

WhereScape Enablement Pack for Azure Synapse RED 9

This is a guide to installing the WhereScape Enablement Pack for Azure Synapse for WhereScape RED 9.0.x

Table of Contents

- [Prerequisites For SQL Server Metadata](#)
- [Prerequisites For Azure Synapse](#)
- [Enablement Pack Setup Scripts](#)
- [Step-By-Step Guide](#)
- [Post Install Steps](#)
- [Source Enablement Pack Support](#)
- [Troubleshooting and Tips](#)

Prerequisites For SQL Server Metadata

Before you begin the following prerequisites must be met:

- Create Database and ODBC DSN :
 - Supported* version of SQL Server or Azure SQL
 - A database to house the RED Metadata Repository.
Note : This needs to be Empty Database with optional permissions SELECT,INSERT,UPDATE,EXECUTE
 - A database for the Range Table DB (Optional)
- Software Installations
 - WhereScape RED version 9.0.1.1 or higher with valid license key entered and EULA accepted
 - WhereScape Enablement Pack for target database version 9.0.1.1 or higher
- Windows Powershell (64 bit) version 4 or higher
 - To check Windows Powershell Version:
 - Run below command in Windows Powershell

```
Get-Host | Select-Object Version
```

- Run below command in Command Prompt

```
powershell $psversiontable
```

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

Prerequisites For Azure Synapse

Before you begin the following prerequisites must be met:

1. Access to an Azure Synapse (Target Environment), with the following connectivity information:
 - a. Server Name
 - b. Database Name
 - c. User Name
 - d. Password
 - e. At least one schema available to use as a RED Data Warehouse Target
 - f. Connection String to connect to Azure Synapse
 - i. e.g Server=tcp:<server>,<port>;Database=<database>;User ID=\$USERNAME\$;Password=\$PASSWORD\$;
Trusted_Connection=False;Encrypt=True;Connection Timeout=30;
 - g. Blob Storage Setup
 - i. Azure Storage Account Name
 - ii. Azure Storage Account Access Key
 - iii. Azure Storage Container Name
 - iv. External Data Source Name
2. Azure Synapse software installed
 - Azure Synapse ODBC driver (64-bit)

- a. Add ODBC DSN for Azure Synapse Target Database
 - .Net Framework 4.8 or higher
 - Windows Powershell version 5 or higher
 - Azure Storage powershell functionality installed.
- a. Run these commands in "Windows PowerShell":
- 3. **Python 3.8 or higher**
 - Select "Add Python 3.8 to PATH" from installation Window
 - Pip Manager Install with command : `python -m pip install --upgrade pip`

```
[Net.ServicePointManager]::SecurityProtocol = [Net.SecurityProtocolType]::Tls12
Install-Module -Name Az.Storage -Force -AllowClobber
```

Note: make sure this is done using a 64-bit powershell terminal

* : 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 Azure Synapse 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 libraries on this machine	Yes	New and Existing installations
4	import_powershell_templates.ps1	Imports or updates the Powershell Templates to a RED Metadata Repository. Also includes any Script or Procedure 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 that on some systems executing Windows Powershell scripts is disabled by default, see troubleshooting for workarounds

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. For Example > `.\Setup_Enablement_Pack.ps1 -help`

Step-By-Step Guide

Setup and configure a new RED Metadata Repository for Azure Synapse

Run Powershell as Administrator:

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

Install or Update WhereScape Powershell Modules

Run Script As Administrator

```
Script 2 > install_WslPowershell_Modules.bat
```

Install or Update WhereScape Python Modules

Run Script As Administrator

```
Script 3 > install_WslPython_Modules.bat
```

There are two steps in this script:

1. Installs the WhereScape WslPython modules to C:\Program Data\WhereScape\Modules\
2. PIP to download/update required Python libraries - for offline install please see the required library list for Python in the Troubleshooting section.

Install or Update WhereScape Powershell Templates (For Existing Installations)

Run Script as Administrator

```
Script 2 > install_WslPowershell_Modules.bat  
Script 4 > import_powershell_templates.ps1  
Script 5 > set_default_templates.ps1
```

Note : Skip this step for new installations.

Set Connection defaults for a Template Set (For Existing Installations)

```
Script 5 > set_default_templates.ps1
```

Note: The upgrade will not overwrite the existing Data Type Mappings, UI configs. The user either needs to manually delete them or use startAtstep for continuing the script execution.

Important Upgrade Notes



If RED repository exists, it will prompt to upgrade the repository. This enablement pack will overwrite any existing Source Enablement Pack UI Configs:

Amazon S3	Load From Amazon S3
Azure Data Lake Storage Gen2	Load From Azure Data Lake Storage Gen2
Google Cloud	Load From Google Cloud

To ensure existing Source Enablement Pack connections and associated Load Tables continue to browse and load:

Go into UI Configuration Maintenance in RED prior to installing this Enablement Pack and rename the affected UI Configurations. While the updated Load Template will work with previous Source Enablement Pack's we recommend moving these previous versions of Load Tables to newly created Parser based connections following this install. The earlier versions of the Source Enablement Pack will be deprecated following this release.

