

Forum:[Hack Forums](#)

Thread: [Trying to make a script to scan a webpage](#)

Slim - October 26, 2015, 10:46 am

I have no idea where to start as i'm new to python. (I know the basics though.)

Here is some psuedo-code to help understand what I want to do.

```
Go to [webpage]
Refresh [webpage] every 10 seconds
if a certain number value is found, email me (repeatedly) [meaning for each unique value it finds,
email me]
else, keep refreshing
```

SIM-1866423 - October 26, 2015, 10:57 am

psuedo-code

PHP Code

```
<?php
import urllib
import time
import re
import smtplib

def gmail( info):
    /*SEND INFO*/
```

```
while True:
    data = urllib.open("URL").read()
    unique = re.findall("REGEXFORVALUE", data)
    for x in unique:
        gmail( x)
    time.sleep(10)
```

Slim - October 26, 2015, 3:08 pm

(Oct 26, 2015 10:57 AM)b3mb4m Wrote: [\[CITE\]](#) psuedo-code

PHP Code

```
<?php
import urllib
import time
import re
import smtplib

def gmail( info):
    /*SEND INFO*/
```

```
while True:
    data = urllib.open("URL").read()
    unique = re.findall("REGEXFORVALUE", data)
    for x in unique:
        gmail( x)
    time.sleep(10)
```

Thank you so much! Question, under "SEND INFO" I don't exactly know what to put. Say my email was helloworld@gmail.com, how would I go about it?

SIM-1866423 - October 26, 2015, 4:34 pm

check smtplib lib on google, you will understand

h0lyl337 - October 26, 2015, 11:48 pm

Hey maybe my example code here will point you in the right direction.

PHP Code

```
<?php
import urllib2
import re
```

```
def getreddit():
    url = "REDDITDOTCOMHERE/"
    request = urllib2.Request(url)
    handle = urllib2.urlopen(request)
    readhandle = handle.read()
    readhandlestr = str(readhandle)
```

```
# Look for any characters in between " <p class="title"><a class="title may-blank ""
# and " a>' " , that match the regular expression pattern, which is " .+? ".
```

```
parse1 = re.findall(r'<p class="title"><a class="title may-blank "(.+?)a>', readhandlestr)
parse1string = str(parse1)
```

```
parse2 = re.findall(r'\s>(.+?)</', parse1string)
for text in parse2:
    print text
```

```
getreddit()
```