



# Create your own Matillion ETL Extract connector

This article is specific to the following platforms - Redshift - Snowflake - BigQuery.

## Reference :

<https://docs.matillion.com/metl/docs/5505487/>

## Overview

Using Matillion ETL's [API Extract component](#) in conjunction with the [Manage Extract Profiles wizard](#) empowers users to create their own Matillion ETL data connector and load data from any data source using the relevant API.

This guide, with a supplementary video guide at the bottom, will explain how to create a custom Matillion ETL data connector.

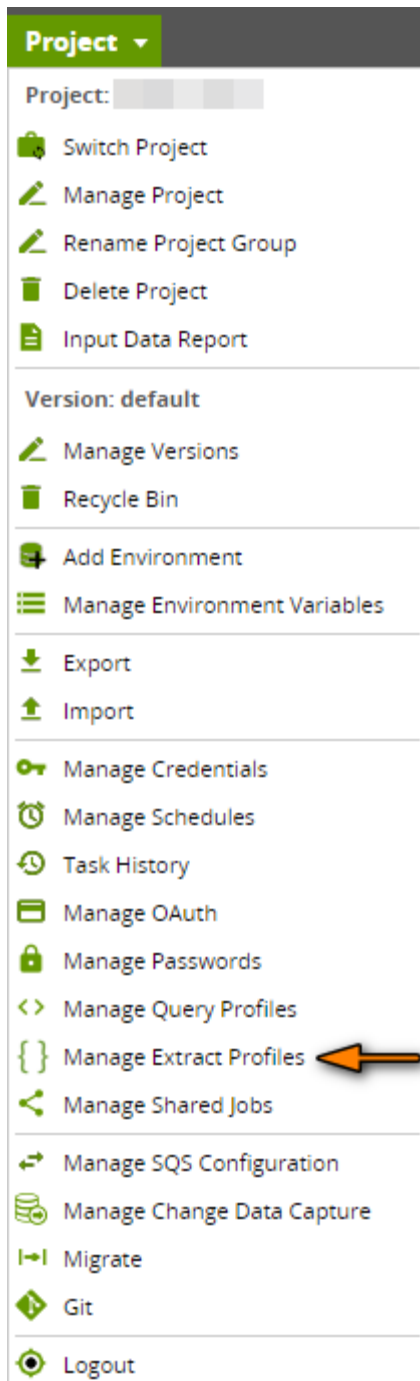
---

## Getting Started - Creating A New Extract Profile

First, an extract profile must be set up.

1. Click *Project*, and then click *Manage Extract Profiles* to open the Manage Extract Profiles dialog.

ANAL



In this dialog, users can manage all of their extract profiles.

2. In the Manage Extract Profiles dialog, click to add a new extract profile. This action launches the Add Extract Profile dialog, in which users must provide a name for their new extract profile. We recommend a human-readable name that is easy to identify.

In this example, the extract profile is named Coin Price.

3. Next, click *Setup* to begin configuring your new extract profile.



**Add Extract Profile**

Profile Name:

Coin Price

OK

Cancel

Setup

4. In the Configure Extract Profiles dialog, click *New Endpoint* to enter the endpoint setup wizard.

**Configure Extract Profiles**

Endpoints

There are currently no endpoints.

New Endpoint

Endpoint Config

Profile Name:

Endpoint Name:

Description:

URI:

Parameters

Parameter Name	Value	Type	Constant
There are currently no parameters.			

Body

OK

Cancel

---

## 1. Source Details

The first page of the wizard requires basic source information:

### *Profile Name*

This setting has already been set, it is the extract profile name—in this case, Coin Price.

---



## Endpoint Name

Specify a name for the endpoint to be created—this is the data source.

## Description

Provide a human-readable description of the endpoint for reference. This setting is optional.

Click *Next*.

Configure Extract Connector

1

2

3

4

Source Details

Endpoint Configuration

Response Configuration

Review

Source Details

Profile Name:

Endpoint Name:

Description:

Information

This dialog will take you through a series of steps to produce an API Extract datasource. We use the options you enter here to obtain sample data from an API and publish the basics of a data source that will be used by the API Extract component to get semi-structured data from the API and store it ready to import into your data warehouse

**Source Details:** Here we just need the name of the data source that you wish to configure. You can also provide a description

**Endpoint Configuration:** In the URI field, enter the URI of the API call you want to make. Configure authentication and any required body and/or header, query or URI parameters. Click send and then validate the returned JSON.

