

Re: Automating search for words in a website using WSH

Source:

<http://www.tech-archive.net/Archive/Scripting/microsoft.public.scripting.wsh/2009-06/msg00010.html>

- *From:* "Hii Sing Chung" <singchung@xxxxxxxxxxx>
 - *Date:* Sun, 7 Jun 2009 08:28:24 +0800
-

Tim,

1. <http://www.rbl.jp/phishing/> At this website, we will look under today's posting at the top portion (June 6, 2009) for the presence of the organization's names in any forms. If we see any match, we will count the number of occurrences.

2. <http://www.fraudwatchinternational.com/phishing/search.php> At this page we will select today's date, then select the organization name from the drop down 'target company', then click search. The result will be displayed as "Result Found: [number of occurrence]".

The number of occurrence we will added up from these 2 sites and key in the e-mail as number of increase in phishing detected.

Let me try to get my team member to furnish an example of the e-mail.

Thanks

Sing Chung

"Tim Harig" <usernet@xxxxxxxxxxx> wrote in message
[news:xxvWl.17770\\$pr6.577@xxxxxxxxxxxxxxxxxxxxxxxxxxxx](news:xxvWl.17770$pr6.577@xxxxxxxxxxxxxxxxxxxxxxxxxxxx)

On 2009-06-06, Hii Sing Chung <singchung@xxxxxxxxxxx> wrote:

"Tim Harig" <usernet@xxxxxxxxxxx> wrote in message
[news:GZrWl.20022\\$hc1.18680@xxxxxxxxxxxxxxxxxxxxxxxxxxxx](news:GZrWl.20022$hc1.18680@xxxxxxxxxxxxxxxxxxxxxxxxxxxx)

On 2009-06-06, Hii Sing Chung
<singchung@xxxxxxxxxxx> wrote:
If I understand you properly, this is the part where you are
having your
difficulty...Right? How are is your script going to the URL
-- or are you
not able to get to the URL at all right now?

The URL is fixed (the one I give here is not the actual website, as I can't
give actual address due to confidentiality), the team members have to check

[SNIP]

Re: Automating search for words in a website using WSH

unicodeword5, are just examples I give because I cannot give the actual words due to confidentiality). I also have to check matching for Chinese and

Thats fine; but, without specifics, I can only give you you general answers.

Loading a URL from WSH is easy using Internet Explorer as a COM server.

You can load an Internet Explorer object as:

```
Wscript.CreateObject("InternetExplorer.Application")
```

Can I then pass this count value back to WSH?

Everything is already being done inside of WSH. You are creating the COM object from WScript (correct spelling, the 'S' should be capitalized above). Once the COM object is created it acts just like a native object for whatever language you are using (Javascript, VBscript, Python, etc.)

You can pull the information as I have shown above. The only thing left is to scrape the information you want from the page. You have been too vague for me to help you in parsing out your informaton so you will have to figure out how to do that yourself.

Vague?? Which is the part not clear? The e-mail is to be sent to Emergency Center to notify if there is any increase in phishing messages detected that use the names of this organization. I strongly feel that it is not wise to manually check against website everyday using human eyes.

It is not that it isn't clear, its that I don't know what the data that you are trying to pull looks like. I have given you enough information to download the web page; but, without being able to see a markup example and that data that you are trying to extract from it, I cannot give you any ideas as to extract that data. Unless the data is preformatted into some kind of delineated data, XML, etc with stateful metadata; you are probably going to have to parse the HTML. HTML scraping is very specific to the page that you are trying to scrape. Without seeing a representative page that you are trying to extract the data from, I cannot tell you how to extract it from that page.

I gave you information on how to download that page. Without being able to see the data, it will then be your responsibility to figure out how to parse it.

Re: Automating search for words in a website using WSH