Scraper - Automatic Content Crawl and Post Plugin for Wordpress

Scraper is a Wordpress plugin that **copies content** and **posts automatically** from any web site. With tons of useful and unique features, Scraper takes content creating process to another level.

This documentation contains details about visual editor, plugin functionalities and use cases with examples. Please contact to our support team, if you need specific questions or have problem with plugin or service.

Getting Started

1. Installing and Plugin Licence
2. Third Party Licences
3. Creating Your First Task
4. Managing tasks

Task Features in Advance Level

5. Single Scraping
6. Serial Scraping
7. RSS Scraping
8. Data Fields
9. WooCommerce and Product Setup
10. Custom Fields
11. Transform Functions
12. Variables
13. Using Google Translate

15. Using Custom Xpath
16. Find and Replace
17. Mathematical Functions and Price Calculations
18. Striping and Cleaning Content
19. Shortcodes
20. Attribute Functions
21. Defining Post Tags
22. Defining Post Categories
23. Updating a Post with Scraper

Other Details About Plugin and Service

Installing and Plugin Licence

After your purchase on Codecanyon, Envato will provide an purchase code for plugin. You could download this code on Codecanyon downloads page as well.

You'll be using this code on plugin licencing after installing plugin.

Download plugin files and upload them to your Wordpress site via plugin upload form or with uploading it to wp-content/plugins directory.

After installing plugin on your Wordpress site, Scraper title will appear on your plugins list. Click "Activate" button for Scraper plugin.

Once you activate Scraper plugin for your web site, you should licence your plugin for your domain. Scraper section will appear on your menu, click on it and go to "Licence" section.

There are different purchase code or access token fields on this page. You should enter "Purchase Code" field to licence Scraper plugin.

Once you enter your purchase code, plugin will be licenced to your web site. If you need to move your domain or reset your token, you could contact to our support team regarding that issue.

Third Party Licences

Scraper has built-in features like content translation. These features requires third-party licences. You could enable and use them with plugin easily.

To enable translation, plugin requires Google Translate API. You could get Google translate API access key on Google API Console.

Please follow this tutorial to get Google API key for your web site.
https://cloud.google.com/translate/docs/quickstart

Creating Your First Task

We'll be creating a simple task on Scraper with Codecanyon Wordpress category scraping.

The page URL that we are going to scrape, is
https://codecanyon.net/category/wordpress/add-ons

To create your first task on Scraper, please follow introductions below :

1. Click **"Create New Task"** on Scraper plugin menu.
2. Paste the URL that you are going to scrape on **"Page URL"** field.
3. Click **"Fetch"** button. Plugin automatically scan the page and get you a preview.
4. On right side of visual editor, there is an option for **"Serial Items"**. You should show plugin your serial items for serial scraping. Click **"crosshair icon"** button besides the **"Item's Path"** field. And click any title on preview area. Plugin will automatically scan pages and get you the content of post.
5. Once you've done with your selection, click **"Jump to content"** button to set your post details. This button will show you the content of any post that has been selected from serial post list.
6. On right side of visual editor, there are post fields. You could define any field for your post on this area easily. Post title, content, categories, gallery images etc. can be defined on this field.
7. Click on **"Post Title"** field and click crosshair button beside path input field.
8. Click on post title from page on preview area. It will automatically get the xpath of post title. With this path, plugin will be able to scan all pages and get title with it's address.
9. Once you are done with selection, click on **"Preview Results"** menu and run your scraping model.
10. You'll see that plugin fetches all pages and all titles easily.

Managing Tasks

After creating your first task, you'll be able to manage tasks. On "Tasks" section of plugin, you could see created tasks by you. There are several functions for every task.

Here is the list of these functions :

1. **Start Task :** This function starts any task that has schedule.
2. **Trigger Now :** This function triggers task to run it now, it also shows the results.
3. **Update :** This button redirects you to task on visual editor for updating task.
4. **Clone :** This button clones task with all of it's properties.
5. **Reset :** This button resets task and it's hash to scan already scan pages again.
6. **Delete :** This button deletes task permanently.

There are also some informations that is useful to track post details.