**View Fields:** From here, you can view the structure of the data returned from the API. We will try to identify the repeat element from the data, or allow you to specify it if there are multiple options. You can also configure paging by selecting one of the supported paging schemes and providing the required parameters.

**Review:** Here you can review the configuration entered that we will use to publish your API Extract Profile data source. Publishing the configuration will make the data source immediately available for your API Extract profile.

Cancel

Back

Next

## 2. Endpoint Configuration

The second page of the wizard requires information about the API call you wish to make.

Specify the method. Users can select between *GET* and *POST*.

Input the endpoint URI. This will be available from the source of the API. In this case, the endpoint is: `https://api.coindesk.com/v1/bpi/currentprice.json`



---

### *Params*

Specify any parameters to validate the API call. Users can pass a Parameter Name, Value, Type, and specify whether the parameter is a Constant or not.

---

### *Auth*

Tick *Enable Authentication* if authentication is required.

Then, select "Username and Password" or "API Key" from the first dropdown. Input the credentials into the correct fields.

Select between Basic and Bearer in the second dropdown.

---

### *Body*

Specify, if required, the body of the request.

---

### *Response*

Here the API call response is displayed. When a user clicks *Send*, the Response tab is automatically brought into view. The following message is displayed for validated API calls:

JSON File validated successfully with no errors.

---

### *Log*

A log of activity.

---

After specifying the method, the URI, and any required elements, click *Send*. As mentioned above, the Response tab will be displayed. Upon validation, click *Next*.



Configure Extract Connector

1

Source Details

2

Endpoint Configuration

3

Response Configuration

4

Review

Endpoint Configuration

GET

https://api.coindesk.com/v1/bpi/currentprice.json

Send

ParamsAuthBodyResponseLog

1

[{"time":{"updated":"Aug 20, 2020 10:07:00 UTC","updatedISO":"2020-08-20T10:07:00+00:00","updateduk":"Aug 20, 2020 at 11:07 BST"},"disclaimer":"This data was produced by the Bitcoin.info team"}]

Validate

JSON File validated successfully with no errors.

Cancel

Back

Next

### 3. View Fields

On the third page of the wizard, users can view the structure of the data returned from the API. Matillion ETL will attempt to identify the Repeating Element from the data. However, users can also specify this themselves.

Users can also configure paging by first toggling *Paging* to "ENABLED" and then selecting one of the supported paging schemes and providing the required parameters.



Click *Next*.

Configure Extract Connector

1

Source Details

2

Endpoint Configuration

3

Response Configuration

4

Review

Response Configuration

Select Fields

Repeating Element:

1

1

bpi STRUCT

1

EUR STRUCT

rate\_float DOUBLE

code TEXT

description TEXT

rate TEXT

symbol TEXT

1

USD STRUCT

rate TEXT

description TEXT

code TEXT

rate\_float DOUBLE

symbol TEXT

1

GBP STRUCT

rate TEXT

description TEXT

symbol TEXT

Paging

☒ DISABLED

Strategy

Select paging type

Cancel

Back

Next

## 4. Review

The final page of the wizard is for reviewing the the configuration of the new endpoint. Publishing the configuration will make the data source immediately available for your API Extract profile.

Click *Back* to make changes. Otherwise, click *Finish* to complete the API extract profile setup.





Configure Extract Connector 1

1 Source Details 2 Endpoint Configuration 3 Response Configuration 4 Review

Review

**Endpoint Config**

Profile Name:

Endpoint Name:

Description:

URI:

**Parameters**

Parameter Name	Value	Type	Constant
There are currently no parameters.			

