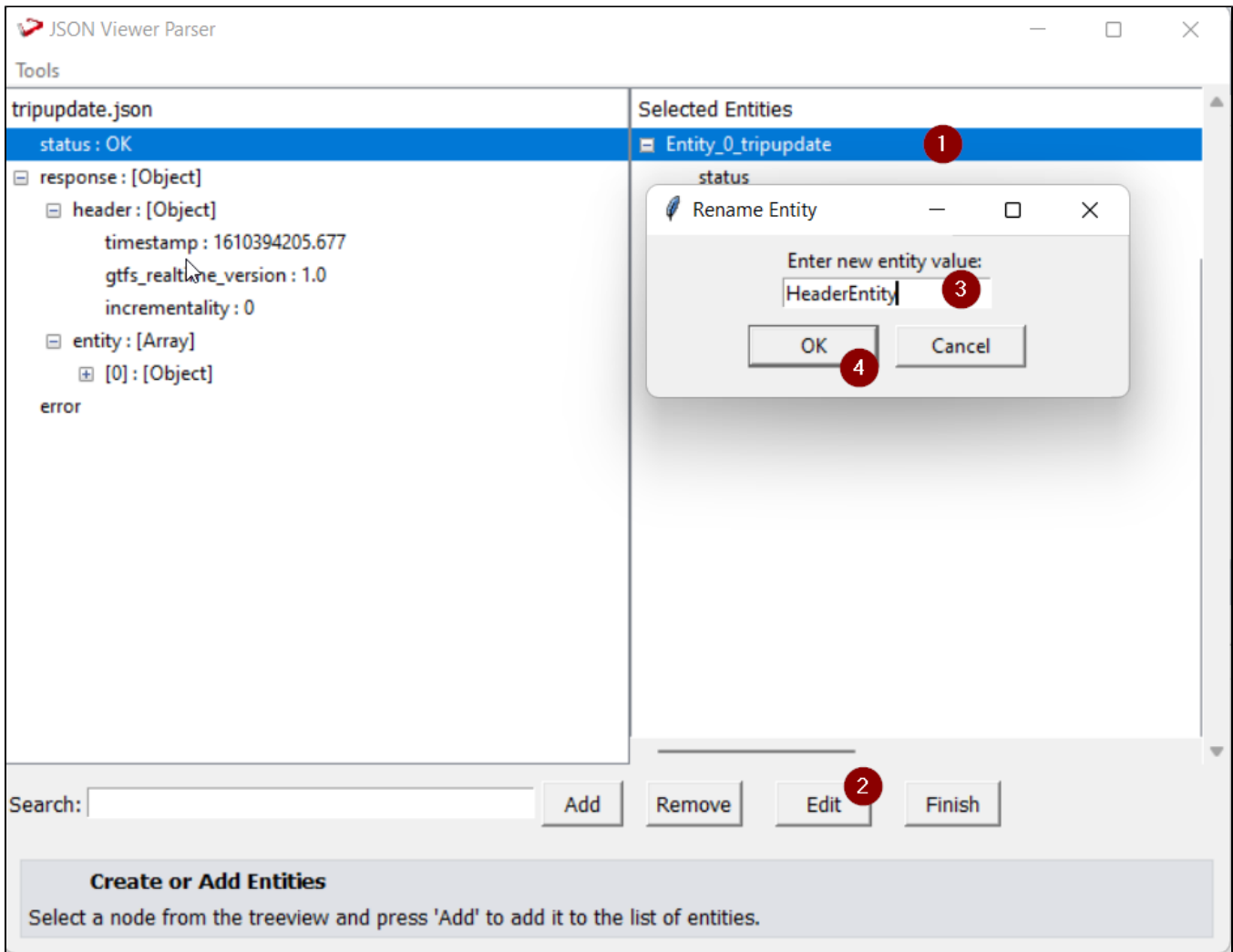
Note: Holding the "Ctrl" key on the keyboard while clicking on different nodes allows the user to select multiple nodes.