1. **Task ID :** This is an unique ID for post that can be useful on support.
2. **Current Page :** If you task that contains pagination, you can track latest scraped page with this number.
3. **Interval :** This column shows information about post schedule.
4. **Last Process :** This column shows information last processed post details.
5. **Total Run :** This column shows completed task runs.

Single Scraping

Single scraping is scans only one page for scraping. It can update any post content or create post. Mostly, it's using to update a page for recent informations like price updates.

You could easily enable single scraping with clicking "Single Post" checkbox.

Serial Scraping

Serial scraping is scans all pages for scraping. It can update any post content or create post. Mostly, it's using to create new products or blog posts.

When you open visual editor, it's automatically on serial scraping mode.

RSS Scraping

Scraper automatically detects if content has RSS or XML and provides all available entries on a list view. You could use serial scraping method on this page easily.

Data Fields

There are several data fields for post structure.

Warning :
Plugin also supports other data fields that is available on your Wordpress web site.

You could enable data fields with creating empty posts with them. With this way plugin will be able to get them from database.

Here is the list of possible data fields :

1. **Post Title :** Post title and content can be defined easily, they contains text.
2. **Post Content :** Post title and content can be defined easily, they contains text.
3. **Featured Image :** For featured image, you should define an image URL or meta tag content.
4. **Post Date :** Post date will automatically parsed by our plugin for Wordpress's date format.
5. **Post Category :** Post category option can create new categories with source data.
6. **Tags :** Tags option can create new tags or define already created tags from source data.
7. **Gallery :** Gallery images will be scraped by plugin with serial image xpath. They all have been tied your post and you could use gallery with {{gallery}} shortcode on your post. For WooCommerce products, they will be automatically defined to product images.

WooCommerce and Product Setup

Plugin supports all tags on WooCommerce plugin. Here is the details for product setup :

1. Gallery images will be defined to your product automatically.
2. Price values should be marked as number on transform menu.
3. All available attributes can be enabled with creating empty product.
4. Product attributes can be used multiple times on data fields. You should mark "is multiple" to use product attributes.

Custom Fields

You can enable any custom field that you require on your web site with filling an empty post. They will appear on visual editor as a data field type. Also there are some fields that comes with plugin built-in.

Some fields requires taxonomy name or variable name. In those cases you could name input to define these attributes.

For example, product attributes of WooCommerce product requires a name for attribute, when you select product_attributes, plugin will automatically shows you the name field.

Transform Functions

This features provides some functions to transform source data. You could use math functions to change price value or use find and replace to remove or update some information on your source text. You could use transform functions to for any data type.

Here is the list of transform functions :

1. Find & Replace
2. Math Functions
3. Clean & Split
4. Clip
5. Translation
6. Attributes
7. Spin
8. Shortcodes

Using transform functions can be useful for interesting cases, here is some examples we would like to introduce you to power of transform functions.

1. Let's say, you have an image URL for 70x70 size information on it. You want to replace these values with 512x512 to get bigger thumbnail. Using find and replace function can solve this problem easily.
2. Let's say, you would like to scrape a data from a script element that contains video URL in it. You could use find and replace function to get it.
3. Let's say you want translate a title to another language, you could achieve it with translate feature easily.
4. Let's say you want add advertising to your content's end and start, you could achieve it with shortcodes feature easily.

There are more example for use cases of transform functions. You could also use transform functions as a pipeline to combine variables together.

Variables

Variables can be used with data fields. You could scrape any element with storing it's value in variable dynamically. They becomes available on shortcodes, so you could use any variable on title or content.

You could scrape a price tag as variable and use this value on title without any problem.

You could use variables for different cases like parsing any attribute from an element and using it's value on an embed tag.

Using Google Translate

Once you enabled Google Translation with your API key, you could translate any content that you scrape to your web page. Translation feature can be used partially. You could only translate title for example.

Translation feature can be enabled on **transform** menu.

Using Custom Xpath

If you have your own custom xpath, you could use it on data field easily. Plugin will automatically scan the page and use it as element selector address.

Find and Replace

Find and replace is a function of transform feature. You could replace multiple value on a data field. It supports regex functions for replacing.

Please be careful while you use special characters like dot, comma or other regex values that handled differently by regex replacer. They should be used with back slash. (For example \.)