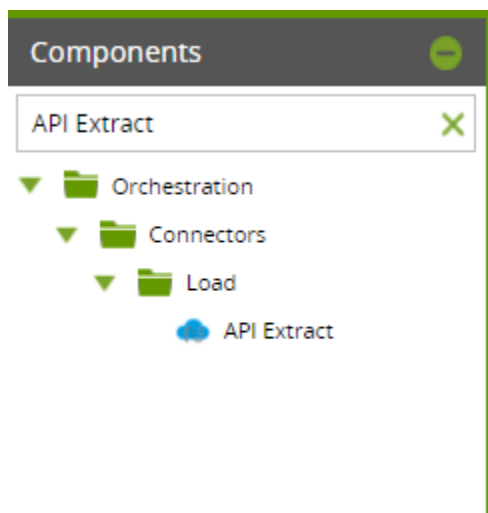
**Body**

Body Enabled:

Cancel Back Finish

## Creating the Orchestration Job

1. In a new Orchestration Job, add the API Extract component onto the Matillion ETL canvas.



2. Click onto the API Extract component, and click into the *API* property. From the dropdown menu, select the newly created API endpoint—in this example, it's Coin Price.





Properties | Export | Help

API Extract

Parameters contain errors.

Name	Value	Status
Name	API Extract	OK
API	Sample	OK
URI Params	end_date, apiversion, gro...	Grid contains empty
Query Params		OK
Header Params		OK
Username		Input required.
Password	*****	OK
Page Limit	1	OK

API

Coin Price

OK Cancel

3. Click into the *Data Source* property and from the dropdown select the data source. The data source is the Endpoint Name from the Manage Extract Profiles wizard—in this case, it is also named Coin Price.

Properties | Export | Help

API Extract

Component requires 1 input

Name	Value	Status
Name	API Extract	OK
API	Coin Price	OK
Data Source	Task History	Unrecognised option
URI Params	end_date, apiversion, gro...	Unrecognised option end_time
Query Params		OK
Header Params		OK
Page Limit	1	OK

Data Source

Coin Price

OK Cancel

Any parameters configured during Endpoint Configuration will be populated in the corresponding Params properties.

5. Select the storage location for your data.

6. Once the remaining properties are configured, right click the API Extract component, and click *Run Job*. Alternatively, with the component selected, press *CTRL + ENTER*.

---

## Transforming the Data

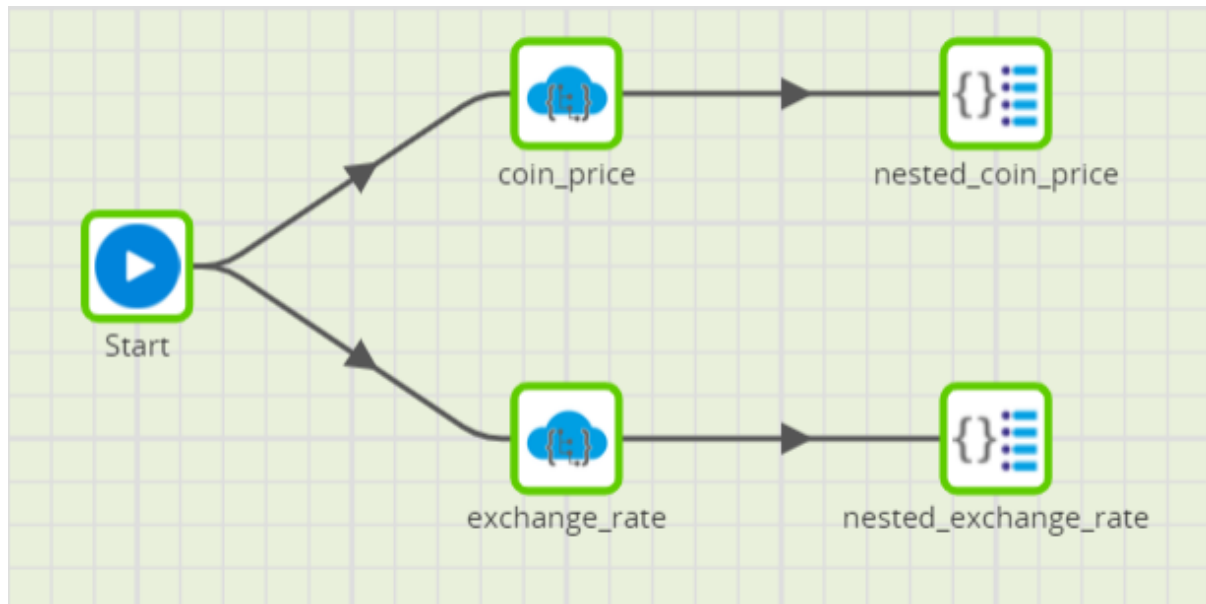
In the next section of this example, we will use a Matillion ETL Transformation Job to calculate the Candian Dollars price of Bitcoin.

