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 xlsxwriter xlrd pyarrow fastparquet pyorc avro avro_python3
jsonpath_ng openpyxl Pillow pyarrow xmltodict 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.
- 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.
- 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: pyd. Command: Python "\$SCRIPT_NAME\$"

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

Add a new connection				×
Provide a name for the new con	nection and select its type			
Connection name: Amazon S Connection type: Amazon				
Connection Browse Script:	Browse_File_Parser			
Load Table UI Configuration:	Load from Windows Data Parser			
Secret Key	****	?		
Access Key	AKIA533YNQGESIVPXJUD	?		
\$3 Bucket Name	iderabucket	•		
\$3 Region	us-east-1	3		
			Test Connection	
			Save Connection	
			<u>N</u> ext > <u>C</u> ance	1

Azure Data Lake Storage Gen2 Connection Setup

ide a name for the new connection and select its type onnection name: Azure Data Lake Storage Gen2 Connection type: Azure Data Lake Storage Gen2 Connection Browse Script Browse_File_Parser Connection Browse Script Load from Windows Data Parser Azure Data Lake Storage Gen2 Account Access Key(Account Key) Azure Data Lake Storage Gen2 Account SAS Token Azure Data Lake Storage Gen2 Account SAS Token Test Connection Save Connection	Connection name: Azure Dat	nection and select its type			
onnection name: Azure Data Lake Storage Gen2 Connection type: Azure Data Lake Storage Gen2 Connection Browse Script Browse_File_Parser Load Table UI Configuration: Load from Windows Data Parser Azure Data Lake Storage Gen2 Account emeadatalake zure Data Lake Storage Gen2 Account Access Key(Account Key) Azure Data Lake Storage Gen2 Account SAS Token Azure Data Lake Storage Gen2 File System emeademo Test Connection	Connection name: Azure Dat	nection and select its type			
Connection type: Azure Data Lake Storage Gen2 Connection Browse Script: Browse_File_Parser Load Table UI Configuration: Load from Windows Data Parser Azure Data Lake Storage Gen2 Account emeadatalake Azure Data Lake Storage Gen2 Account SAS Token Azure Data Lake Storage Gen2 File System emeademo Test Connection					
Connection type: Azure Data Lake Storage Gen2 Connection Browse Script: Browse_File_Parser Load Table UI Configuration: Load from Windows Data Parser Azure Data Lake Storage Gen2 Account emeadatalake Azure Data Lake Storage Gen2 Account SAS Token Azure Data Lake Storage Gen2 File System emeademo Test Connection					
Connection type: Azure Data Lake Storage Gen2 Connection Browse Script: Browse_File_Parser Load Table UI Configuration: Load from Windows Data Parser Azure Data Lake Storage Gen2 Account emeadatalake Azure Data Lake Storage Gen2 Account SAS Token Azure Data Lake Storage Gen2 File System emeademo Test Connection		a Lake Storage Gen2			
Connection Browse Script Browse_File_Parser Load Table UI Configuration: Load from Windows Data Parser Azure Data Lake Storage Gen2 Account emeadatalake Zure Data Lake Storage Gen2 Account Key) Azure Data Lake Storage Gen2 Account SAS Token Azure Data Lake Storage Gen2 File System emeademo Test Connection					
Load Table UI Configuration: Azure Data Lake Storage Gen2 Account emeadatalake azure Data Lake Storage Gen2 Account Key) Azure Data Lake Storage Gen2 Account SAS Token Azure Data Lake Storage Gen2 File System emeademo Test Connection	connection type.				
Azure Data Lake Storage Gen2 Account emeadatalake ? zure Data Lake Storage Gen2 Account Access Key(Account Key) ? ? Azure Data Lake Storage Gen2 Account SAS Token ? ? Azure Data Lake Storage Gen2 File System emeademo ? Test Connection ************************************				•	
zure Data Lake Storage Gen2 Account Access Key(Account Key) Azure Data Lake Storage Gen2 Account SAS Token Azure Data Lake Storage Gen2 File System emeademo Test Connection		Load Table UI Configuration:	Load from Windows Data Parser	•	
Azure Data Lake Storage Gen2 Account SAS Token Azure Data Lake Storage Gen2 File System Test Connection Test Connection	Az	ture Data Lake Storage Gen2 Account	emeadatalake		۲
Azure Data Lake Storage Gen2 File System emcademo (7) Test Connection	-				۲
Test Connection	Azure Data La	ike Storage Gen2 Account SAS Token		••••	۲
	Azure	e Data Lake Storage Gen2 File System	emeademo		0
Save Connection					Test Connection
					Save Connection

