

Forum:[Hack Forums](#)

Thread: [PHP curl](#)

redmoon90211 - January 15, 2020, 11:07 am

I am trying to scrape item numbers (eg. 33FR410) from this page

<https://www.tavor-trade.com/%D7%A7%D7%98%D7%92%D7%95%D7%A8%D7%99%D7%95%D7%AA/%D7%97%D7%9C%D7%A4%D7%99%D7%9D-%D7%9C%D7%9E%D7%95%D7%A6%D7%A8%D7%99-%D7%97%D7%A9%D7%9E%D7%9C/%D7%9E%D7%96%D7%92%D7%A0%D7%99%D7%9D/aa-%D7%91%D7%9C%D7%95%D7%A0%D7%99-%D7%92%D7%96-%D7%9C%D7%A7%D7%99%D7%A8%D7%95%D7%A8/>

But I can't find them anywhere in the source HTML or HAR file
Completely puzzled ?

Any help is appreciated

Nachfolger - January 15, 2020, 11:17 am

(Jan 15, 2020 12:07 PM)redmoon90211 Wrote: [\[CITE\]](#) I am trying to scrape item numbers (eg. 33FR410) from this page

<https://www.tavor-trade.com/%D7%A7%D7%98%D7%92%D7%95%D7%A8%D7%99%D7%95%D7%AA/%D7%97%D7%9C%D7%A4%D7%99%D7%9D-%D7%9C%D7%9E%D7%95%D7%A6%D7%A8%D7%99-%D7%97%D7%A9%D7%9E%D7%9C/%D7%9E%D7%96%D7%92%D7%A0%D7%99%D7%9D/aa-%D7%91%D7%9C%D7%95%D7%A0%D7%99-%D7%92%D7%96-%D7%9C%D7%A7%D7%99%D7%A8%D7%95%D7%A8/>

But I can't find them anywhere in the source HTML or HAR file
Completely puzzled ?

Any help is appreciated
You can't find them because they're loaded by a XHR request, which is not sent if you view the source or fetch it via cURL.

redmoon90211 - January 15, 2020, 11:19 am

(Jan 15, 2020 12:17 PM)Nachfolger Wrote: [\[CITE\]](#) You can't find them because they're loaded by a XHR request, which is not sent if you view the source or fetch it via cURL.

I can't find them in the HAR file saved from Chrome browser either

Nachfolger - January 15, 2020, 11:32 am

(Jan 15, 2020 12:19 PM)redmoon90211 Wrote: [\[CITE\]](#) I can't find them in the HAR file saved from Chrome browser either
They are all there, it's the XHR request made to "https://shop4.wizsoft.com/VSHOP/WSHOP.wzx". Block the request URI and the items will no longer load.

redmoon90211 - January 15, 2020, 11:57 am

(Jan 15, 2020 12:32 PM)Nachfolger Wrote: [\[CITE\]](#) They are all there, it's the XHR request made to "https://shop4.wizsoft.com/VSHOP/WSHOP.wzx". Block the request URI and the items will no longer load.

Thanks ! Still can't find them tho, may be encrypted or in base64 format, I will look into that

Charger - January 15, 2020, 1:46 pm

(Jan 15, 2020 12:57 PM)redmoon90211 Wrote: [\[CITE\]](#) Thanks ! Still can't find them tho, may be encrypted or in base64 format, I will look into that
See if there's any JavaScript that decodes it.

AFAIK you can see every JS call made in chrome profiles

Nachfolger - January 15, 2020, 2:20 pm

(Jan 15, 2020 12:57 PM)redmoon90211 Wrote: [\[CITE\]](#) Thanks ! Still can't find them tho, may be encrypted or in base64 format, I will look into thatBased on what [Charger](#) just said, I decided to dig further. Here you go:

Code eval("Tmp = " + unescape(response.responseText));It produces a JSON object.

Aka:

```
Code eval("Tmp = " +
unescape('%u007b%u0020%u0022%u0053%u0074%u0061%u0074%u0075%u0073%u0022%u00
3a%u0022%u004f%u004b%u0022%u002c%u0022%u004e%u006f%u0074%u0065%u0073%u005
4%u0061%u0062%u0022%u003a%u005b%u005b%u0022%u0031%u0022%u002c%u0022%u05d
e%u05d5%u05ea%u05d2%u0022%u002c%u0022%u0030%u0022%u005d%u005d%u002c%u002
2%u0043%u006f%u006c%u0075%u006d%u006e%u0073%u0022%u003a%u005b%u0022%u004e
%u006f%u0074%u0065%u0049%u0044%u0022%u002c%u0022%u004e%u006f%u0074%u0065%
u004e%u0061%u006d%u0065%u0022%u002c%u0022%u0049%u0073%u004e%u0075%u006d%
u0065%u0072%u0069%u0063%u0022%u005d%u0020%u007d`'))
```

Results in:

Quote:{"Status":"OK","NotesTab":[["1","????","0"]],"Columns":["NoteID","NoteName","IsNumeric"]}

Charger - January 15, 2020, 2:43 pm

(Jan 15, 2020 03:20 PM)Nachfolger Wrote: [\[CITE\]](#) Based on what [Charger](#) just said, I decided to dig further. Here you go:

Code eval("Tmp = " + unescape(response.responseText));It produces a JSON object.

Aka:

```
Code eval("Tmp = " +
unescape('%u007b%u0020%u0022%u0053%u0074%u0061%u0074%u0075%u0073%u0022%u00
3a%u0022%u004f%u004b%u0022%u002c%u0022%u004e%u006f%u0074%u0065%u0073%u005
4%u0061%u0062%u0022%u003a%u005b%u005b%u0022%u0031%u0022%u002c%u0022%u05d
e%u05d5%u05ea%u05d2%u0022%u002c%u0022%u0030%u0022%u005d%u005d%u002c%u002
2%u0043%u006f%u006c%u0075%u006d%u006e%u0073%u0022%u003a%u005b%u0022%u004e
%u006f%u0074%u0065%u0049%u0044%u0022%u002c%u0022%u004e%u006f%u0074%u0065%
u004e%u0061%u006d%u0065%u0022%u002c%u0022%u0049%u0073%u004e%u0075%u006d%
u0065%u0072%u0069%u0063%u0022%u005d%u0020%u007d`'))
```

Results in:

I'm a genius I know

Nachfolger - January 15, 2020, 2:46 pm

(Jan 15, 2020 03:43 PM)Charger Wrote: [\[CITE\]](#) I'm a genius I knowCompared to a aboriginal? Yes.
Compared to me? No.

redmoon90211 - January 15, 2020, 4:42 pm

Thank you all ! Especially [Nachfolger](#), you rock !

newQuery - January 16, 2020, 6:24 am

When it comes to scrapping, I'd always use python.

See: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

Charger - January 16, 2020, 10:42 am

(Jan 16, 2020 07:24 AM)newQuery Wrote: [\[CITE\]](#) When it comes to scrapping, I'd always use python.

See: <https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

simple_html_dom is available for PHP which is almost the same as BS.