

WhereScape Enablement Pack for Google BigQuery

This is a guide to installing the WhereScape Enablement Pack for Google BigQuery for WhereScape RED 8.6.1.x

Prerequisites

Before you begin the following prerequisites must be met:

1. A supported* version of SQL Server or Azure SQL with
 - Database created to house the RED Metadata Repository with 'SELECT','INSERT','UPDATE','EXECUTE' permissions
 - (optionally) a database for the Range Table DB
 - ODBC DSN's created for these DB's
2. A BigQuery environment with
 - At least one schema available to use as a RED Data Warehouse Target
3. BigQuery software installed
 - BigQuery ODBC driver (64-bit) Currently only Simba ODBC driver 2.3.2.1003 is supported.
 - Google Cloud SDK for utilities GSUTIL and BQ installed and on the Windows system path
 - Download and Install GoogleCloudSDKInstaller from <https://cloud.google.com/sdk/docs/quickstart>
 - ODBC DSN created to connect to your RED Data Warehouse on BigQuery
 - Google Big Query Catalog Name
 - At least one storage bucket needs to be created on Google Cloud Storage.
 - Permission Required : Upload,Read and Delete File
4. WhereScape RED version 8.6.1.0 or higher
 - Must be pre-installed with valid license key entered and EULA accepted
5. WhereScape Enablement Pack for BigQuery version 8.6.1.0 or higher
 - Downloaded and unpacked to a local temp folder
6. Python 3.8 or higher
 - 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
```

* : RED supports the following versions for the metadata repository: MS SQL SERVER 2012 to 2019 and Azure SQL DB

Enablement Pack Setup Scripts

The Enablement Pack Install process is entirely driven by scripts. The below table outlines these scripts, their purpose and if "Run as Administrator" is required.

#	Enablement Pack Setup Scripts	Script Purpose	Run as Admin	Intended Application
1	Setup_Enablement_Pack.ps1	Setup and configure a RED Metadata Repository for target database If RED repository exists then updates the repository with 1.Templates 2.Scripts 3.Extended Properties 4.Datatype Mappings 5.UI Configurations	Yes	New and Existing installations
2	install_WslPowershell_Modules.bat	Installs or updates WslPowershell Modules on this machine	Yes	New and Existing installations
3	install_WslPython_Modules.bat	Installs or updates WslPython Modules and required Python libs on this machine	Yes	New and Existing installations
4	import_python_templates.ps1	Imports or updates the Python Templates (in BETA) to a RED Metadata Repository. Also includes any Script Imports	no*	Existing installations
5	set_default_templates.ps1	Applies the RED Connection defaults in a RED Metadata Repository for Python or Powershell templates	no*	Existing installations

* Note

- Some systems executing Windows Powershell scripts are disabled by default, see troubleshooting for workarounds
- install_WslPowershell_Modules.bat is required for creation of pilot range table as scripts required for pilot table i.e create,pause and restart are currently in powershell only

Each Powershell script in the list above provides some help at the command line, this can be output by passing the "-help" parameter to the script.

Step-by-step guides

Setup and configure a new RED Metadata Repository for BigQuery

1. Double check Prerequisites section
2. Run Windows Powershell as Administrator :

```

Install Enablement Pack

Script 1 > Powershell -ExecutionPolicy Bypass -File .\Setup_Enablement_Pack.ps1
Script 3 > install_WslPython_Modules.bat
    
