

Admin Area » Manage Price Crawler


Price crawler is used to populate prices for merchant products.



















Each merchant registered can choose for his website to be spidered for price updation in his merchant area. The crawler spiders' merchant website in search of product and its price. It searches **for the Chunk Start and Chunk End to get the html between them. Then it uses the other Start and End parameters to get the Product Title, Manufacturer, Code and Price.** Then it updates the price of that product for that merchant in the price-comparison database.

Once updation is done, crawler then goes into the links on that page for their crawling. Therefore, page by page, whole website is spidered for that merchant.

Once a merchant's whole website is spidered, it moves ahead to spider another merchant's website.

Note: Crawling is resource hungry task and requires time, cpu and other resources of the server.

Click on **Merchants > Current Merchants** In the merchant listing click on the Spidering URLs icon  (see image below).

Page Stores                  

MerchantId	Name	Email	
<input type="checkbox"/>	lorusso	ecows	rlorussobulk@hotmail.com
<input type="checkbox"/>	kelkoo	kelkoo uk	ecowscomparison@live.com
<input type="checkbox"/>	cmania	roberto	ecowscomparison@yahoo.co.uk

Page 1 of 1
Total Record(s) : 3

This will display a listing of urls for the merchant. You can use the **Add icon on the toolbar** to add a new url, or you can edit a previous url.

Select Merchant: Kaon Softwares

ID	URL
<input type="checkbox"/> 1000	http://audiovisual.kelkoo.co.uk/c-100278823-gps.html

Page 1 of 1

1. URL

It is the url from where the crawler will start crawling for the prices of this merchant.

2. Spider URLs Containing

It should contain a phrase which should be present in the subsequent urls to be spidered. Meaning, when you will go to the above url, you will notice there are links to the other categories and products on that page as well as links to about us and privacy policy. You dont want your crawler to waste time in unnecessary pages so you instruct it to spider only those urls which contains that particular phrase.

This will ensure that crawler will only go into pages where price info will be present. Therefore saving time and resources. If left blank it will go into all links like about us and others.

3. Process URLs Containing

It should contain a phrase which is present in the urls of the product price page. So that only those pages are processed for price extraction.

4. Ignore URLs Containing

You can provide a phrase which if present in the urls of the product price page, they will be ignored and not processed.

5. Chunk Start & Chunk End

This is the starting html and ending html of the chunk you want to process. Once this chunk is found, it will be processed with other start and end parameters to get the fields of a product.

6. Product Title Start & End

These two fields will be used to find product title from the chunk, the data between these two fields will supposed to be the product title.

7. Manufacturer Start & End

These two fields will be used to find product manufacturer from the chunk.

8. Manufacturer Code Start & End

These two fields will be used to find product code from the chunk.

9. Comment Start & End

These two fields will be used to find merchant comments on the product from the chunk.

10. Shipping Start & End

These two fields will be used to find product shipping from the chunk.

11. Price Start & End

These two fields will be used to find product price from the chunk. This price is exclusive of the currency symbol.

12. Currency Symbol

You can specify the currency symbol here. You can write plain text, html or Unicode.

13. Buy Url Affiliate String

You can provide a string to attach to the buy url. This feature can be used to send an affiliate tracking id to the product buy link.

14. Can/Cannot add new product

Use this check to specify what should the crawler do if the product price is not found in our database, whether to add the price for that merchant or not.

Add Price Crawler URLs for 'salman' >>>

Merchant ID : *

URL : *
(Tip: Start of spidering)

Spider URLs containing : *
Comma (,) separated
(Tip: Only go into links which contain above text)

Process URLs containing : *
(Tip: Only process a url for product which contains above text)

Ignore URLs containing : *
Comma (,) separated
(Tip: Will ignore a url if it contains any of above text)

Chunk Start : *

Chunk End : *

Product Title Start :

Occurance count: (type 1 if not sure)

Product Title End :

Manufacturer Start :

Occurance count: (type 1 if not sure)

Manufacturer End :

Manufacturer Code Start :

Occurance count: (type 1 if not sure)

Manufacturer Code End :

Comment Start :

Comment End :

Shipping Start :

Shipping End :

Price Start :

Occurance count: (type 1 if not sure)

Price End :

Currency Symbol : * Unicode/HTML tag/plain text

Buy URL Affiliate String :

It should be carefully entered so that it makes a valid url.
It can be ?**affid=XEseeEGF** or &**affid=XEseeEGF**

Can add new prices to products (if not found in database).
 Cannot add new prices to products (if not found in database).
Only update current products.

Active :



Now click on the Run Price Spider to start spidering.
You can view log of the actions perform once task is finished by the spider.

The screenshot shows the Comparison V6 administrator interface. At the top, the logo 'Comparison V6' is on the left, and 'Administrator: admin' is in the center. On the right, there are 'Home' and 'Chan' links. Below the header is a navigation bar with tabs for 'Parameters', 'Content', 'Merchants', 'Members', 'Affiliates', and 'Reports'. The 'Merchants' tab is active, and a dropdown menu is open over it, listing options: 'Current Merchants', 'Import Datafeeds', 'Export Datafeed', 'Manage Autofeeds', 'Register a Merchant', 'Run Price Crawler' (highlighted in blue), and 'Search Merchants'. To the right of the menu, there are icons for 'Refresh', 'Search', 'Add', and 'Edit'. Below the menu, a table is partially visible with a header row containing 'URL' and a data row with the URL 'http://computing.kelkoo.co.uk/c-1100'. At the bottom right, there is a pagination indicator showing '1 of 1' and 'Record(s) : 1'.