

Function details : esta función almacena todos los datos referentes a los ecommerce que en rubik se registran lo hace mediante la captura de tags html , Javascript o objetos json los cuales son almacenados por los imports mediante expresiones regulares que este contiene los cuales son:

```
// URL
$import->url($url);

// Código
$import->code(variable);

// Título
$import->title(variable);

// Descripción
$import->description(variable);

// Precio
$import->price(variable);

// Stock
$import->stock(variable);

// Categoría
$import->categorytags_insert(variable);

// Marca
$import->add_brand(variable);

// Ean, ref, gtin, sku, upc
$import->add_code(variable);

// Shipping
$import->shipping(variable);

// Año
$import->add_year(variable);

// Formato
$import->add_format(variable);
```

### **Expresiones regulares**

Para capturar la información contenida entre los tags o variables usamos la expresión (.\*?)  
Que se traduce a captura todo lo que se encuentre dentro de.

Ejemplo expresión Tag HTML

```
'<h1>(.*?)</h1>'
```

Ejemplo expresión variable Javascript

```
'var price =(.*?);'
```

Para omitir contenido que usamos la expresión `.*` Que se traduce a omitir hasta, la usamos para escalar contenido de un Tag a otro como por ejemplo

Tenemos este contenido

```
<div>abcd</div>
<span>fgrtdfd</span>
<h1>título</h1>
```

Y deseamos capturar el contenido del h1 partiendo desde el div usamos la siguiente expresión que se traduce a captura todo lo que está en el h1 anclando desde el div omitiendo todo lo que hay entre el div y el h1

```
'<div>.*?<h1>(.*?)</h1>'
```

Generalmente se usan como anclas los id y las clases cuando el Tag o la variable dónde se encuentra la información que deseamos capturar se encuentra más de una vez en el HTML de las webs.

## Funciones de sistema rubik

Función `nr_match`: está función recibe tres parámetros los cuales son la expresión regular que usaremos para capturar la información, la variable caché y la variable dónde almacenaremos la información capturada con la expresión regular , está función devuelve un string.

Ejemplo:

Tenemos el Tag `<h1>hola</h1>`

```
nr_match('<h1>(.*?)</h1>', $cache, $variable);
```

En este caso `$variable` almacena hola

Función `nr_match_all`: está función recibe tres parámetros los cuales son la expresión regular que usaremos para capturar la información, la variable caché y la variable dónde almacenaremos la información capturada con la expresión regular , está función devuelve un array.

Ejemplo:

Tenemos tres veces en el HTML el Tag h1 `<h1>hola</h1>`

```
<h1>como</h1>
```

```
<h1>estas</h1>
```

```
nr_match_all('<h1>(.*?)</h1>', $cache, $variable);
```

En este caso `$variable` almacena

```
array {
```

```
"hola",  
"como",  
"Estas"  
}
```

Función `nr_get_content`: esta función hace la misma función de `file_get_content` del php nativo el cual es extraer el contenido un un archivo , este contenido puede ser un string como el HTML de una URL , un objeto json generado por una petición get o post , recibe un parámetro que es la URL a la cual le vamos a extraer la información, junto a esto podemos enviarle un arreglo con cookies, cabeceras, encoding entre otros. Se debe almacenar en una variable.

Ejemplo

```
$content = nr_get_content($url);
```

Función `nr_get_cookies_from_headers`: esta función tiene como propósito extraer las cookies generadas por las páginas webs, recibe un parámetro que es la URL de la web a la cual le vamos a extraer las cookies , un array con el parámetro `'header'=>true` y `'usecache'=>false` , devuelve un array con las cookies generadas, se usa frecuentemente en aquellas webs que tengan que hacer cambios de moneda para distintos países. Se debe almacenar en una variable.

Ejemplo

```
$cookies = nr_get_cookies_from_headers($url,array('header'=>true , 'usecache'=>false));
```

Función `dump`: hace la misma función de `var_dump` del php nativo la usamos frecuentemente para ver qué nos devuelven las variables en las cuales almacenamos la información.

Ejemplo

```
dump($variable);
```

Función `nr_remove_params`: esta función nos sirve para remover parámetros de las URL en caso de necesitar hacerlo, recibe un array con los parámetros que deseamos remover y la URL.

Ejemplo

```
$array = array('param1','param2');  
$url = nr_remove_params($url, $array);
```

Función `offers`: esta función sirve para capturar las ofertas en el caso de los ecommerce que son del tipo marketplace como Amazon, Ebay, entre otras. La usamos si y solo si el cliente nos dice que debemos extraer esta información.