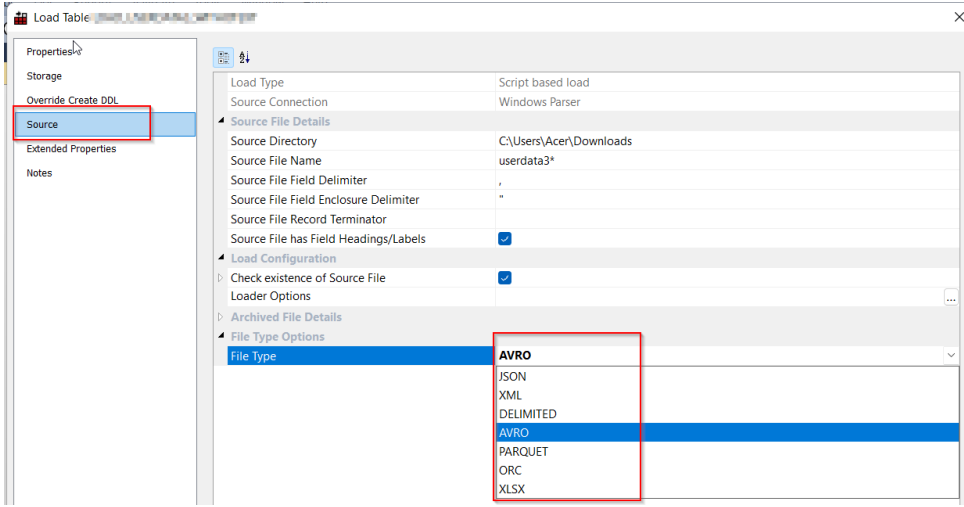
Post Install Steps

Configure Connections

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

1. Azure Target Connection : Set below extended properties
 - Blob Storage Account
 - Blob Storage Access Key
 - Blob Storage Container
 - Blob Data Source
2. Connection: 'Database Source System' - this connection was setup as an example source connection --Optional
 - open it's properties and set it up for a source DB in your environment
 - or you can remove it if not required

Source Enablement Pack Support

Source Pack Name	Supported By Azure Synapse	Supported Features
Amazon S3	Yes	Install Powershell package: <ul style="list-style-type: none"> [Net.ServicePointManager]::SecurityProtocol = [Net.SecurityProtocolType]::Tls12 Install-Module -Name AWSPowerShell.NetCore Install-Module -Name AWS.Tools.Installer
Azure Data Lake Storage Gen2	Yes	Load file from Azure Data Lake Storage Gen2 bucket to Azure Synapse Install Python package: <ul style="list-style-type: none"> pip install azure-storage-file-datalake Required Properties: <ul style="list-style-type: none"> Azure Data Lake Storage Gen2 Account Name Azure Data Lake Storage Gen2 Access Key Azure Data Lake Storage Gen2 SAS Token Azure Data Lake Storage Gen2 Container Name Azure Data Lake Storage Gen2 Directory Name (Created in Container)
Google Drive	No	None
Google Cloud	Yes	Please refer the official link for steps to download gsutil: https://cloud.google.com/storage/docs/gsutil_install
Windows Parser	1.CSV 2.Excel 3.JSON 4.XML 5.AVRO 6.ORB 7.PARQUET	Refer to Windows Parser Guide. Load Template ,Source Properties will have option to select parser type to load the files. 

NOTE:

File Types : ORC

For copying ORC data to Synapse, the complex data types are not currently supported (STRUCT, MAP, LIST, UNION), this is a current limitation in Azure Synapse.

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_WslPython_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 .\Setup_Enablement_Pack.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 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
- databricks
- databricks-sql-connector

If a valid RED installation can not be found

If you have Red 8.5.1.x or higher installed but the script (Setup_Enablement_Pack.ps1) fails to find it on you 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.

Amazon AWS PowerShell command not found

Some systems require a force initialization of AWS modules after installing Amazon S3 PowerShell modules, whereas other systems only require a restart. Run the PowerShell command below as an administrator to initialize Amazon modules.

1. Import-Module AWS.Tools.Installer
2. Set-AWSCredentials

PowerShell module error

Restart the system once all module installation is complete and PowerShell module errors persists

Export Table

Export file is created based on selected file format, and not from file attributes (export file delimiter and optionally enclosed by)

Authentication methods for advance connect

The screenshot shows the Synapse workspace interface for Microsoft Entra ID settings. The left sidebar contains navigation options: Overview, Activity log, Access control (IAM), Tags, Diagnose and solve problems, Settings (Microsoft Entra ID, Properties, Locks), and Analytics pools. The main content area is titled 'Microsoft Entra ID' and includes a 'Microsoft Entra admin' section with fields for 'Microsoft Entra admin' (set to 'SQL Server Access') and 'Admin Object/App ID' (set to '4ce47c64-724a-4caa-98d9-858a39bbcadf'). Below this is the 'Microsoft Entra ID authentication only' section, which contains a checkbox labeled 'Support only Microsoft Entra ID authentication for this workspace' that is currently checked. A red box highlights this checkbox.

1. When using different authentication methods other than SQL Server authentication, this option needs to be enabled from Synapse workspace which will allow only Azure Active Directory(AAD) authentication and SQL Server authentication will no longer work for the same.
2. If we want to user SQL Server authentication, this option needs to be unchecked.