First, another API profile must be set up for the current US Dollars to Canadian Dollars exchange rate. The setup of this extract profile follows the same steps as outlined so far in this article for the previous API extract profile. The endpoint for this second extract profile is <https://api.exchangerate-api.com/v4/latest/USD>



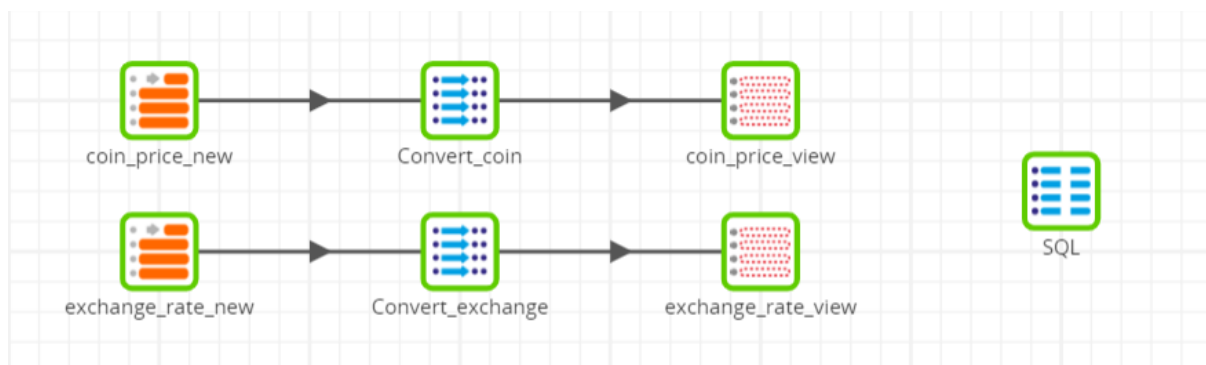
For the Orchestration Job, we have a [Start component](#) connected to two [API Extract](#) components. The first is for the Coin Price API; and the second is for the Exchange Rates API.

Since these data staging components are Extract components, they may return nested data that requires flattening.



Next, a Transformation Job is created. In this job, there are two [Table Input](#) components, one calling each new table made in the Orchestration Job. Then, a [Convert Type](#) component is used to convert the data type of the columns required to later run an SQL query that will calculate the Canadian Dollars price of Bitcoin comparative to USD. Finally, a [Create View](#) component is connected to each Convert Type component, outputting a view definition to our database rather than writing the data to a physical table.

Finally, there is an [SQL component](#) to run our SELECT query.



The SQL query and the data returned can be found below, where we have calculated the price of Bitcoin in Canadian Dollars.

SQL Query

Variables

Name	Default value
bpi_usd_rate_float	
rates_cad	
cad_coin_price	

Run

bpi_usd_rate_float	rates_cad	cad_coin_price
11786.215399999999	1.3165990000000001	15517.719409424601

Manage Variables

Update Component

OK

Cancel

```
1 SELECT
2 t1."bpi_usd_rate_float",
3 t2."rates_cad",
4 (t1."bpi_usd_rate_float" * t2."rates_cad") as CAD_Coin_Price
5 FROM "exchange_rate_new" t2, "coin_price_new" t1
6 group by 1,2
```

Below, users can copy the SQL query for their own work if required.

SELECT

```
t1."bpi_usd_rate_float",
```

```
t2."rates_cad",
```

```
(t1."bpi_usd_rate_float" * t2."rates_cad") as CAD_Coin_Price
```

FROM "exchange\_rate\_new" t2, "coin\_price\_new" t1

group by 1,2