Google Cloud Connection Setup

(******		~
Add a new connection		×
Provide a name for the new connection and select its type		
Connection name: Google Cloud Storage		
Connection type: Google Cloud Storage		
Connection Browse Script: Browse_File_Parser		Ì
Load Table UI Configuration: Load from Windows Data Parser		
Google Storage Bucket Name wtd-bucket	•	
Google Storage Project Name wtd-development	•	
	Test Connection	
	Save Connection	
		1
	Next > Cance	
		21

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

Cloud Browser

Amazon S3 File Browser		- 🗆 X
File Help		
File Path: ./		Back
File	Туре	Size
cities_LtoR.txt	txt	21147
cities_StoZ.txt	txt	19875
X customer-changes.xml	xml	1263
customer.csv	CSV	283
X customer.xml	xml	1237
customer.xsd	xsd	891
] dataassingleline.json	json	219
] dataastree.json	json	321
file_example_CSV_5000 - Copy.csv	CSV	284042
file_example_CSV_5000.csv	csv	284042
Add Remove R	efresh	Browse Details: 46 File and 22 Folder Found
File	Туре	Size
Cities.txt t	xt	48096
customer.xml x	ml	1237
file_example_CSV_5000.csv of	SV	284042
	Downloading 1 of 3	Cancel OK

Select and click Add to copy files to staging area.
 Click Back to navigate to previous directory
 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: Windows Data Parser Connection type: Windows Data Parser Connection Browse Script: Browse_File_Parser Load Table UI Configuration: Load from Windows Data Parser	Test Connection Save Connection
	<u>N</u> ext > <u>C</u> ancel

Browse Parser

Choose parser as per file type

Options to Parse File: Cities.txt		_	\times
Field Delimiter Record Delimiter Field Enclosure Delimiter Row Limit For Profiling SheetName	 \\n " 100	Parser Skip Lines Header Record Header Lines Coefficient	_
Save Options For All Files city_code city_name state_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 2001 state d 14711		OK	

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		_		\times
Field Delimiter		Parser	DELIMITED	~
Record Delimiter	\n	Skip Lines	0	
Field Enclosure Delimiter	"	Header Record	V	
Row Limit For Profiling	100	Header Lines	0	_
SheetName	None 🗸	Coefficient		_
Save Options For All File	s	ОК		
situ code leitu name letate cod				_
city_code city_name state_code 278 Assen 1	ercountry_code			
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.

SON Viewer Parser	_	×
Tools		
tripupdate.json		-
status : OK		
error		
		 Ŧ
Search: Add		
Create or Add Entities		
Select a node from the treeview and press 'Add' to add it to the list of entities.		

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.

SON Viewer Parser	– o ×
Tools	
tripupdate.json	Selected Entities
status : OK	Entity_0_tripupdate
response : [Object]	response.header.timestamp
header: [Object]	
timestamp : 1610394205.677	
gtfs_realtime_version : 1.0	
incrementality : 0	
entity: [Array]	
error	
Search: Add	Remove Edit Finish
Create or Add Entities	
Select a node from the treeview and press 'Add' to add it to the	list of entities.

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.