```

Note : To get help for parameter values run the script for help. For Example Powershell.exe -ExecutionPolicy Bypass -File .\Setup_Enablement_Pack.ps1 -help

Install or Update WhereScape Python Templates

1. Double check Prerequisites (specifically the Python prerequisites)
2. Run as Administrator - Script 3 - install_WslPython_Modules.bat, there are two steps in this script:
 - Step 1 - Installs the WhereScape WslPython modules to C:\Program Data\WhereScape\Modules\
 - Step 2 - Uses PIP to download/update required Python libraries - for offline install please see the required library list for Python in the Troubleshooting section.
3. Run - Script 5 - import_python_templates.ps1
4. Run - Script 6 - set_default_templates.ps1
 - choose "Python" when prompted

Set Connection defaults for a Template Set

1. Run - Script 6 - set_default_templates.ps1
 - choose "Python" if prompted

Source Enablement Pack Support

Source Pack Name	Supported By Google Big Query	Supported Features	Permissions Required Google Big Query
Google Cloud Storage	Yes	<ul style="list-style-type: none"> • RED version 8.6.1.0 supports load files from Google Cloud Storage to Google BigQuery tables . <ul style="list-style-type: none"> • Files are loaded from Google storage bucket to BigQuery tables 	
Google Drive	Yes	<ul style="list-style-type: none"> • RED version 8.6.1.0 supports load files from Google Drive to Google BigQuery tables . <ul style="list-style-type: none"> • Supports only if the authentication id is same for Google Drive and Google Big Query • Supports only CSV files load for any field delimiter. • Load is done through external table as supported by Google Big Query. • Python Package Required : pip install google-cloud-bigquery <p>Refer to the user guide for WhereScape Source Enablement Pack - Google Drive for more information</p>	<p>Following permissions are required to create and query an external table in BigQuery.</p> <ul style="list-style-type: none"> • bigquery.tables.create • bigquery.tables.getData • bigquery.jobs.create <p>Following predefined IAM roles include both bigquery.tables.create and bigquery.tables.getData permissions:</p> <ul style="list-style-type: none"> • bigquery.dataEditor • bigquery.dataOwner • bigquery.admin <p>Following predefined IAM roles include bigquery.jobs.create permissions:</p> <ul style="list-style-type: none"> • bigquery.user • bigquery.jobUser • bigquery.admin
Azure Data Lake Storage Gen2	No	None	
Amazon S3	No	None	

Post Install Steps - Optional

If you used the script 'install_New_RED_Repository.ps1' then the following optional post install steps are available

Configure Connections

There were two connections added that will optionally require your attention:

1. Connection: 'Database Source System' - this connection was setup as an example source connection,
 - open it's properties and set it up for a source DB in your environment
 - or you can remove it if not required

Enable Script Launcher Toolbar

There are a number of stand-alone scripts which provide some features, these scripts have been added to the Script Launcher menu but you will need to enable the menu toolbar item to see them.

To enable the Script Launcher menu in RED: Select menu item 'View->Toolbars->Script Launcher'

Troubleshooting and Tips

Run As Administrator

Press the Windows Key on your keyboard and start typing cmd.exe, when the cmd.exe icon shows up in the search list right click it to bring up the context menu, select "Run As Administrator"

Now you have an admin prompt navigate to to the folder where you have unpacked your WhereScape Red Enablement Pack to using the 'cd' command:

```
C:\Windows\system32> cd <full path to the unpacked folder>
```

Run batch (.bat) scripts from the administrator prompt by simply typing the name at the prompt and hit enter, for example:

```
C:\temp\EnablementPack>install_WsIPowershell_Modules.bat
```

Run Powershell (.ps1) scripts from the administrator prompt by typing the Powershell run script command, for example:

```
C:\temp\EnablementPack>Powershell -ExecutionPolicy Bypass -File .\install_New_RED_Repository.ps1
```

Notes: In the event you can not bypass the Powershell execution policy due to group policies you can instead try "-ExecutionPolicy RemoteSigned" which should allow unsigned local scripts.

Windows Powershell Script Execution

On some systems Windows Powershell script execution is disabled by default. There are a number of workarounds for this which can be found on the internet by searching the term "Powershell Execution Policy".

Here is the most common workaround which WhereScape suggests, which does not permanently change the execution rights:

Start a Windows CMD prompt as Administrator, change directory to your script directory and run the WhereScape Powershell scripts with this command:

- cmd:>Powershell -ExecutionPolicy Bypass -File .\<script_file_name.ps1>

Restarting failed scripts

Some of the setup scripts will track each step and output the step number when there is a failure. To restart from the failed step (or to skip the step) provide the parameter "-startAtStep <step number>" to the script.

Example:

```
Powershell -ExecutionPolicy Bypass -File .\<script_file_name.ps1> -startAtStep 123
```

Tip: to avoid having to provide all the parameters again you can copy the full command line with parameters from the first "INFO" message from the beginning of the console output.

Python requirements for offline install

Additionally to the base Python installation being required, the WhereScape Python Template set also requires certain additional Python libraries. The install scripts uses the PIP (package manager) to download these libraries, however for offline installs you will need to install the required libraries yourself.

Required Python libraries/add-ons:

- pywin32-ctypes
- python-tds

- pywin32
- glob2
- gzip-reader
- regex
- pyodbc

If a valid RED installation can not be found

If you have Red 8.5.1.x or higher installed but the script (install_New_RED_Repository.ps1) fails to find it on your system then you are most likely running PowerShell (x86) version which does not show installed 64 bit apps by default. Please open a 64 bit version of PowerShell instead and re-run the script.