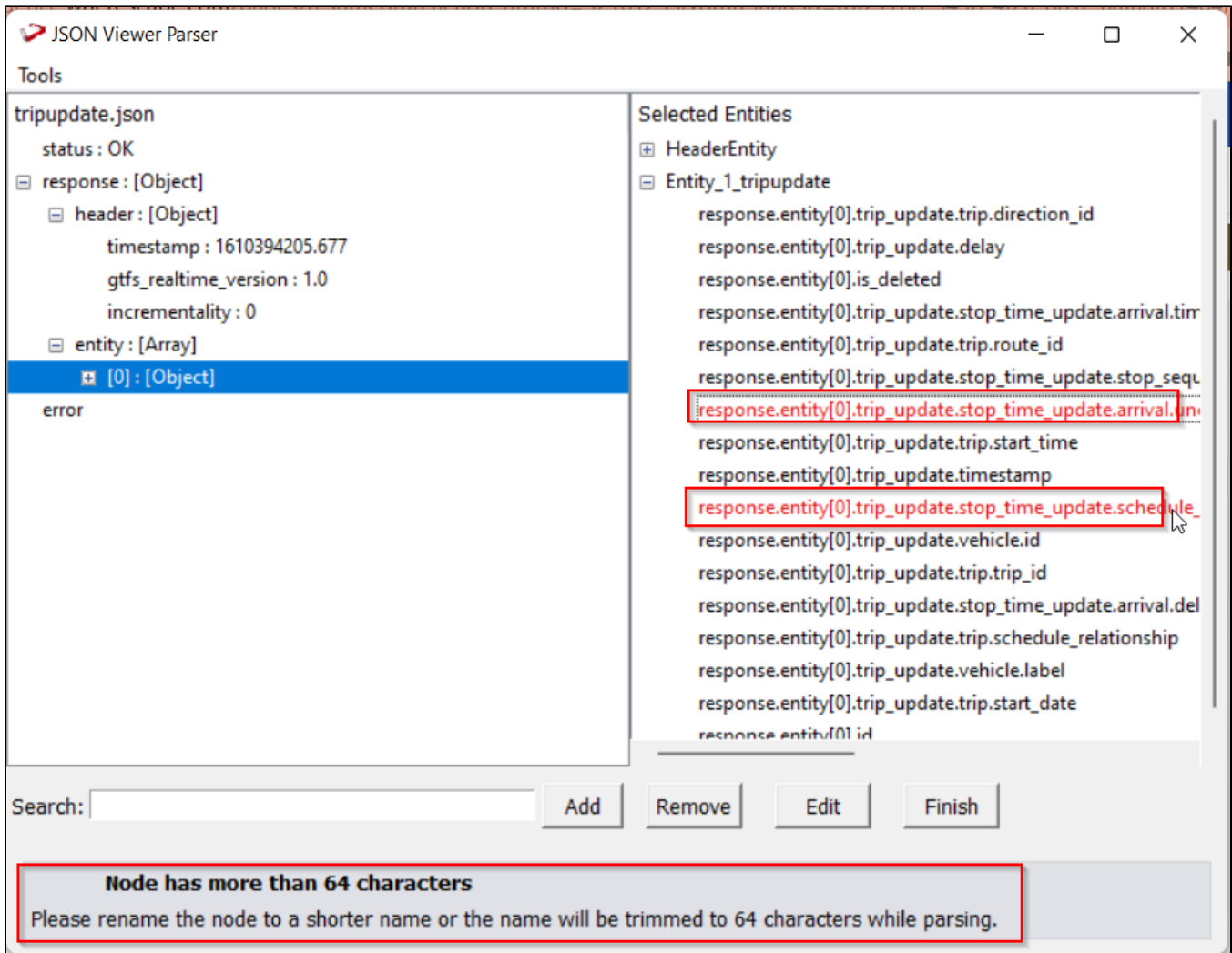
To add a new node from a JSON tree to an entity that has already been created. Select the node in the JSON tree to which the node should be added (Example: Entity_0), select one or more nodes, and click "Add" to add the selected node to the selected entity.



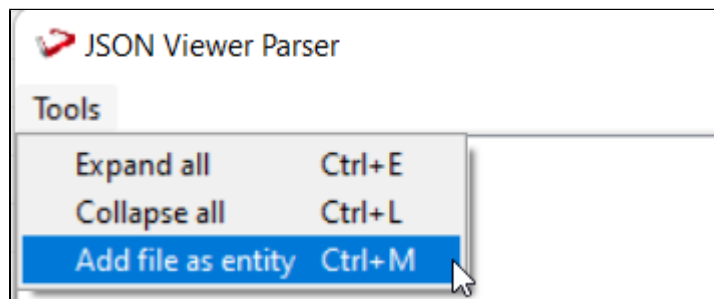
To edit name of Entity, Select the entity and press the "Edit" button to change the name. This will open a window with a text box where you may type in the new name for that object and then click "Ok."



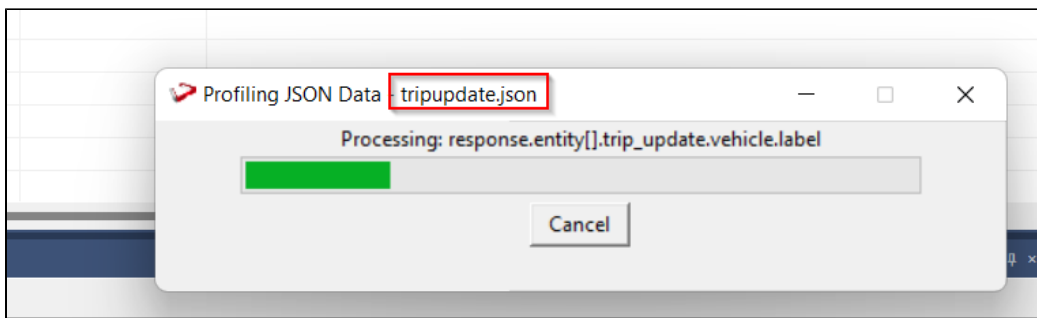
In the selected entities pane, nodes in entities with more than 64 characters are colored "red." In WhereScape RED, the names of these "red" colored nodes are trimmed.



To add complete file for profiling select below option from Tools Menu



After selecting all the entities and files options, progress of the profiling will appear with the progress bar and can be canceled at any point.



The working for XML parser is similar to JSON parser explained above.

Known issues

1. Load table properties are not automatically transferred from source models to RED export models. Updated MCRs in 3D 9.0.1.0 solve this issue.
2. Script language settings need to be set manually.
3. Requires manual installation process.