Date :  26-02-2024

web scraper Api

version: 1.0.0

# **Overview**

The Web Scraper API provides web scraper services to retrieve HTML and plaintext data from any website URL.

The Web Scraper API enables developers to build applications that extract HTML and plaintext data from websites, facilitating tasks such as news aggregation, price monitoring for e-commerce, or gathering information for research purposes. It streamlines data acquisition processes, offering businesses valuable insights into market trends, competitor activities, and user behavior with minimal manual intervention.

# **Getting Started**

## **Prerequisites**

* Volt Foundry
* [**Api Ninjas Account**](https://api-ninjas.com/api/webscraper)
* API Key associated with your account.

## **Importing the adapter**

**To import the Data Adapter to Volt Foundry, do the following:**

1. Sign in to the  [HCL Foundry](https://manage.hclvoltmx.com/).
2. From the left navigation menu, select **API Management**.
3. In **API Management**, select **Custom Data Adapters**.  
   A screenshot of a computer

   Description automatically generated
4. Click **IMPORT** to import a custom data adapter.  
   A blue and white box with text

   Description automatically generated
5. On the Import Data Adapter dialog box, click browser to import.  
   A screen shot of a computer

   Description automatically generated
6. Select **Web Scraper API zip** file and click **IMPORT**.

After you import the data adapter, Volt Foundry opens a window that shows the metadata of the data adapter.

A screenshot of a web scraper

Description automatically generated

After you import the data adapter, you can view it on the Custom Data Adapters page and use it to create services on Volt Foundry.

A screenshot of a computer

Description automatically generated

## **Creating an Integration service**

After you import the data adapter into Volt Foundry, you can use it to create an Integration Service.

Follow the given steps to create an Integration service using the **Web Scraper API** Adapter.

1. Sign in to the [HCL Foundry](https://manage.hclvoltmx.com/).
2. From the left navigation menu, select **API Management**.
3. In **API Management**, select **Integration**.

A screenshot of a computer

Description automatically generated

1. To create a new service, click the **+** button or the **CONFIGURE NEW** button.

A screenshot of a computer

Description automatically generated

1. On the Service Definition tab, select the service type as**Web Scraper API,** and click **SAVE**.

A screenshot of a computer

Description automatically generated

Alternatively, you can also create a Foundry app and create an Integration service inside it.

## **Creating and Executing operations**

After you create an integration service, you can create and execute operations using the service.

#### **Creating an Operation**

* In **API Management/Foundry app you created**, in the **Integration** section, select the service that you created.
* After you select the service, navigate to the **Operation List** tab.  
  A screenshot of a computer

  Description automatically generated
* From the drop-down list, select an operation that you want to execute, and click **ADD OPERATION**.

A screenshot of a computer

Description automatically generated

#### **Executing an Operation**

* From the **Operations List** tab, in the **Configured Operations** section, select the operation you want to execute.

A screenshot of a computer

Description automatically generated

* On the Operation Page, in the **Request Input** tab, in the **Body** section, for the following parameters, specify the values in the respective boxes under the **Test Value** column.

**Mandatory Parameters**

* + **X-Api-Key:**  API Key associated with your account.

To get api key – Login in to [**Api Ninjas Account**](https://api-ninjas.com/profile)-> Go to My Account-> API Statistics -> You can view your api key.

* + **url:** URL to scrape.

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

* Select a run-time environment and click **Save and Fetch Response** to get a response based on your inputs.

A screen shot of a computer code

Description automatically generated

## **Publishing your application**

If you want to use the services in client applications, you need to publish an app to a run-time environment. You can create the service (as described above) in an application or import the service into an application and publish the application.

# **References**

## Endpoint Documentation

## **Requests Input Parameters**

|  |  |
| --- | --- |
| **Parameter** | **Description** |
| X-Api-Key | **[Required] [Header]** API Key associated with your account. |
| url | **[Required]** URL to scrape. |
| text\_only | **[Optional]** whether to only extract visible text (ignores HTML tags and metadata). Must be either true or false. Default is false. |
| user\_agent | **[Optional]** User-Agent string to use in the request header. |

### **Response Output Parameters**

|  |  |  |
| --- | --- | --- |
| **Object Property** | **Data Type** | **Description** |
| data | String | The HTML content retrieved from the requested URL. |

**Reference Document:**  [Web Scraper API - API Ninjas (api-ninjas.com)](https://api-ninjas.com/api/webscraper)

# **Revision History**

Adapter version 1.0.0:

## **Known Issues**

* No Known Issues

## **Limitations**

* No limitations