SON Viewer Parser			_	×
Tools				
tripupdate.json		Selected Entities		4
status : OK		Entity_0_tripupdate		
response : [Object]		response.header.gtfs_realtime_v	/ersion	
header: [Object]		response.header.timestamp		
timestamp : 1610394205.677		response.header.incrementality		
gtfs_realtime_version : 1.0				
incrementality : 0				
entity : [Array]				
error				
2				
00				
		4		
			-	
Search:	Add	Remove Edit Finis	h	
Create or Add Entities				
Select a node from the treeview and press 'A	Add' to add it to the	list of entities.		

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.

SON Viewer Parser	– – ×	
Tools		
tripupdate.json	Selected Entities	-
status : OK	Entity_0_tripupdate	
response : [Object]	status	
header: [Object]	response.header.gtfs_realtime_version	
timestamp : 1610394205.677	response.header.timestamp	
gtfs_realtime_version : 1.0		
incrementality : 0		
entity : [Array]		
error		
Search: Add	Remove Edit Finish	
Create or Add Entities		
Select a node from the treeview and press 'Add' to add it to the	e list of entities.	

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."

SON Viewer Parser					\times
Tools					
tripupdate.json		Selected Entities			-
status : OK		Entity_0_tripupdate	1		
 response : [Object] header : [Object] timestamp : 1610394205.677 gtfs_realtime_version : 1.0 incrementality : 0 entity : [Array] (0) : [Object] 		status Rename Entity Enter new entity HeaderEntity OK 4	value: 3 Cancel	×	
Search:	Add	Remove Edit	Finish		-
Create or Add Entities Select a node from the treeview and press 'Add' to add	d it to the	list of entities.			

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.

Tools tripupdate.json status : OK □ response : [Object] □ header : [Object] timestamp : 1610394205.677 gtfs_realtime_version : 1.0 inscrementality: 0	Selected Entities			
status : OK response : [Object] header : [Object] timestamp : 1610394205.677 gtfs_realtime_version : 1.0	 HeaderEntity Entity_1_tripupdate response.entity[0].trip_update.trip.direction_id response.entity[0].trip_update.delay response.entity[0].is_deleted 			
 □ response : [Object] □ header : [Object] timestamp : 1610394205.677 gtfs_realtime_version : 1.0 	Entity_1_tripupdate response.entity[0].trip_update.trip.direction_id response.entity[0].trip_update.delay response.entity[0].is_deleted			
⊨ header : [Object] timestamp : 1610394205.677 gtfs_realtime_version : 1.0	response.entity[0].trip_update.trip.direction_id response.entity[0].trip_update.delay response.entity[0].is_deleted			
timestamp : 1610394205.677 gtfs_realtime_version : 1.0	response.entity[0].trip_update.delay response.entity[0].is_deleted			
gtfs_realtime_version : 1.0	response.entity[0].is_deleted			
in cromontality (0				
incrementality : 0	response.entity[0].trip_update.stop_time_update.arrival.tim			
entity: [Array]	response.entity[0].trip_update.trip.route_id			
	response.entity[0].trip_update.stop_time_update.stop_sequ			
error	response.entity[0].trip_update.stop_time_update.arrival.un			
	response.entity[0].trip_update.trip.start_time			
	response.entity[0].trip_update.timestamp			
	response.entity[0].trip_update.stop_time_update.schedule_			
	response.entity[0].trip_update.vehicle.id			
	response.entity[0].trip_update.trip.trip_id			
	response.entity[0].trip_update.stop_time_update.arrival.del			
	response.entity[0].trip_update.trip.schedule_relationship			
	response.entity[0].trip_update.vehicle.label			
	response.entity[0].trip_update.trip.start_date			
	resnonse entity/01 id			
accent.				
Search: Add	Remove Edit Finish			
Node has more than 64 characters				
Please rename the node to a shorter name or the name will be	trimmed to 64 characters while parsing.			

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

> JSON Viewer Pars	ser	
Tools		
Expand all	Ctrl+E	
Collapse all	Ctrl+L	
Add file as entity	Ctrl+M	

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.

Profiling JSON Data tripupdate.json –	×	
Processing: response.entity[].trip_update.vehicle.label		
Cancel		ą

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

Known issues

- 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.
 Script language settings need to be set manually.
 Requires manual installation process.