Función `nr_utf8_decode`: esta función

Parsea los caracteres a entidades como por ejemplo es frecuente encontrar en las URL de países como alemania que contienen letras fuera de lo común como diéresis y éstas en la URL vienen codificadas como `%D22` y esta función la parsea a `ç` .

Ejemplo

```
$url = nr_utf8_decode($url);
```

Función session\_refresh: nos sirve para extraer cookies de sesión las cuáles tienen cierto tiempo de expiración y está función refresca las cookies cada 5 minutos, se usa con la función nr\_get\_cookies\_from\_headers, recibe como parámetro la URL que deseamos extraer las cookies y devuelve un array con las cookies.

### **Capturando peticiones get y post**

Para capturar la información de una petición get , abrimos el inspector de elementos luego vamos a la sección xhr , hacemos la petición en la página web que estamos trabajando y en esta se mostrará la URL a la cual hace la petición junto con los parámetros que nos solicita, una vez tengamos está información armamos la URL y se la pasamos a la función nr\_get\_content y está nos devolverá el contenido.

Para capturar la información de una petición post abrimos el inspector de elementos luego vamos a la sección xhr , hacemos la petición en la página web que estamos trabajando y en esta se mostrará la URL a la cual hace la petición junto con los parámetros que nos solicita, una vez tengamos está información armamos la URL y se la pasamos a la función nr\_get\_content junto con un array que contendrá los parámetro post que le envía al servidor y está nos devolverá el contenido, ejemplo:

Petición get

```
$urlGet = 'https://www.laulralaqueenvia.php?parametro1=1234&parametro2=test';
```

Se envía de la siguiente manera

```
$peticion = nr_get_conten($urlGet);
```

Petición post

\$urlPost

```
'https://www.laulralaqueenvia.php';
```

Parámetros post

```
$post = 'PARAMETRO1=1234&PARAMETRO2=test'
```

Se envía de la siguiente manera

```
$peticion = nr_get_conten($urlPost, array('post'=>$post, 'usecache'=>false));
```

### **Cabeceras o custom headers**

Estás cabeceras son requeridas por algunas peticiones para devolver su contenido, ya que contienen información enviadas a los servidores que son de uso exclusivo para las mismas tales como autorizaciones, token o sesiones , especifican objetos del tipo json entre otras.

Para localizar estás cabeceras custom headers vamos al inspector de elementos luego a la sección all o xhr , nos colocamos sobre la URL que deseamos obtener las cabeceras hacemos click derecho seleccionamos la opción copy as Bash y está nos copia en el

portapapeles la información, luego almacenamos en arreglos las que comienzan con H - y las enviamos en un arreglo de la siguiente forma

```
$url = nr_get_content($_url,array('custom_headers'=>$custom_headers));
```

## **Funciones y ciclos usados frecuentemente**

### **Funciones**

json\_decode, parse\_url, str\_replace, substr\_count, count, str\_pos, trim, is\_numeric, array\_search, array\_pop, array\_shift, unset, time, switch.

### **Ciclos iterativos**

foreach, while, for.

### **Que es la store\_info?**

La store\_info es un arreglo que contiene toda la información acerca de la store que estamos trabajando en determinado momento, contiene información como currency , countryCode y lenguaje

Que son los más usados frecuentemente.

Se invocan mediante el nombre de la posición del arreglo y se almacenan en variables.

Ejemplo

```
$currency = $store_info['Currency'];
```

Currency hace referencia al tipo de moneda que se usan en determinados países, como por ejemplo euro, dolar, libras entre otras.

countryCode es el código de país como por ejemplo para Alemania sería de , España es, Francia fr entre otros.

DefaultLanguage es el idioma que se habla por defecto en cada país, para reino unido sería 'en' España 'es' ya así sucesivamente.

Existen otras posiciones en el arreglo store\_info , para ver cuáles son se debe hacer  
dump(\$store\_info);

Sitemaps y crawler.

La ruta crawler de una store es la URL que nos muestra al hacer click en Home de La store casi siempre es la URL limpia por defecto

Ejemplo:

<https://www.lawebdelpinguino.com>

La ruta sitemap de una store es el mapa del sitio que contiene todas las URL archivadas en un archivo xml y se consigue de la siguiente forma, escribiendo sitemaps.xml al final de la URL.

Ejemplo

<https://www.lawebdelpinguino.com/sitemap.xml>