

Re: view page source or save after load

Source: <http://coding.derkeiler.com/Archive/Python/comp.lang.python/2006-09/msg03018.html>

- *From:* "alex23" <wuwei23@xxxxxxxxxx>
 - *Date:* 20 Sep 2006 22:26:39 -0700
-

zephron2000 wrote:

I need to either:

1. View the page source of a webpage after it loads
- or
2. Save the webpage to my computer after it loads (same as File > Save Page As)
- urllib is not sufficient (using urlopen or something else in urllib isn't going to do the trick)

You don't really say `_why_ urllib.urlopen "isn't going to do the trick"`. The following does what you've described:

```
import urllib
page = urllib.urlopen('http://some.address')
open('saved_page.txt','w').write(page).close()
```

If you're needing to use a browser directly and you're running under Windows, try the Internet Explorer Controller library, IEC:

```
import IEC
ie = IEC.IEController()
ie.Navigate('http://some.address')
page = ie.GetDocumentHTML()
open('saved_page.txt','w').write(page.encode('iso-8859-1')).close()
```

(You can grab IEC from <http://www.mayukhbose.com/python/IEC/index.php>)

Hope this helps.

-alex23