Mathematical Functions and Price Calculations

Mathematical functions can be enabled easily on transform menu. You could also strip source content as a number. With this way, all values that has special character or not number will be cleaned.

On transform function, you could perform mathematical expressions. The "value" variable will be your source content. You could perform anything on that value.

For example : value + value * 0.5   expressions adds value itself with it's half of it.

Striping and Cleaning Content

With striping and cleaning content feature, you could remove any tag from source content, or strip HTML elements easily. It's also possible clean advertising scripts on this feature.

Shortcodes

Shortcodes feature provides a special field that you could write any variable that has been collected from source web site.

{{content}} is the source value that comes from transform's element. You could combine any value on that field easily. It also shows you the available tags.

Attribute Functions

Attribute functions provides some options to parse various attributes. If you would like to parse a background image from style attribute. You could do it easily on this menu with background image method.

Defining Post Tags

Post tags can be easily defined on "tag" field type. You need to specify a path that has serial expression on it.

If your tags are in a container, you could also use split feature to split content with delimiter.

Defining Post Categories

Post categories can be easily defined on "post_category" field type. You need to specify a path that has serial expression on it.

If your categories are in a container, you could also use split feature to split content with delimiter.

Updating a Post with Scraper

Once you are done with your task, you could set your post to update content easily. On project save menu, you could select any post from your recent task list. Plugin will automatically update the post content with data fields.

Scraper Visual Editor

Our support team can reach task's content from service logs if you provide us any. This method makes easier support process and with this way our support team assist you quickly on any case.

Scraper's plugin has functions that scrapes elements from given URLs. Our **visual editor is a cloud based editor that hosted by us** on our web service. Plugin doesn't contains visual editor's files in it. It only gets JSON data that built on visual editor.

Plugin Structure

Our plugin has clean and simple file structure. It only 3 different files in it's folder.

The files are :
- scraper.php
- plugin.php
- assets (folder)

Please download plugin only from Envato sources.

Licence and Third party libraries

Here is the list of used framework or libraries on plugin :

1. jQuery
2. Vue.js
3. PHP Proxy script (https://github.com/heiswayi/web-proxy-script)

How do I install plugin to my Wordpress web site?

Please follow the introduction below to install Wordpress plugin to your web site.

Method #1
- Extract the zip file that you've downloaded from Envato.
- Extract the plugin zip file to a temporary folder.
- Upload this folder to your Wordpress web site's **wp-content/plugins** folder path.
- Open web browser and go to Wordpress admin panel.
- Go to **Plugins / Installed** plugins page.
- You must see that **Scraper** title on plugins page.
- Click **"Activate"** button and that's it.

Method #2
- Extract the zip file that you've downloaded from Envato.
- Copy plugin zip file to a temporary folder.
- Open web browser and go to Wordpress admin panel.
- Go to **Plugins / Add New** page.
- Select zip file from temporary folder and click "Install Now" button.
- Click **"Activate"** button and that's it.

Are there any technical requirements to use this product?

Yes, like every Wordpress plugin, there are some technical requirements for Scraper as well.

Please make sure your hosting or server provides these technical requirements. Otherwise plugin cause some problems on your Wordpress site or it basically doesn't work properly.

- PHP 5 or higher PHP versions on your web site
- cURL or fopen support on your hosting
- Outgoing and incoming data transfer should be allowed on your server for API connections

If your server provider is not willing to allow outgoing and incoming network transfers, you could allow these IP addresses and URLs specifically.

- http://scraper.site
- https://scraper.site

Please note that our IP addresses might change in some cases. At this point, it's better to allow these domains only instead of IP addresses.

How do I contact support?

You could contact with our support any time. Please send email to support@imagets.com if you need support.

Please note that you could also write to us on Envato.

Refund policy for plugin

You could refund our plugin on Envato market in some cases. Please check out this page if you need to refund plugin :
https://themeforest.net/page/customer_refund_policy

Refund will be approved by our support team and your access token will be disabled on our systems.

Licences

This theme or plugin item is entirely licensed under the General Public License (GPL). You will find a copy of the license text in the same directory as this text file. Read more about licensing here:

http://codecanyon.net/licenses